

4. Theoretische Konstrukte und Indikatoren

4.1 Konzeptspezifikation

Die in einer Untersuchung verwendeten Begriffe (theoretischen Konstrukte) müssen zunächst definiert werden. Wie in 2.1.4.1 ausführlich begründet wurde, ist eine solche Definition als Nominaldefinition zu verstehen - sie kann im Gegensatz zur Realdefinition nicht wahr oder falsch sein (sie trifft also nicht ein wie auch immer verstandenes „Wesen“ des Gegenstandes), sondern nur mehr oder weniger brauchbar. In empirischen Untersuchungen verwendete Begriffe sollten daher präzise definiert sein, konsistent verwendet werden und theoretisch fruchtbar sein. Entscheidend für die Brauchbarkeit eines Begriffes ist die theoretische Fruchtbarkeit, d.h. mit dem Begriff müssen sich Hypothesen bilden lassen, die einen hohen Informationsgehalt haben und sich empirisch bewähren (vgl. die Kriterien für eine „gute“ Theorie in 2.1.4.1). Um zu theoretisch fruchtbaren Begriffen zu kommen, ist eine sorgfältige Konzeptspezifikation erforderlich. Dazu bedient man sich des Verfahrens der Begriffsexplikation, also der systematischen sprachlichen Präzision von Begriffen. Zunächst wird die Verwendung des Begriffes (z.B. Entfremdung) in den verschiedensten Theoriekontexten untersucht (Literaturstudium: Was verstehen z.B. Hegel, Marx, Weber, Seeman, Luhmann, usw. unter Entfremdung). Auf diese Weise werden unterschiedliche Vorstellungsinhalte zusammengestellt und präzisiert (dimensionale Analyse). In Hinblick auf die Forschungsziele wird dann entschieden, welche Vorstellungsinhalte in die Definition des Begriffes aufgenommen werden sollen (Kriterium der theoretischen Fruchtbarkeit). Im Gegensatz zur Begriffsinterpretation ist es also nicht entscheidend, ob die begriffliche Präzisierung mit der Auffassung eines bestimmten Autors übereinstimmt (z.B. mit der Auffassung von Entfremdung bei Marx). Gefordert ist eine rekonstruktiv-kritische Vorgehensweise, die den zu definierenden Begriff vor dem Hintergrund der vermuteten theoretischen Zusammenhänge (z.B. einer zu entwickelnden Theorie der Entfremdung) präzisiert. Das Ergebnis der Konzeptspezifikation ist also eine Nominaldefinition des zugrundeliegenden theoretischen Konstrukts, die geeignet ist, dieses optimal in eine zu entwickelnde Theorie einzupassen.

4.2 Operationalisierung

Für die empirische Forschung genügt es nicht, daß die verwendeten Begriffe explizit definiert sind. Es müssen vielmehr Forschungsoperationen angegeben werden, mit deren Hilfe entscheidbar ist, ob in einer konkreten Situation ein mit dem betreffenden Begriff bezeichnetes Phänomen vorliegt oder nicht (ob z. B. eine bestimmte Person Anzeichen von Entfremdung

zeigt oder nicht). Eine solche Festlegung von Forschungsoperationen nennt man die Operationalisierung eines Begriffs.

Insbesondere für Begriffe ohne direkten empirischen Bezug ist die Operationalisierung ein zweistufiger Prozeß. In diesem Zusammenhang knüpfen wir wieder an die Diskussion der Zweisprachentheorie in 2.1.4.2 an. Für Begriffe der theoretischen Sprache (Begriffe ohne direkten empirischen Bezug, sog. theoretische Konstrukte) muß zunächst festgelegt werden, welche Begriffe der Beobachtungssprache (Begriffe mit direktem empirischen Bezug) als Indikatoren gelten sollen. Dem entspricht die Angabe von Korrespondenzregeln oder die Spezifikation einer Meßtheorie im Sinne des kausalanalytischen Ansatzes (siehe unten).

Wir bezeichnen den Vorgang der Indikatorenauswahl als Operationalisierung im weiteren Sinne. Da die Menge aller möglichen Indikatoren nie genau spezifiziert werden kann, ist die Auswahl von Indikatoren immer selektiv und notwendigerweise unvollständig. Das zentrale Problem besteht also darin, wie die Zuordnung eines Indikators zu einem theoretischen Begriff gerechtfertigt werden kann. Man unterscheidet folgende Lösungen des Korrespondenzproblems (vgl. auch Schnell/Hill/Esser: Methoden der empirischen Sozialforschung, Kap. 4.2):

- Operationalismus
- typologisch-induktiver Ansatz (wird hier nicht weiter behandelt)
- kausalanalytischer Ansatz

Die operationalistische Lösung besteht in der definitiven Gleichsetzung von theoretischem Konstrukt und Indikator(en). Dies bedeutet, daß man z.B. unter Intelligenz genau das versteht, was die in einem Intelligenztest enthaltenen Aufgaben (Indikatoren) messen. Damit wird die Zweisprachentheorie faktisch wieder aufgegeben. Es gibt im Grunde gar kein unabhängig definiertes theoretisches Konstrukt „Intelligenz“, sondern genau so viele empiristische Begriffe der Intelligenz, wie es unterschiedliche Intelligenztests gibt. Damit wird es unmöglich, eine einheitliche Theorie der Intelligenz aufzustellen. Vielmehr gibt es genau so viele, im Sinne von Kuhn untereinander inkommensurable, Aussagen über Intelligenz, wie es verschiedene Untersuchungen über Intelligenz mit unterschiedlichen Intelligenztests gibt. Aus diesem Grunde ist die operationalistische Lösung des Korrespondenzproblems eine Scheinlösung, die den Anforderungen der Sozialwissenschaften nicht gerecht wird.

Der kausalanalytische Ansatz hält an der Zweisprachentheorie fest und betrachtet die Indikatoren als beobachtbare Folgen der latenten Variablen (Wirkungsindikatoren). [Daneben besteht auch die Möglichkeit, Indikatoren als Ursachen von latenten Konstrukten zu konzipieren (Ursachenindikatoren). In einem erweiterten Verständnis des kausalanalytischen Ansatzes sind sowohl Wirkungsindikatoren als auch Ursachenindikatoren zugelassen. Wir beschränken uns im folgenden aber auf die Diskussion von Wirkungsindikatoren.] Es wird also eine explizite Meßtheorie spezifiziert, die die Beziehung zwischen theoretischen Konstrukten und Indikatoren als Kausalbeziehung konkretisiert. Vorteile des kausalanalytischen Ansatzes sind

das Festhalten an der Zweisprachentheorie und die explizite Formulierung einer Meßtheorie (und auch einer Meßfehlertheorie) zusätzlich zur Kerntheorie. Komplementär dazu ergibt sich als (unvermeidlicher) Nachteil (die Welt ist nun einmal kompliziert) ein komplexes mathematisches Modell, zu dessen Lösbarkeit oft stark vereinfachende Annahmen gemacht werden müssen (für eine Formalisierung und genauere Diskussion des kausalanalytischen Ansatzes siehe 4.3).

Auch im kausalanalytischen Ansatz stellt die Auswahl von validen Indikatoren in erster Linie ein theoretisches Problem dar (Meßtheorie). Indikatoren sind genau dann eine valide Operationalisierung eines theoretischen Konstruktes, wenn eine kausale Beziehung zwischen ihnen und dem zugehörigen theoretischen Konstrukt theoretisch einsichtig ist und empirisch bestätigt werden kann. (Nach den Überlegungen in 2.1.4.2 ist klar, daß eine solche empirische Überprüfung immer nur im Verbund von Kern-, Meß- und Meßfehlertheorie stattfinden kann.)

Zur Konkretisierung der Problematik der Indikatorenauswahl gehen wir kurz auf die Messung des Postmaterialismus ein. Die Postmaterialismus-These von R. Inglehart besagt, daß in den westlichen Industriegesellschaften seit dem 2. Weltkrieg eine Entwicklung von materialistischen hin zu postmaterialistischen Einstellungen zu beobachten ist. Zur empirischen Überprüfung dieser These ist die Messung postmaterialistischer Einstellungen erforderlich. Dazu werden die folgenden vier Items (Auswahl von Indikatoren aus dem potentiell unendlichen Indikatorenuniversum) der Postmaterialismusskala verwendet:

- A: Aufrechterhaltung von Ruhe und Ordnung
- B: Mehr Einfluß der Bürger auf Entscheidungen der Regierung
- C: Kampf gegen steigende Preise
- D: Schutz des Rechtes auf freie Meinungsäußerung

Die Indikatoren der Postmaterialismusskala stellen also politische Ziele dar, die von den Befragten entsprechend ihrer Wichtigkeit in eine Rangreihe gebracht werden sollen. Dabei sind die Items A und C Indikatoren für eine materialistische Einstellung, während die Items B und D eine postmaterialistische Einstellung anzeigen sollen. Je nachdem, ob der Befragte eher die postmaterialistischen oder die materialistischen Ziele für wichtiger hält, wird ihm ein entsprechender Wert auf der Postmaterialismusskala zugeordnet (eine genauere Diskussion des zugrundeliegenden Meßmodells erfolgt in 5.1).

In dieser Form wird die Postmaterialismusskala üblicherweise eingesetzt. Es existiert aber auch eine andere Version der Postmaterialismusskala, die andere Items verwendet, nämlich:

- E: Erhaltung eines hohen Grades von wirtschaftlichem Wachstum
- F: Sicherung von starken Verteidigungsstreitkräften
- G: Verstärktes Mitspracherecht der Menschen an ihrem Arbeitsplatz und in ihrer Gemeinde

H: Versuche, unsere Städte und ländlichen Gebiete zu verschönern

In dieser zweiten Version sind E und F die materialistischen und G und H die postmaterialistischen Items.

Diese unterschiedlichen Operationalisierungen (im weiteren Sinne) des Postmaterialismus führen zu deutlich verschiedenen empirischen Aussagen (vgl. Hagstotz 1985). (Die folgenden Ergebnisse beziehen sich auf Daten aus dem ALLBUS 1982.) So sind nach Version I 70% der Bundesbürger materialistisch orientiert, nach Version II aber nur 58%. Die Korrelation zwischen den beiden Skalen (hier wird τ_b als Assoziationsmaß verwendet; siehe auch 5.2) beträgt nur .27, d.h. die beiden Versionen messen weitgehend unterschiedliche Sachverhalte. Um zwischen den beiden Versionen zu entscheiden, muß auf die genaue Bedeutung des theoretischen Konstruktes Postmaterialismus in der Postmaterialismustheorie Ingleharts eingegangen werden. Zu fragen ist, welche Indikatoren (in welchem länderspezifischen soziokulturellen Kontext (die Skala wird international vergleichend eingesetzt)) die Inhalte der im Postmaterialismus definierten Werthaltungen am validesten wiedergeben. Neben theoretisch-inhaltlichen Überlegungen im Hinblick auf die in den einzelnen Items formulierten Sachverhalte (Inhaltsvalidität) ist dabei auch die Überprüfung der theoretisch postulierten Beziehung des Postmaterialismus (P) mit anderen relevanten Variablen der Theorie von Bedeutung (Kriteriumsvalidität, siehe 5.4.1.2). Beschränkt man sich dabei auf die zentralen Variablen Alter (A) und Bildung (B), so ergeben sich für Version I substantiell bedeutsame Korrelationen, die mit den Erwartungen der Theorie im Einklang stehen (r_{PA} ist negativ und r_{PB} ist positiv) während sich für Version II lediglich eine schwache Beziehung für r_{PA} ergibt. Diese Ergebnisse legen, zusammen mit zusätzlichen Überlegungen, die die inhaltliche Bedeutung der einzelnen Items betreffen, den Schluß nahe, daß es sich bei der ersten Version um eine validere Umsetzung des Postmaterialismuskonstruktes aus der Theorie Inglehards handelt.

Unter Operationalisierung im engeren Sinne versteht man die Operationalisierung der einzelnen Indikatoren, also die Angabe von konkreten Meßvorschriften für die Indikatoren. Die Operationalisierung ist also ein zweistufiger Prozeß: Zunächst müssen für ein theoretisches Konstrukt (z. B. die normative Geschlechtsrollenorientierung) Indikatoren gesucht werden. Im Falle einer Likert-Skala zur nGRO (siehe 5.3.1) sind die Items der Likert-Skala die Indikatoren. Zusätzlich müssen diese Indikatoren in konkrete Meßvorschriften umgesetzt werden. Im Falle der Items der Likert-Skala ist etwa über die konkrete Gestaltung des Fragebogens, die Wahl einer 5-stufigen bzw. 6-stufigen Rating-Skala, die genaue Vorgehensweise bei der Befragung, usw. zu entscheiden.

Im Zusammenhang mit jeder Messung müssen die folgenden Gütekriterien beachtet werden (siehe genauer in 5.4) :

- **Objektivität:** Die Objektivität gibt an, in welchem Ausmaß eine Messung unabhängig vom Untersucher ist (mangelnde Objektivität kann z. B. durch Interviewereffekte in Ab-

hängigkeit vom Geschlecht des Interviewers bei der Erhebung der nGRO (vgl. 6.1) verursacht werden).

- **Reliabilität (Zuverlässigkeit):** Die Reliabilität betrifft die Genauigkeit oder Präzision einer Messung. Insbesondere sollten sich bei einer Meßwiederholung genau die gleichen Werte ergeben, wie bei der ersten Messung. Mangelnde Reliabilität wird durch Zufallsfehler bei einer Messung hervorgerufen, z. B. durch Ablesefehler oder mangelnde Konzentration bzw. mangelndes Interesse von Befragten im Interview (vgl. 6.1).
- **Validität (Gültigkeit):** Die Validität einer Messung hängt davon ab, ob wirklich das gemessen wird, was eigentlich inhaltlich gemessen werden soll. Eine Messung kann präzise im Sinne der Reliabilität sein, ohne jedoch eine Information über den zugrundeliegenden Sachverhalt zu geben (so z. B., wenn ein Befragter sozial erwünschte Antworten gibt oder tendenziell allen Fragen zustimmt (Ja-Sage-Tendenz; siehe auch 6.1)).

Operationalisierung von theoretischen Konstrukten als zweistufiger Prozeß

(die jeweiligen Gütekriterien sind in eckigen Klammern beigefügt).

L_T :	<u>theoretisches Konstrukt</u> (z.B. Intelligenz)		[theoretische Fruchtbarkeit des Begriffes]
	Operationalisierung im weiteren Sinne (Lösung des Korrespondenzproblems)	Zuordnung von Indikatoren (Auswahl aus dem Indikatorenuniversum)	[Validität der Indikatoren]
L_O :	<u>Indikatoren</u> (z.B. Aufgaben des Intelligenztests)		
	Operationalisierung im engeren Sinne	Umsetzen in eine konkrete Meßvorschrift: Wahl des Meßmodells, Gestaltung und Durchführung des Tests	[Objektivität, Reliabilität, Validität]
Messung:	<u>konkrete Meßvorschrift</u> (konkrete Vorgaben für die Gestaltung, Durchführung und Auswertung des Intelligenztests)		

4.3 Der kausalanalytische Ansatz

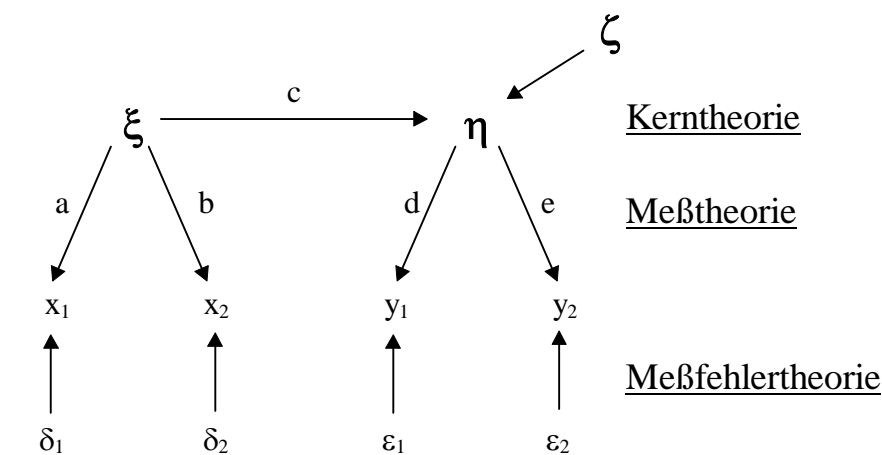
Der kausalanalytische Ansatz bietet die Möglichkeit, die Koeffizienten der Kerntheorie, der Meßtheorie und der Meßfehlertheorie in einem integrierten Kausalmodell zu schätzen - und zwar allein auf der Grundlage der empirisch erhobenen Korrelationen zwischen den Indikato-

ren. Allerdings müssen dazu einige zusätzliche Annahmen (insbesondere über die Meßfehler) gemacht werden, damit das formale Modell lösbar ist. Im folgenden diskutieren wir das einfachste Modell (eine bivariate Beziehung als Kerntheorie; jedes theoretische Konstrukt wird durch zwei Indikatoren erfaßt) in allen Einzelheiten.

In der Praxis verwendet man Modelle mit komplexen multivariaten kausalen Zusammenhängen und einer Vielzahl von Indikatoren, die mit Hilfe von speziellen Analyseprogrammen, wie z. B. LISREL (vgl. Kap. 8), gelöst werden. Ziel unserer Darstellung kann es daher nur sein zu zeigen,

1. daß solche Modelle überhaupt gelöst werden können und
2. daß mit der zusätzlichen expliziten Berücksichtigung eines Meßmodells (Meßtheorie und Meßfehlertheorie) genauere Aussagen gemacht werden können, die zu inhaltlich bedeutsamen Modifikationen in den Schlußfolgerungen empirischer Untersuchungen führen. Deutlicher ausgedrückt: Untersuchungen ohne die Berücksichtigung von Meßmodellen führen zu verzerrten Ergebnissen und damit auch zu inhaltlich falschen Schlußfolgerungen.

Das Kausalmodell im einzelnen:



$$\left. \begin{aligned}
 x_1 &= a\xi + \delta_1 \\
 x_2 &= b\xi + \delta_2 \\
 y_1 &= d\eta + \varepsilon_1 \\
 y_2 &= e\eta + \varepsilon_2
 \end{aligned} \right\} \text{Meßtheorie} \quad [\text{Strukturgleichungen}]$$

$$\eta = c\xi + \zeta \quad \text{Kerntheorie}$$

Zur Lösung des Modells werden die folgenden Annahmen gemacht:

- Standardisierte Variablen: Alle im Modell enthaltenen Variablen sind standardisiert, d.h. es handelt sich um intervallskalierte Variablen mit einem Mittelwert von 0 und einer Standardabweichung von 1.
- Lineare Beziehungen: Alle Beziehungen im Modell sind linear (vgl. die Gleichungen des Modells).

- Unabhängige Meßfehler: Als Meßfehlertheorie spezifizieren wir unabhängige Zufallsfehler, d.h. die Fehlervariablen $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ sind untereinander und mit den anderen latenten Variablen unkorreliert. Außerdem sind ξ und ζ unkorreliert. Diese Voraussetzung ist die weitreichenste und zugleich am wenigsten realistische Annahme, die wir zur Lösung des Modells machen müssen (vgl. auch 3.1). Üblicherweise enthalten solche Modelle weitere Indikatoren für die in der Kerntheorie enthaltenen theoretischen Konstrukte und damit auch zusätzliche empirische Informationen (Korrelationen in der zugehörigen Korrelationsmatrix). Ist dies der Fall, können die hier gemachten restriktiven Annahmen schrittweise gelockert werden, so daß eine realitätsnähere Meßfehlertheorie entsteht, die z.B. auch korrelierte Meßfehler zuläßt (vgl. auch die Diskussion der Annahmen der klassischen Testtheorie in 5.4.1).

Grundlage der Berechnung der zu schätzenden fünf Modellparameter a, b, c, d, e ist die empirisch gegebene Korrelationsmatrix der Indikatoren x_1, x_2, y_1, y_2 . Da das Modell mehr empirische Informationen als zu schätzende Parameter enthält, das Modell also überidentifiziert ist (6 empirische Korrelationskoeffizienten bei nur 5 zu schätzenden theoretischen Parametern), bleibt ein Freiheitsgrad zu Testzwecken. Es kann also eine Testgleichung abgeleitet werden, d.h. ein Zusammenhang zwischen den empirisch ermittelten Korrelationen, der erfüllt sein muß, wenn die Voraussetzungen des Modells gelten sollen.

Nach den Regeln der Pfadanalyse (vgl. insbesondere 3.1.1.3) ergeben sich aus den obigen Voraussetzungen die folgenden Gleichungen (Effektzerlegung):

$$r_{x_1x_2} = ab$$

$$r_{y_1y_2} = de$$

$$r_{x_1y_1} = acd$$

$$r_{x_1y_2} = ace$$

$$r_{x_2y_1} = bcd$$

$$r_{x_2y_2} = bce$$

Als Testgleichung ergibt sich:

$$r_{x_1y_1} r_{x_2y_2} = r_{x_1y_2} r_{x_2y_1}$$

Ersetzt man in der Testgleichung die Korrelationen durch die zugehörigen Produkte von Pfadkoeffizienten, so erkennt man, daß die Testgleichung tatsächlich aus den Modellvoraussetzungen abgeleitet werden kann:

$$acd \cdot bce = ace \cdot bcd$$

$$abc^2 de = abc^2 de \quad .$$

Für die Parameter des Modells ergibt sich daraus:

$$c^2 = \frac{r_{x_1y_2} r_{x_2y_1}}{r_{x_1x_2} r_{y_1y_2}} = \frac{r_{x_1y_1} r_{x_2y_2}}{r_{x_1x_2} r_{y_1y_2}}$$

$$a^2 = r_{x_1x_2} \frac{r_{x_1y_2}}{r_{x_2y_2}} = r_{x_1x_2} \frac{r_{x_1y_1}}{r_{x_2y_1}}$$

$$b^2 = r_{x_1x_2} \frac{r_{x_2y_2}}{r_{x_1y_2}} = r_{x_1x_2} \frac{r_{x_2y_1}}{r_{x_1y_1}}$$

$$d^2 = r_{y_1y_2} \frac{r_{x_2y_1}}{r_{x_2y_2}} = r_{y_1y_2} \frac{r_{x_1y_1}}{r_{x_1y_2}}$$

$$e^2 = r_{y_1y_2} \frac{r_{x_2y_2}}{r_{x_2y_1}} = r_{y_1y_2} \frac{r_{x_1y_2}}{r_{x_1y_1}}$$

Die Pfadkoeffizienten sind also nur bis auf das Vorzeichen bestimmt. Die Vorzeichen der Pfadkoeffizienten müssen aus theoretischen Überlegungen über die Richtung der Kausalbeziehungen und die Polung der Indikatoren erschlossen werden. Die Polung erfolgt konventionell, ist aber im Rahmen der Überprüfung des Modells fest vorgegeben. Wichtig ist, daß die die Vorzeichen der Pfadkoeffizienten betreffenden theoretischen Überlegungen vor der empirischen Überprüfung des Modells angestellt werden, so daß ein zusätzlicher Modelltest möglich ist, indem die Vorzeichen der empirischen Korrelationen zwischen den Indikatoren (über die Effektzerlegungen) mit den postulierten Vorzeichen der Pfadkoeffizienten verglichen werden.

Beispiel zum kausalanalytischen Ansatz

Kerntheorie: Soziale Distanz $\xrightarrow{+}$ Ausländerfeindlichkeit

Korrelationsmatrix der Indikatoren aus dem ALLBUS 1980

(6 empirische Informationen):

	r_{ij}	x_1	x_2	y_1	y_2	
Arbeitsplatz	x_1	1.00				(negative) Indikatoren für soziale Distanz
Freundeskreiss	x_2	+ .315	1.00			(ξ)
Ehepartner	y_1	-.195	-.370	1.00		(positive) Indikatoren für Ausländerfeindlichkeit (η)
politische Tätigkeit	y_2	-.094	-.241	+.563	1.00	

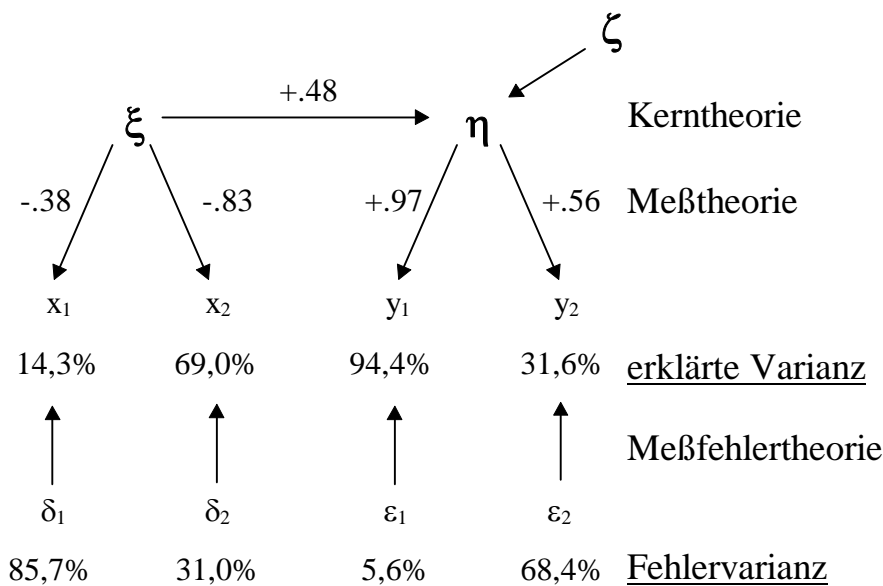
Die Indikatoren für soziale Distanz sind jeweils so gepolt, daß ein positiver Wert das Vorliegen von sozialen Beziehungen anzeigt. Die Variablen Arbeitsplatz und Freundeskreis stellen daher negative Indikatoren für soziale Distanz dar. Wir erwarten daher einen negativen Wert für die entsprechenden Faktorladungen a und b. Die Polung für die Indikatoren zur Ausländerfeindlichkeit erfolgte derart, daß ein hoher Wert eine starke Ausländerfeindlichkeit anzeigt. Die Variablen Ehepartner und politische Tätigkeit sind daher positive Indikatoren für Ausländerfeindlichkeit. Dementsprechend erwarten wir positive Werte für die Faktorladungen d und e. Damit können Vorhersagen über die Korrelationen zwischen den Indikatoren abgeleitet werden. Wir erwarten positive Beziehungen zwischen den Indikatoren des gleichen Konstrukts und negative Beziehungen zwischen den Indikatoren verschiedener Konstrukte. Wie ein Blick auf die empirischen Korrelationen aus der ALLBUS-Studie zeigt, bestätigen sich diese theoretischen Voraussagen in unserem Beispiel. Die Überprüfung der Testgleichung ergibt folgendes Ergebnis:

$$r_{x_1y_1} r_{x_2y_2} \stackrel{!}{=} r_{x_1y_2} r_{x_2y_1}$$

$$(-.195)(-.241) \stackrel{?}{=} (-.094)(-.370) \quad \text{Testgleichung}$$

$$.0469 \approx .0349$$

Für die folgenden Überlegungen gehen wir davon aus, daß die Testgleichung erfüllt ist, d.h. daß die beiden Seiten der Testgleichung nicht signifikant voneinander abweichen. Damit können die Parameter des Modells aus den obigen Gleichungen berechnet werden, wobei zu berücksichtigen ist, daß die beiden Versionen der Parametergleichungen nur dann zu den gleichen Resultaten führen, wenn die Testgleichung exakt erfüllt ist. (In unserem Beispiel wurde das geometrische Mittel aus den beiden Parameterschätzungen gebildet):



Die Kerntheorie wird durch das Modell bestätigt - der Pfadkoeffizient ist mit .48 sogar relativ hoch im Vergleich zu anderen empirischen Untersuchungen (obwohl dies nur einem Anteil

der erklärten Varianz der Ausländerfeindlichkeit von 23% entspricht, während der Rest von 77% auf andere Einflüsse (zusammengefaßt in dem Fehlerterm ζ) zurückgeführt werden muß). Wie an den Faktorladungen (Pfadkoeffizienten) zu erkennen ist, sind die verwendeten Indikatoren von durchaus unterschiedlicher Qualität. Während der Indikator x_2 (Freundeskreis) mit einer Faktorladung von $-.83$ überwiegend von dem zu messenden theoretischen Konstrukt abhängt, gilt dies für x_1 nur zu einem sehr geringen Teil, nämlich zu 14,3%. Genauer ausgedrückt: Die Bestätigung (bzw. Verneinung) der Frage nach einer Bekanntschaft mit Ausländern im Freundeskreis (x_2) durch die Befragten läßt sich zu 69% ($-.83^2$) auf deren soziale Distanz zu Ausländern zurückführen. Zu 31% hängt die Beantwortung dieses Items von anderen (im Sinne der Meßfehlertheorie zufälligen!) Einflüssen (δ_2) ab. Die Bestätigung (bzw. Verneinung) der Frage nach einer Bekanntschaft mit Ausländern am Arbeitsplatz (x_1) kann dagegen nur zu 14,3% ($-.38^2$) auf die soziale Distanz der Befragten zurückgeführt werden - zu 85,7% spielen hier andere (zufällige!) Einflüsse (δ_1) eine Rolle. Die erklärte Varianz entspricht gerade der Reliabilität der Indikatoren (siehe die Ausführungen zur klassischen Testtheorie in 5.4.1): Die höchste Reliabilität mit $.944$ hat demnach der Indikator y_1 (Ehepartner). Das zugrundeliegende theoretische Konstrukt Ausländerfeindlichkeit kommt in y_1 in fast reiner Form zum Ausdruck - lediglich ein Varianzanteil von 5,4% ist auf den zugehörigen Zufallsfehler ε_1 zurückzuführen.

Hätte man versucht, die Beziehung zwischen sozialer Distanz und Ausländerfeindlichkeit ohne Spezifikation eines Meßmodells allein aus den Korrelationen der beteiligten Indikatoren zu erschließen, so hätte man je nach Art der verwendeten Indikatoren Werte zwischen $.094$ und $.37$ erhalten (vgl. die Korrelationsmatrix). Wie das Modell zeigt, ist die um Meßfehler bereinigte Korrelation zwischen den theoretischen Konstrukten mit $.48$ deutlich höher. Die empirischen Korrelationen sind also im Vergleich zu der theoretischen Korrelation vermindert, da in ihnen auch die Meßfehler (Zufallsfehler) zum Ausdruck kommen (mangelnde Reliabilität der Indikatoren). Die Korrelationen zwischen den empirischen Indikatoren sind dabei um so höher, je höher die Faktorladungen (theoretische Validität der Indikatoren im Rahmen des Modells; siehe 5.4.1.2) und die zugehörigen Reliabilitäten sind. So ist z.B. die Korrelation zwischen den beiden reliabelsten Indikatoren r_{x_2, y_1} mit $.37$ nur geringfügig kleiner als die zugehörige theoretische Korrelation von $.48$.

Im Rahmen der klassischen Testtheorie versucht man, die Reliabilität von Skalen, die aus einer Vielzahl von Indikatoren zusammengesetzt sind (z.B. von Likert-Skalen), zu schätzen; eine Möglichkeit ist z.B. die Berechnung von Cronbachs α (vgl. 5.4.1.1). Sind die Reliabilitäten bekannt, so kann mit Hilfe einer mathematischen Formel eine Minderungskorrektur vorgenommen werden, die die durch die Einwirkung von Zufallsfehlern (mangelnde Reliabilität) in den Skalen verursachte Abschwächung der theoretischen Korrelationen zwischen den Skalen wieder ausgleicht (siehe 5.4.1.1). Der kausalanalytische Ansatz ist demgegenüber der fortgeschrittenere und elegantere, da er auf eine Skalenbildung im Sinne von additiven Indizes (wie z.B. bei der Likert-Skalierung; vgl. 5.3.1) verzichtet und die

einzelnen Items explizit im Sinne einer Meß- und Meßfehlertheorie in das Kausalmodell integriert. Eine Minderungskorrektur ist beim kausalanalytischen Ansatz unnötig, da in den berechneten theoretischen Korrelationen die mangelnde Reliabilität der Indikatoren „quasi automatisch“ berücksichtigt wird.

Einfluß unterschiedlicher Reliabilitäten

Die Berücksichtigung von Meßmodellen und die damit verbundene Berücksichtigung unterschiedlicher Reliabilitäten kann die inhaltlichen Schlußfolgerungen, die aus einer empirischen Untersuchung gezogen werden müssen, deutlich verändern. Generell gilt, daß die Korrelationen zwischen den theoretischen Konstrukten größer sind als die entsprechenden empirischen Korrelationen zwischen den zugeordneten Indikatoren - und zwar um so mehr, je geringer die Reliabilität der Indikatoren ist. Vergleicht man also verschiedene Gruppen, die sich durch unterschiedliche Reliabilitäten (z.B. durch ein gruppenspezifisches Antwortverhalten bedingt) auszeichnen, so ist dieser Vergleich notwendigerweise verzerrt, solange die erforderlichen Minderungskorrekturen, die in beiden Gruppen unterschiedlich stark ausfallen werden, nicht durchgeführt sind.

Beispiel: Empirische Untersuchung des Statuszuweisungsprozesses in den USA für Schwarze und Nicht-Schwarze (vgl. Bielby/Hauser/Featherman 1977):

Kerntheorie: Statuszuweisungsmodell (vgl. auch 3.1.1.3)

Meßmodell mit 3 Meßzeitpunkten (z. B. fo_1, fo_2, fo_3 für FO)

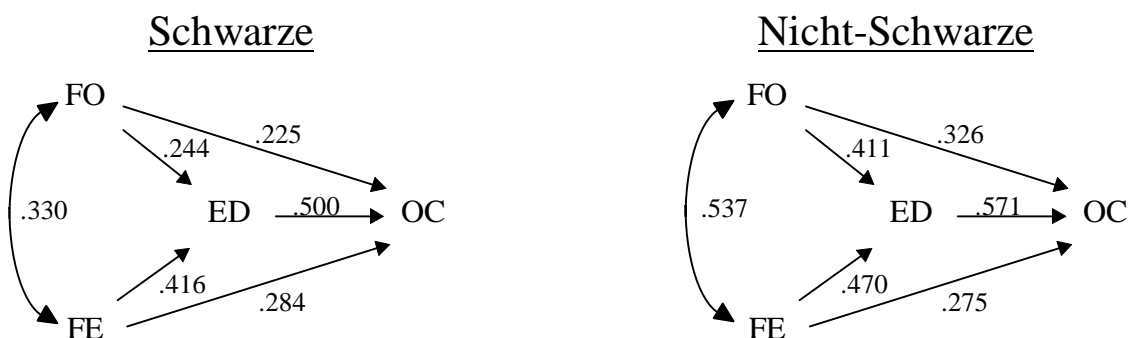
FO = Fathers Occupation

ED = Respondents Education

FE = Fathers Education

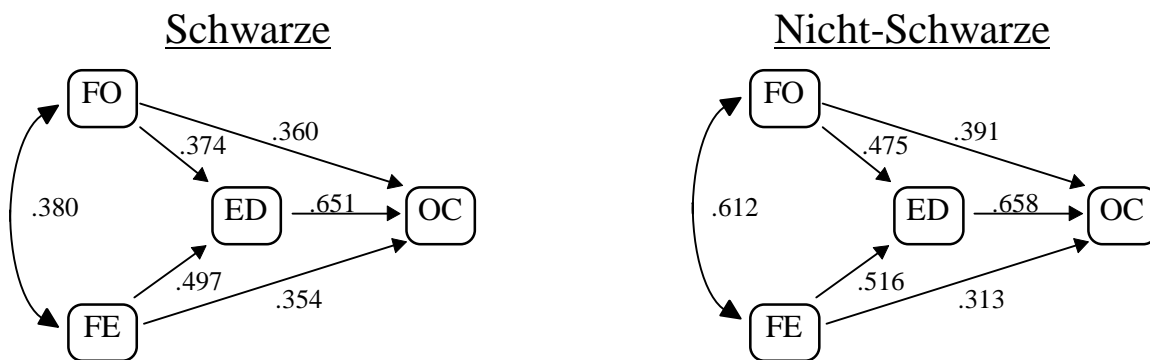
OC = Respondents Occupation

a) unkorrigiertes Modell (nur ein Indikator, z.B. $FO \equiv fo_1$)

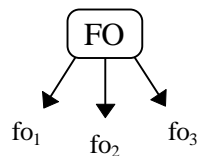


b) korrigiertes Modell (Modell, das auch Meßfehler berücksichtigt):

Die geringere Reliabilität bei den Angaben der Schwarzen bedingt eine stärkere Minderungskorrektur.



Die Meßmodelle werden nicht explizit mit aufgeführt, z. B.:



Schlußfolgerung: Berücksichtigt man Meßfehler (größere Ungenauigkeit (geringere Reliabilität) in den Antworten der Schwarzen), so ist die Wirkung der Bildung (ED) auf den erreichten Berufsstatus (OC) in beiden Gruppen gleich, d.h. Bildungsinvestitionen sind für Schwarze genauso wirkungsvoll wie für die übrige Bevölkerung in den USA.

[Bemerkung: Genaugenommen ist ein Gruppenvergleich mit standardisierten Variablen nicht so aussagekräftig wie ein entsprechender Vergleich mit unstandardisierten Variablen auf der Grundlage von Kovarianzstrukturen, der auch Unterschiede in den Varianzen zwischen den Gruppen berücksichtigt.]