

GESIS

ZUMA

45

November 1999

NACHRICHTEN

AMU

Zentrum für Umfragen, Methoden und Analysen (ZUMA)

ZUMA ist Mitglied der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen e.V. (GESIS). Die GESIS ist eine Einrichtung der „Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz“ (WGL).

Vorsitzender des Trägervereins ZUMA e.V.: Prof. Dr. Max Kaase

Direktor: Prof. Dr. Peter Ph. Mohler

Hausanschrift

B 2, 1
68 159 Mannheim

Postanschrift

Postfach 12 21 55
68 072 Mannheim

Telefon 0621/1246 - 0

Fax 0621/1246 - 100

Redaktion: 0621/1246 - 268

E-Mail zuma-nachrichten@zuma-mannheim.de

Internet <http://www.zuma-mannheim.de/> ZUMA

<http://www.social-science-geis.de/> GESIS

Redaktion: Dr. Paul Lüttinger

ISSN 0941-1670 23. Jahrgang

Die ZUMA-Nachrichten informieren über Forschungsarbeiten und Dienstleistungen des Instituts. Beiträge externer Autoren werden dann aufgenommen, wenn sie aus Kooperationsprojekten mit ZUMA resultieren oder von besonderem wissenschaftlichen Interesse für die Arbeiten im Hause und die Empirische Sozialforschung sind. Alle Beiträge unterliegen einem internen förmlichen Begutachtungsverfahren durch jeweils zwei Gutachter. Der Herausgeber entscheidet auf der Grundlage der Gutachten über die Annahme eines Beitrages. Hinweise zur Manuskriptgestaltung sind auf der letzten Seite des Heftes.

Namentlich gekennzeichnete Beiträge geben die Meinung der Autoren wieder. Der Nachdruck von Beiträgen ist nach Absprache möglich. Die ZUMA-Nachrichten erscheinen im Mai und November eines Jahres. Sie werden Interessenten auf Anforderung kostenlos zugesandt. Die ZUMA-Nachrichten finden Sie auch im WWW unter der Adresse: <http://www.zuma-mannheim.de/publications/periodicals/zuma-nachrichten/>.

Druck: Verlag Pfälzische Post GmbH, Neustadt/Weinstraße.

Gedruckt auf chlorfrei gebleichtem Papier.

© ZUMA

25 JAHRE ZUMA

Die Entwicklung der Umfrageforschung in der Bundesrepublik Deutschland
in den siebziger und achtziger Jahren.

Erwin K. Scheuch

Sozialberichterstattung in der Bundesrepublik Deutschland.

Wolfgang Zapf

Empirische Sozialforschung in den neunziger Jahren.

Karl Ulrich Mayer

FORSCHUNGSBERICHTE

Die Befragung von Arbeitsmigranten: Einwohnermeldeamt-Stichprobe
und telefonische Erhebung?

Nadia Granato

Messung von internen und externen Kontrollüberzeugungen in allgemeinen
Bevölkerungsumfragen.

Nina Jakoby und Rüdiger Jacob

Thematik oder Incentives? Zur Erhöhung der Rücklaufquoten bei
postalischen Befragungen.

Rolf Porst

Interaktionseffekte in Strukturgleichungsmodellen mit der Theorie des geplanten
Verhaltens: Multiple Gruppenvergleiche und Produktterme mit latenten Variablen.

Jost Reinecke

Die Stichprobenziehung bei Zuwandererbefragungen.

Kurt Salentin

MITTEILUNGEN

Kostengünstiger Zugang zu amtlichen Mikrozensusdaten der Jahre 1989
1991, 1993 und 1997

Thesen zur Verbesserung der informationellen Infrastruktur in der
Bundesrepublik Deutschland

Förderpreis des Statistischen Bundesamtes für ZUMA-Dissertation

NSDstat - ein Programm zur schnellen Datenanalyse

Testdaten zum IAB-Betriebspanel erhältlich

BUCHBESPRECHUNGEN

P. H. HARTMANN: Lebensstilforschung. Darstellung, Kritik und Weiterentwicklung

N. GILBERT/K. G. TROITZSCH: Simulation for the Social Scientist

IN EIGENER SACHE

Das Zentrum für Umfragen, Methoden und Analysen (ZUMA) feierte am 4. November sein 25jähriges Bestehen. Wir haben dieses Jubiläum im Kreise vieler Kolleginnen und Kollegen und früherer wie heutiger Mitarbeiterinnen und Mitarbeiter begangen. Auf dem Programm standen mehrere Festvorträge, die sich mit der Entwicklung der Empirischen Sozialforschung befassen und die wir im vorliegenden Heft veröffentlichen. *Erwin K. Scheuch* referierte über die Entwicklung der Umfrageforschung in der Bundesrepublik Deutschland in den siebziger und achtziger Jahren. Die Sozialindikatorenforschung in der Bundesrepublik Deutschland war das Thema von *Wolfgang Zapf*, und *Karl Ulrich Mayer* beschäftigte sich in seinem Vortrag mit der „Empirischen Sozialforschung in den neunziger Jahren“. Wir danken den Referenten für die Genehmigung des Abdrucks ihrer Vorträge.

Außerdem berichten in dieser Ausgabe der ZUMA-Nachrichten *Nina Jakoby* und *Rüdiger Jacob* über die Entwicklung zweier Kurzskalen zur Messung von internen und externen Kontrollüberzeugungen, die in allgemeinen Bevölkerungsumfragen eingesetzt werden können. Die Items wurden im Rahmen des ZUMA Sozialwissenschaften-BUS getestet.

Auch bei der Befragung spezieller Populationen wie z.B. von Arbeitsmigranten empfiehlt es sich, wann immer möglich die Stichprobe aus dem Einwohnermelderegister zu ziehen. Inwieweit eine telefonische Befragung möglich ist, hängt entscheidend von der Telefondichte ab sowie von der Identifizierungsquote, d.h. dem Anteil der Stichprobenelemente, für die eine Telefonnummer ermittelt werden kann. Die Ergebnisse einer Studie, die sich mit der Analyse dieser beiden Größen bei Migranten und speziell bei Türken befaßt, werden von *Nadia Granato* vorgestellt.

Mit einem klassischen Thema der Empirischen Sozialforschung beschäftigt sich *Rolf Porst*, der im Rahmen einer postalischen Befragung sowohl den Einfluß aktueller und interessanter Einstiegsfragen als auch die Wirkung von Anreizen (incentives) auf die Teilnahmebereitschaft untersuchte. Es zeigt sich, daß sowohl ein Anreiz in Form von schönen Briefmarken als auch „interessante“ Einstiegsfragen die Ausschöpfungsquote bei schriftlichen Befragungen steigern können.

Jost Reinecke diskutiert in seinem Beitrag zur Erweiterung der Theorie linearer Strukturgleichungsmodelle Modellierungsmöglichkeiten von Interaktionseffekten in solchen Modellen. WertxErwartungsprodukte zu den Konstrukten Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle, formuliert in der Theorie des geplanten Ver-

haltens, werden hierzu herangezogen. Anhand einer repräsentativen Stichprobe von Jugendlichen und jungen Erwachsenen kann mit multiplen Gruppenvergleichen und latenten Produktmodellen gezeigt werden, daß für das Wert×Erwartungsprodukt der wahrgenommenen Verhaltenskontrolle ein signifikanter Interaktionseffekt vorliegt. Der Stellenwert unterschiedlicher Schätzverfahren (ML, GLS und WLS) wird in bezug auf die latenten Produktmodelle diskutiert.

Kurt Salentin zeigt in seinem Beitrag, welchen Einfluß verschiedene Stichprobenverfahren - Melderegisterauszüge, Auswahlen auf der Grundlage von Gebietseinheiten, Schneeballverfahren und namensorientierte Verfahren - auf die Repräsentativität von Zuwandererstichproben haben und welchen Aufwand sie verursachen.

Vor zwei Jahren konnten wir erstmals darüber berichten, daß das Statistische Bundesamt amtliche Mikrodaten in Form von Scientific Use Files der Empirischen Sozialforschung gegen eine geringe Bereitstellungsgebühr von 130 Mark zur Verfügung stellt. Die Reihe der verfügbaren Daten wurde nun beträchtlich erweitert: Beim Statistischen Bundesamt können mittlerweile auch die faktisch anonymisierten Mikrozensusdaten der Jahre 1989, 1991, 1993 und 1997 zu den gleichen Sonderkonditionen für Forschungszwecke bestellt werden. Ermöglicht wurde dies durch eine Vereinbarung zwischen dem Statistischen Bundesamt und der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS e.V.), welche die Grundkosten für die Bereitstellung der Mikrozensus übernimmt.

Heike Wirth, Soziologin am Zentrum für Umfragen, Methoden und Analysen, erhielt für ihre Doktorarbeit zum Thema „Bildung, Klassenlage und Partnerwahl. Eine empirische Analyse der bildungs- und klassenspezifischen Heiratsbeziehungen“ einen Förderpreis des Statistischen Bundesamtes. Wir gratulieren Frau Wirth herzlich zu dieser Auszeichnung.

Ausgeschieden bei ZUMA ist Klaudia Erhardt.

Die Auslieferung der ZUMA-Nachrichten erfolgt diesmal erst zu Beginn des neuen Jahres. Ich hoffe, Sie hatten friedvolle Feiertage und einen angenehmen Übergang ins Jahr 2000, für das ich Ihnen alles Gute wünsche.

Max Kaase
Vorsitzender des ZUMA e.V.

DIE ENTWICKLUNG DER UMFRAGEFORSCHUNG IN DER BUNDESREPUBLIK DEUTSCHLAND IN DEN SIEBZIGER UND ACHTZIGER JAHREN¹

ERWIN K. SCHEUCH

1. Zur Situation vor der Konsolidierung²

Erst mit Ende der sechziger Jahre hatte sich die Umfrageforschung als Vorgehensweise in der Bundesrepublik voll etabliert. Das war das Ergebnis einer Entwicklung vornehmlich nach 1951, und als deren Beginn darf eine Tagung in Weinheim im Jahre 1951 angesehen werden.

Wie auch in den USA und einigen anderen Ländern reicht die empirische Forschung weiter zurück und kann auch für Deutschland, in Einzeluntersuchungen praktiziert, in den zwanziger Jahren ausgemacht werden. Unabhängig von der Forschung mit sozialwissenschaftlicher Ausrichtung hatte sich damals eine betriebswirtschaftlichen Zwecken verpflichtete Marktforschung ausgebildet, die bereits 1934 mit der Gründung der Gesellschaft für Konsumforschung (GfK) eine erste institutionelle Verfestigung erhielt. Parallele Entwicklungen sind auch in England auszumachen, wo die Beobachtung der öffentlichen Meinung durch das Institut Mass Observations besonders hervorgehoben werden kann.

-
- 1** Vortrag anlässlich des 25jährigen Bestehens von ZUMA, Mannheim, 4. November 1999.
 - 2** Etwas ausführlicher hierzu E. K. Scheuch: „Von der Pioniertat zur Institution - Beobachtungen zur Entwicklung empirischer Sozialforschung“. In: D. Franke/J. Scharioth (Hg.): 40 Jahre Markt- und Sozialforschung in der Bundesrepublik Deutschland. München 1990: 43-67. Zur internationalen Einbettung dieser Entwicklung finden sich - bei aller Einseitigkeit ihrer Aufmerksamkeit - interessante Einzelheiten bei Elisabeth Noelle-Neumann: „Political Opinion Polling in Germany“. In: Robert M. Worcester (Hg.): Political Opinion Polling - An International Review. London 1983: 44-60. Siehe auch - allerdings bei einigen Beiträgen mit kritischer Aufmerksamkeit - Dieter Kaufmann et. al. (Hg.): Empirische Sozialforschung im vereinten Deutschland. Frankfurt 1992.

All das war selbstverständlich durch den Krieg unterbrochen worden. Gefördert durch die Besatzungsmächte nach 1945, dabei insbesondere englische und amerikanische Initiativen, lernten Deutsche die Techniken der Forschung als Importe³. Dabei gab es interessante Einzeluntersuchungen, aber von einer eigentlichen Entwicklung kann man erst mit der Tagung in Weinheim sprechen. Auf Einladung der amerikanischen High Commission Germany trafen sich dort 38 Vertreter von Universitätsinstituten und von 30 privatwirtschaftlich verfaßten Einheiten - wobei bezeichnenderweise das einzige Institut mit einer zurückreichenden Tradition, nämlich die GfK, eine Teilnahme nicht für angezeigt hielt. Es war eine Tagung mit lebhaften Auseinandersetzungen, wobei die damals berühmtesten Professoren Leopold von Wiese und Theodor W. Adorno ihre Skepsis gegenüber den quantifizierenden Verfahren ausdrückten.

Damit teilten sie die vorherrschenden Ansichten in der qualifizierten Öffentlichkeit Deutschlands. Noch einige Jahre später konnte Adenauer auf ein ihm mißfallendes Ergebnis Zustimmung mit seiner abwertenden Äußerung finden, es seien lediglich etwa 2.000 Menschen befragt worden, und das sei doch keine Grundlage für eine Verallgemeinerung. Es dauerte Jahrzehnte, bis der Gedanke mehrheitsfähig wurde, eine nach methodischen Prinzipien ausgewählte Zahl von 2.000 oder weniger Personen sei eine zureichende Grundlage für einen Schluß vom Teil der Befragten auf das Ganze der Bevölkerung. Es war dies eben eine kontra-intuitive Behauptung, die erst durch Gewöhnung und eine Behandlung als Selbstverständlichkeit durch Medien und Wirtschaft Akzeptanz fand.

Auf der Weinheimer Tagung waren - bis auf die Abwesenheit von GfK - Sozialwissenschaftler und Marktforscher noch nicht getrennt. Sie vereinte der Gebrauch des gleichen Instrumentariums, was zum damaligen Zeitpunkt weitgehend die Befragung eines repräsentativen Querschnitts mit einem standardisierten Fragebogen bedeutete. Aufgrund von Anregungen in anderen Ländern Westeuropas gründeten die Marktforscher dann am 20. Juni 1955 in Würzburg den „Arbeitskreis für betriebswirtschaftliche Markt- und Absatz-

3 Es gibt eine ziemlich desinformative Spezialität: „Die Soziologie, insbesondere die empirische Forschung hat auch im NS-Staat prosperiert“, deren neuerdings lautstärkster Adept ein C. Klingemann ist. Hier ist nicht der Ort, auf die Willkür bei der Zurechnung von Personen zur empirischen Soziologie einzugehen, noch auf die Blindheit - oder Ignoranz - von Klingemann gegenüber der „deutschen Soziologie“ um Hans Freyer in Leipzig, oder die Ahnungslosigkeit des „Antifa“ Klingemann über den Alltag im NS-Staat. Bis auf Helmut Schelsky haben Personen mit NS-Kontakten nach 1945 in der Entwicklung der Empirie an Universitäten keine Rolle gespielt. Sozialforschung wurde als Importe rezipiert.

forschung“. Voraussetzung für die Mitgliedschaft war eine Institutionalisierung mit mindestens zehn Mitarbeitern. Aus dieser Vereinigung wurde inzwischen der „Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute“ (ADM). Die Änderung des Titels ist Ergebnis grundsätzlicher Überlegungen gewesen.

Parallel hierzu organisierten sich die Sozialforscher, soweit sie in institutionalisierten Kontexten tätig sein konnten, in der „Arbeitsgemeinschaft Sozialwissenschaftlicher Institute“ (ASI). Mit der Gründung der ASI wurden hochschulfreie und Hochschulinstitute in einem gemeinsamen Kontext zusammengeführt, was keineswegs selbstverständlich war.

Bei der weiteren Entwicklung der quantitativen Forschung für wirtschaftliche und sozialwissenschaftliche Fragestellungen war es im deutschen Sprachbereich von großer Bedeutung, daß die zentrifugalen Tendenzen in der Entwicklung von Hochschulinstituten, hochschulfreien gemeinnützigen Einheiten und Instituten auf privatwirtschaftlicher Grundlage persönlich und organisatorisch konterkariert wurden. Bei späteren Auseinandersetzungen, auf die noch eingegangen wird, erwies sich dies als eine Voraussetzung für die Verteidigung des rechtlichen und wirtschaftlichen Rahmens der Forschung.

An den Hochschulen wurde die quantitative Forschung zunächst nur an einigen Plätzen akzeptiert. Soweit sich die Sozialwissenschaften als Geisteswissenschaften verstanden, wirkte diese Art quantitativer Forschung wie ein Fremdkörper. Mit diesen Problemen hatte die wirtschaftlichen Zwecken dienende Marktforschung weniger zu tun; aber auch hier erwies sich zunächst die Meinung vieler Marketing- und Werbeleute als hinderlich: Man wisse aus dem Bauch heraus mehr, als da statistisch errechnet würde. Das verblaßte aber schneller als die reservierten Einstellungen an den Hochschulen, und auch die meist angewandte Forschung der hochschulfreien Institute hatte weniger Schwierigkeiten der Akzeptanz.

Einen Durchbruch in der qualifizierten Öffentlichkeit für die Umfrageforschung als Institution brachten die Wahlforschung und hierbei insbesondere die Wahlvorhersagen sowie die Erklärung von Wahlergebnissen am Abend eines Wahltages im Fernsehen. Das war zunächst von den beiden Nestoren der Umfrageforschung, der journalistisch ausgebildeten Elisabeth Noelle-Neumann und dem Wirtschaftsjournalisten Stackelberg als eine Art PR-Veranstaltung organisiert. Vor dem Wahltag wurden in verschlossenen Umschlägen von beiden Voraussagen bei Fernsehanstalten hinterlegt und dann vor laufenden Kameras mit dem Wahlergebnis verglichen.

Wichtiger war es, daß Forscher die Gründe für ein Wahlergebnis aufgrund quantitativer Erhebungen zu deuten in der Lage waren. Ein wichtiges Beispiel sind die gemeinsamen Fernsehsendungen von Wildenmann und Scheuch bei den Wahlen 1961. In den weiteren

Jahren übernahmen dann Vertreter der privatwirtschaftlich verfaßten Institute diese Aufgabe der Erklärung - wie beispielsweise Tacke von EMNID. Inzwischen sind die entsprechenden Kenntnisse hierfür auch Journalisten vertraut. In dieser Abfolge des Personals für Kommentare - Professoren, Demoskopon, Journalisten - kann ein Prozeß zunehmender Institutionalisierung der Umfrageforschung als Dienstleistung gesehen werden.

Wir haben einmal Institutionalisierung definiert als den Moment, wo ein Vorgehen oder ein Sachverhalt nicht mehr sachlich erklärt werden muß, sondern durch die Benennung als erklärt gilt. Seit Anfang der siebziger Jahre reicht es für die Autorität eines Untersuchungsberichtes aus, wenn gesagt wird, er beruhe auf einer repräsentativen Umfrage.

2. Die siebziger und achtziger Jahre als die Zeit der Konsolidierung und methodischen Differenzierung

Bis in die siebziger Jahre war das Instrumentarium, das in Umfragen verwendet wurde, verhältnismäßig einfach. Es wurden entweder Quotensamples oder Wahrscheinlichkeitsstichproben einfacher Art benutzt, um die Erhebungspersonen zu bestimmen. Von nun an aber werden auch die Stichproben komplizierter, nehmen Schichtungen und bewußte Verzerrungen als Design zu. Als Gemeinschaftsunternehmen wurde seither ein „ADM-Mastersample“ entwickelt, das hohen methodischen Ansprüchen genügt und in vielen Marktforschungen auch praktisch eingesetzt wird. Auffälliger noch ist die Komplizierung der Stimuli. Fragebatterien werden standardisiert, die Anwendung von Skalen so häufig, daß ZUMA ein Skalenhandbuch zusammenstellen kann⁴. Das Institut für Demoskopie ist führend bei der Verwendung von Bildern als Stimuli, wobei sich eine gewisse Vorliebe für längere Satzfolgen in Sprechblasen beobachten läßt.

Einer solchen Komplizierung der Stimuli entspricht auch eine Weiterentwicklung bei der Analyse. Üblich waren bis in die sechziger Jahre lediglich Randauszählungen und Aufgliederungen nach einigen demographischen Merkmalen. Durchweg wurde dabei auf statistische Maßzahlen verzichtet. Das ist zwar auch heute noch bei einigen Instituten kennzeichnend, insbesondere auch beim Institut für Demoskopie. Allgemein aber wurden aussagekräftige Verfahren der Statistik durch die Zugänglichkeit von Computern und

⁴ ZUMA-Handbuch Sozialwissenschaftlicher Skalen. Bonn. Informationszentrum Sozialwissenschaften, Band 1-3, 1983.

mehr noch durch die damit verbundene Entwicklung von Programmpaketen gefördert. Insbesondere in der akademischen Forschung ist die Verwendung von Auswertungspaketen wie SPSS auch für Diplomarbeiten bereits Routine. Nichtparametrische Verfahren wie die Korrespondenzanalyse sind keine exotischen Vorgehensweisen mehr.

Eine dritte Besonderheit, die mit der Etablierung der Forschung in den beiden Jahrzehnten verbunden ist - nach der Komplizierung der Stimuli und der Verfeinerung der Auswertung -, ist die Vielfalt benutzter Verfahren. In allen entwickelten Ländern wuchsen die Schwierigkeiten bei der Anwendung des persönlichen Interviews. Die Ausschöpfungsraten der Stichproben sanken, und die Korrektur durch Gewichtung mußte immer voraussetzungsvoller werden. Die Ausschöpfungsraten sind in verschiedenen Ländern zwar unterschiedlich hoch - besonders niedrig in den USA und relativ hoch in Japan -, aber die Entwicklungsrichtung hin zu immer größeren Schwierigkeiten ist in allen entwickelten Ländern die gleiche. Das führte zu einem Durchbruch der Telefoninterviews, die man vordem nur für geeignet hielt, bei speziellen Zielpopulationen verwandt zu werden - allerdings in den achtziger Jahren durchweg nur in privatwirtschaftlich arbeitenden Instituten. Zunächst war die im Vergleich zu den USA geringere Telefondichte in europäischen Ländern ein Argument gegen Telefoninterviews, aber dieses ist mit der Vollversorgung vieler europäischer Bevölkerungen und der Japans mit Telefon inzwischen überwunden.

Insbesondere in den achtziger Jahren wurde besondere Computer-Software für Telefoninterviews entwickelt, welche die automatische Identifizierung von Befragungspersonen ermöglichen ohne den Umweg über Telefonverzeichnisse, was angesichts der Zunahme von Personen, die einer Eintragung ins Register widersprechen, von Bedeutung ist. Zugleich werden durch entsprechende Software die Interviewer in ihrem Vorgehen kontrolliert, der Umgang mit Filterfragen - eine bis dahin sehr ärgerliche Fehlerquelle - automatisiert und die Aufzeichnungsgenauigkeit erhöht. Spezielle Programme für Telefoninterviews ermöglichen eine sofortige Eingabe der Daten und deren automatische Auswertung. Die apparative Ausrüstung für diese Entwicklungen hatten in Deutschland zunächst nur die kommerziellen Institute; in akademischen Instituten setzte sich diese Entwicklung erst mit zehn Jahren und mehr Verspätung allmählich auch durch.

Selbstverständlich erforderte diese Entwicklung methodische Prüfung. Das Telefoninterview hat keineswegs das persönliche Interview obsolet werden lassen, aber heute wissen wir ziemlich viel über die Eignung des einen oder anderen Verfahrens je nach Thema der Untersuchung. Hier sind die methodologischen Untersuchungen von ZUMA hervorzuheben. Daß dabei gelegentlich das Telefoninterview auch da angewandt wird, wo es trotz seiner Vorzüge von Schnelligkeit und Interview-Kontrolle besser nicht benützt würde, hängt damit zusammen, daß die Kosten für das persönliche Interview explodiert sind.

Quantitative empirische Forschung beschränkt sich selbstverständlich nicht auf Umfragen, aber was eine Umfrage ist, das ist auch vielfältiger geworden. In diesem Zusammenhang sei verwiesen auf die Übernahme der Exit polls in Deutschland: die Befragung von Wählern beim Verlassen des Wahllokals. Hiermit werden erstaunlich hohe Genauigkeiten bei Hochrechnungen in kürzester Zeit nach Abschluß des Wahlvorgangs erzielt, was die Bekanntheit der Umfrageforschung und ihr Ansehen in der Bevölkerung erheblich vergrößert hat.

Auch das Befragen von ausgewählten Personen in Laborsituationen kann noch zur Umfrageforschung gerechnet werden. Als „Baggern“ dient es insbesondere der Ermittlung von Reaktionen gegenüber neuen Produkten.

Ein recht selten benutztes Instrument der Sozialforschung sind „Panels“ geblieben. Zwar erheben Datenschützer keine grundsätzlichen Einwände mehr, aber forschungstechnisch und finanziell sind solche Wiederholungsbefragungen doch zu aufwendig, um mehr als besondere Kraftakte zu sein. Das gilt dann a fortiori für Lebenslaufbeobachtungen, wie sie Carl Gunnar Janson an einer Alterskohorte seit über 35 Jahren vornimmt. In der kommerziellen Forschung sind inzwischen allerdings Panels nicht mehr so selten, wobei sich der Beobachtungszeitraum kaum je über ein Jahr erstreckt - was für die dort üblichen Problemstellungen auch ausreicht.

Umfragen der herkömmlichen Art sind jedoch nach wie vor bestimmend für den ganzen Dienstleistungsbereich und noch stärker für die Wirksamkeit in der qualifizierten Öffentlichkeit. Als Beispiel für die Nützlichkeit, die im Verlaufe des Prozesses der Institutionalisierung der Umfrageforschung zuerkannt wurde, seien Umfragen über extremistisches Verhalten in der Politik hervorgehoben. Durch Umfragen konnte den politisch wichtigen Kreisen aufgezeigt werden, daß der Linksextremismus, der sich in den späten sechziger Jahren an den Hochschulen so lautstark durchsetzte, auf eine Minderheit der dort Tätigen beschränkt blieb. Besonders nervös reagierte das politische Establishment auf das in den sechziger Jahren vorübergehend rapide ansteigende Potential an NPD-Wählern. Wiederum durch Umfragen konnte nachgewiesen werden, daß die Wahl für die NPD überwiegend nicht Ausdruck rechtsextremer Gesinnung, sondern eine Form der Protestwahl bedeutete. Auch auf der Grundlage dieser Befunde konnte davon abgeraten werden, die Wähler der NPD als unverbesserliche Nazis propagandistisch anzugehen und vielmehr in der Öffentlichkeit darauf abzustellen, daß diese politische Gruppierung weder ein diskutables Programm für praktische Arbeit noch ein diskutables Personal für diese Aufgaben hatte. Nicht methodische Überzeugungsarbeit, sondern praktischer Erfolg war entscheidend dafür, daß bis Ende der sechziger Jahre die Umfrageforschung als institutionalisiert gelten konnte.

3. Die Institutionalisierung der Umfrageforschung im akademischen Milieu

Für Auskünfte über die Entwicklung der empirischen Sozialforschung sowohl der hochschulfreien gemeinnützigen Institute wie auch der Hochschulinstiute ist die wichtigste Grundlage die jährliche Befragung durch das Zentralarchiv und das Informationszentrum. Die Angaben der Institute werden dabei ergänzt durch systematische Auswertung von Universitätsjahrbüchern und vergleichbaren Veröffentlichungen sowie durch Auskünfte von forschungsfördernden Institutionen. Seit 1968 erscheint jährlich ein Band mit Beschreibungen der erhobenen Projekte unter dem Titel „Empirische Sozialforschung“⁵. Nach dieser Erhebung sind etwa zwischen dem Jahr 1970 840 und dem Jahr 1990 1.100 Erhebungen an hochschulfreien und Hochschulinstiuten gezählt worden, deren Material in maschinenlesbarem Format aufgezeichnet ist und für Sekundäranalysen verfügbar sein könnte. Die Zahl der insgesamt mitgeteilten Forschungsarbeiten ist wesentlich größer, liegt über 4.000.

Das finanzielle Volumen dieser Forschung ist nicht einfach zu bestimmen, weil an den Hochschulen vom gleichen Personal nebeneinander Forschung, Verwaltung und Lehre erwartet wird. Nach eigenen zugegebenen gewagten Schätzungen, bei denen unterstellt wird, daß etwa ein Drittel der Zeit des Stammpersonals für Forschung verwendet wird, kommen wir auf ein Finanzvolumen von etwas über hundert Millionen DM pro Jahr. Die Ausgaben für die Forschung an den hochschulfreien gemeinnützigen Instiuten dürften nicht darunter liegen, so daß wir auf eine Schätzung des Volumens von um die 200 Millionen DM pro annum kommen. Zwischen 1970 und 1990 ist ein Anstieg von vielleicht 20 Prozent zu beobachten, aber eben kein Umbruch oder bemerkenswerte Expansion.

Vielleicht ist eines der wichtigsten Indizien für die Konsolidierung der Umfrageforschung das Festschreiben der Methodenausbildung als Teil der Magister- und Diplomstudiengänge insbesondere in der Soziologie, zunehmend aber auch in Politologie und Erziehungswissenschaften. Die dort gelehrteten Verfahren sind dabei der praktischen Forschung durchweg um einiges an Verfeinerungen und intellektuellem Anspruch voraus.

Bei den Erhebungstechniken ist die Rangfolge in der Bedeutung der verschiedenen Techniken über einen Beobachtungszeitraum von 20 Jahren bemerkenswert stabil:

⁵ Früher im Verlag Dokumentation Sauer, München, später im Campus-Verlag Frankfurt.

Anteil von Verfahren an quantitativen Erhebungen insgesamt⁶

	1970	1975	1980	1985	1990
Einzelinterview	55%	48%	48%	45%	43%
Telefoninterview	-	-	-	5%	6%
Expertengespräch	-	-	18%	22%	18%
Gruppendiskussion	-	-	13%	13%	9%
Schriftliche Befragung	43%	47%	48%	46%	50%
Beobachtung	16%	21%	17%	15%	10%
Dokumentenanalyse	-	-	24%	28%	22%
Inhaltsanalyse	15%	22%	18%	13%	7%
Tests	18%	19%	11%	7%	7%

Bundesweite Erhebungen sind in dieser großen Zahl von Untersuchungen sehr selten. Bevorzugter Gegenstand der Untersuchung sind räumlich oder sozial spezialisierte Populationen. Dabei nehmen die Versuchsanordnungen im Verlauf der Zeit an Raffinement etwas zu. Insbesondere wird im Zeitverlauf die Auswertung anspruchsvoller. Für letzteres liegt eine Auszählung aus der Erhebung für 1980 vor⁷. Hierbei werden Vorgehensweisen der universitären Einrichtungen mit denen der hochschulfreien Institutionen verglichen. Dies ist das Ergebnis⁸:

	Hochschulfreie Institute	Universitäten
Qualitative Auswertungsverfahren	10%	9%
Quantitative Verfahren:		
-Tabellenanalyse	26%	35%
-Korrelationsverfahren	27%	26%
-Multivariate Verfahren	55%	53%
-Analysepakete	11%	14%
Sonstiges	26%	30%

Zwischen den Fächern gibt es einige Unterschiede. Darunter sind Korrelationsverfahren bei Untersuchungen durch Psychologen etwas häufiger, während Soziologen stärker als Angehörige anderer Fächer Tabellenanalysen vorziehen. Generell ist über die Jahre hin-

⁶ Wir haben die Jahrbücher der erwähnten Erhebung für 1971, 1976, 1981, 1986 und 1991 ausgewertet. Die dort erfaßten Berichte beschreiben überwiegend Tätigkeiten des vorausgegangenen Jahres. Die Jahrbücher selbst erscheinen im Durchschnitt jeweils zwei Jahre nach den empirischen Arbeiten.

⁷ Empirische Sozialforschung 1981, München 1982, S. XXVIII.

⁸ *ibid.*

weg eine Zunahme anspruchsvollerer Verfahren bis hin zu den nichtparametrischen Versionen von Faktorenanalysen, wie der Korrespondenzanalyse, zu beobachten. Der Zugang zu Analysepaketen verleitet allerdings auch zur Verwendung von Statistiken, die ein Messniveau voraussetzen, das meist in den Erhebungen der Sozialwissenschaften nicht erreicht wird.

Die Sozialwissenschaftler haben zwei große Zeitreihenuntersuchungen organisieren können, die zunehmend für den Nachwuchs an Universitäten Grundlage von Examensarbeiten sind. Das ist zunächst der ALLBUS, der als wiederkehrende Erhebung von Franz Urban Pappi angeregt wurde. Er ist eine Kombination von wiederkehrenden Fragemodulen und wechselnden Inhalten, wobei durchweg der Abstand bei der Wiederholung eines Moduls fünf Jahre beträgt. Hinzu gekommen ist das International Social Science Program (ISSP), in dem inzwischen in 33 Ländern ein Modul abgefragt wird, das dann für Sozialforscher in allen 33 Ländern zum internationalen Vergleich verfügbar ist. ISSP enthält ebenfalls ein Element der Replikation und eine veritable Zeitreihe durch die aufwendige Demographie, die für die beteiligten Institute verpflichtend ist. So ist auch in der Sozialforschung die Tendenz weg von der lediglichen Begründung von Aussagen durch Momentaufnahmen hin zu Beobachtungen über Zeitabläufe institutionalisiert.

Nicht für alle Erhebungsjahre wurden die Themen verschlüsselt. Wenn dies geschah, dann war zudem das Kategoriensystem nicht einheitlich. Aber vorwiegend kann eine ziemliche Stabilität in der Rangfolge der verschiedenen Themen behauptet werden.

Themen (Mehrfachnennungen):	1970	1975
Erziehung, Schule	24%	48%
Beruf	19%	29%
Wissenschaft	15%	18%
Arbeit und Betrieb	14%	14%
Medien und öffentliche Meinung	10%	15%
Politik	6%	9 %
Recht	3%	5 %
Wirtschaft	3%	0 %

Wir haben aus der Vielzahl der Kategorien hier nur die wiedergegeben, die besonders häufig gezählt wurden, und jene, die überraschend selten waren. Selbstverständlich gibt es sehr viele weitere Themen, die seltener bearbeitet wurden - wie Religion, Sozialpolitik oder Militär, aber für diese Topoi war das auch zu erwarten. Bemerkenswert ist, daß insbesondere das Thema Politik von den Medien besonders beobachtet und gern wiedergegeben wird - was für die Häufigkeit in dieser Zählung nicht gilt. Es gibt eben eine außer-

ordentliche Diskrepanz zwischen der Wiedergabe der Themen in öffentlichen Medien und ihrer Häufigkeit im Alltag der Forschung.

Das folgt auch daraus, daß ein größerer Teil der Themen der akademischen Höherqualifikation der Beteiligten dient. Im Schnitt gilt ein Fünftel aller Untersuchungen dem Erwerb eines Doktorgrades. In den weitaus meisten Fällen sind Forschungen an hochschulfreien Instituten und insbesondere an Hochschulinstituten Eigenprojekte, geben also das lokale Interesse der akademisch orientierten Forschung an sich selber wieder. Überraschend ist bei diesen Auszählungen, daß die Förderung durch die großen Stiftungen quantitativ von nachrangiger Bedeutung ist. Für das Jahr 1975 wurden von den berichteten Projekten durch die Deutsche Forschungsgemeinschaft neun Prozent und durch die Stiftung Volkswagenwerk drei Prozent finanziert.

Allgemeine Schlußfolgerung: Diese quantitativen Forschungen im akademischen Milieu dienen in erster Linie dem eigenen Milieu, und zwar innerhalb der fachlich und nach Themenbereichen aufgespalteten internen Forschungsöffentlichkeit.

4. Die Institutionalisierung der Umfrageforschung im privatwirtschaftlichen Milieu

Verallgemeinernde Aussagen über Forschung von Instituten mit privatwirtschaftlicher Grundlage sind erst für die neunziger Jahre quantitativ zu begründen. Vorher gab es vereinzelte Erhebungen wie die von Viggo Graf Blücher 1981 und die von Thomas Bausch 1988/89⁹. Die wichtigste Quelle für die eigenen Aussagen sind aber neuere Statistiken und andere Angaben des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute (ADM) im Internet¹⁰. Die Angaben beziehen sich auf die Mitglieder in diesem Fachverband, der jedoch nur 41 der etwa über 80 der ernster zu nehmenden Institute organisiert. Dies sind dann allerdings mit einer Ausnahme die wichtigsten Institute, die zusammen rund 65 Prozent des Umsatzes der Marktforschung und 72 Prozent der Beschäftigten in diesem Dienstleistungsgewerbe repräsentieren.

Nach einer Statistik von ESOMAR wurde gegen Ende unserer Beobachtungsperiode weltweit mit Marktforschung etwa 13 Milliarden DM umgesetzt. Bis Ende der neunziger Jahre dürfte sich dieser Betrag sogar verdoppelt haben. Etwas mehr als 40 Prozent der

⁹ Thomas Bausch: „Stichprobenverfahren in der Marktforschung“. München 1990.

¹⁰ Die Adreses der homepage: www.adm-ev.de.

Umsätze werden in Ländern der Europäischen Union erreicht, die damit der weltweit bedeutendste Markt für diese Dienstleistungen sind. Deutschland nimmt hier eine führende Stellung ein, allerdings im Vergleich zur Bevölkerungszahl unterdurchschnittlich gegenüber England.

Viggo Graf Blücher ermittelte 1981, daß es in der Bundesrepublik 165 Voll-Service-Umfrageinstitute gibt¹¹; demgegenüber rechnet KONTEXT für 1989 mit 117 Instituten, in denen 5.174 Personen fest angestellt sind. Der gesamte Umsatz wird hierbei auf eine Milliarde DM geschätzt. Erstaunlich ist, daß nach Auskunft der ADM-Institute ein Drittel des Umsatzes im Ausland erzielt wird, und zwar überwiegend durch dort ansässige Auftraggeber. Sekundärforschung, qualitative Primäruntersuchungen und sogenannter Desk-Research sind von abnehmender Bedeutung im Vergleich zu quantitativen Primäruntersuchungen. Bei diesen haben Paneluntersuchungen inzwischen einen erstaunlichen Anteil von etwa einem Drittel des gesamten Umsatzes erreicht. Die Mehr-Themen-Erhebungen, an die man sonst bei dem Verweis auf Sozialforschung oft denkt, haben dagegen nur einen Anteil von etwa fünf Prozent am Gesamtumsatz. Das bedeutet: Die Marktforschung geht zunehmend von der Untersuchung als Momentaufnahme zu Zeitreihenbetrachtungen über.

Von besonderer Bedeutung für die sozialwissenschaftlichen Kenntnisse der eigenen Gesellschaft sind große Forschungsreihen, die sich als Gemeinschaftsunternehmen von Forschergruppen und Instituten seit Anfang der siebziger Jahre entwickelt haben. Hier sei insbesondere verwiesen auf die Analyse der Mediennutzung, die bis 1971 mit einer großen Stichprobe als Leseranalyse für die Druckmedien durchgeführt wurden. Die hohe Zahl von Interviews ist notwendig angesichts des Vorherrschens von Regionalzeitungen in der Bundesrepublik. Ab 1972 werden auch die elektronischen Medien berücksichtigt, und die Stichprobengröße wurde angesichts der vielen Regionalsender auf 17.500 Personen erweitert. 1987 wurden aus dieser Gemeinschaftserhebung dann zwei parallele Studien: die Leseranalyse für Druckmedien (LA) und die Medienanalyse (MA) für elektronische Medien. Die Materialien dieser Untersuchungen stehen der Allgemeinheit für Sekundäranalysen zur Verfügung.

Auch der Tourismus, der als Wirtschaftsfaktor immer wichtiger wird, wird durch regelmäßige Erhebungen als Zeitreihe abbildbar. Jahrzehnte hindurch war hierfür der „Arbeits-

¹¹ Dieter Franke und Joachim Scharioth (Hg.): „40 Jahre Markt- und Sozialforschung in der Bundesrepublik“. München 1990, S. 58.

kreis für Tourismus“ verantwortlich, den inzwischen jedoch die großen Tourismusunternehmen nicht mehr tragen. Erfreulicherweise gibt es jedoch eine Nachfolgeinstitution mit eingeschränktem Umfang der Erhebungen.

Auch für Markenartikel gibt es Zeitreihenerhebungen, die allerdings bisher für die Öffentlichkeit nicht in gleicher Weise verfügbar sind. Weitere Erhebungsgegenstände auch mit Zeitreihencharakter zeigen den Umgang der Deutschen mit Geld und Wertpapieren.

Die wichtigste der Zeitreihen für die Sozialwissenschaften dürfte aber die Wahlforschung sein. Beginnend mit der Bundestagswahl 1957 ist inzwischen jede Bundestagswahl durch eine aufwendige Erhebung begleitet worden. Das Zentralarchiv hat hieraus einen als Zeitreihe auswertbaren Datensatz erstellt.

Gegen Ende der beiden Dekaden herrschte das persönliche Interview noch eindeutig vor - aber bei rascher Zunahme des Telefoninterviews. Die Befragung über das Internet spielte damals noch keine und heute eine lediglich außerordentlich untergeordnete Rolle. Es zeichnet sich aber ab, daß in persönlichen Interviews Papier und Bleistift durch tragbare Computer (Lap-Top und Pen-Top) ersetzt wird. Es gab bereits damals bei den privatwirtschaftlichen Instituten über tausend CATI(Computer Assisted Telefon Interviews)- und über 2.000 CAPI(Computer Assisted Personal Interviews)-Geräte im Einsatz.

Wichtigster Auftraggeber der Umfrageforschung für wirtschaftliche Zwecke ist die Konsum- und Gebrauchsgüterindustrie - allerdings mit sinkender Tendenz. Der Anteil der öffentlichen Auftraggeber am Umsatz geht auch angesichts der Kassenlage der öffentlichen Hand zurück, wogegen die Bedeutung der Medien als Klienten erheblich zunimmt. Überraschend ist die geringe Bedeutung der Werbeagenturen für die Auftragslage der Umfrageinstitute.

Ungeachtet des Gesamtvolumens sind Umfragen auf privatwirtschaftlicher Basis eine Dienstleistung auf der Grundlage von Klein- und Mittelbetrieben. Infratest als eines der größten Institute hat etwa 600 Mitarbeiter und einen Umsatz von etwas über 100 Millionen DM. Die Gesellschaft für Konsumforschung dürfte wesentlich größer sein. Die meisten bekannteren Institute haben lediglich zwischen 20 und 60 Angestellte. Der Bekanntheitsgrad hat dabei mit der Größe weniger zu tun und ist in erster Linie eine Folge der Beachtung von Themen durch Medien für ein allgemeines Publikum. So sind in der Öffentlichkeit besonders bekannt das Institut für Demoskopie, das EMNID-Institut, die Forschungsgruppe Wahlen und FORSA wegen ihrer Erhebungen mit politischen Themen. Sie sind alle wirtschaftlich wesentlich weniger bedeutsam als die in den Medien kaum beachteten GETAS und GfK.

Zur Konsolidierung hat entscheidend auch beigetragen, daß sowohl an Hochschulen als auch gefördert durch Berufsverbände wie BVM und ADM regelmäßige Ausbildungsgänge angeboten werden. Hier ist das Wissen in allen entwickelten Ländern inzwischen auf vergleichbarem Niveau. Internationale Vereinigungen wie ESOMAR, WAPOR, CESSDA und IFDO sowie ICPSR haben zu diesem hohen Entwicklungsstand der Ausbildung von Meinungsforschern Wichtiges beigetragen.

Eine Aussage über die Häufigkeit von Themen bei der privatwirtschaftlich verfaßten Umfrageforschung kann sich leider nicht auf Zählungen stützen, wie sie von der ASI veranlaßt wurden. Die Institute behandeln die Themen ihrer Aufträge als Geschäftsgeheimnis. Sicher ist dennoch, daß von Sozialforschung keines der größeren Institute leben könnte. Das wirtschaftliche Wohlergehen hängt von der Marktforschung ab (im weiten Sinn verstanden, also auch unter Einbezug der Medienforschung). Indirekt läßt sich die Bedeutung von Problemstellungen wohl aus den Themen der Schriftenreihen und Seminare erschließen. Für die Seminare sind die Erweiterung des vertrauten Arsenal von Techniken einerseits um Intensiv- und Tiefenuntersuchungen und andererseits um multivariate Verfahren dringlich. Die Fragen des Designs werden in mehreren Kursen behandelt¹². Schließlich ist bei teilweise unkundigen Auftraggebern die Präsentation für den Geschäftserfolg wichtig. Die Internationalisierung der Märkte ist in einer ganzen Reihe von Aufsätzen Thema. In den Schriften wird zudem die Bedeutung der Forschung als Voraussetzung für Planen wiederholt angesprochen. Das impliziert die Einbeziehung der Marktforschung in Entscheidungsprozesse der Auftraggeber. Die dürfte aber keinesfalls durchweg Realität sein, sondern ist als Versuch der Erziehung von Klienten zu bewerten, sollte also spiegelverkehrt gelesen werden.

Mit der Behauptung, hier gehe es um Marktforschung, versuchen schon seit langem viele verschiedene Vertriebsformen sich Zugang zu Haushalten zu verschaffen. Dieses Dauerproblem war besonders virulent beim Zeitschriftenvertrieb. Gegenwärtig tarnt sich auf diese Weise der Versuch von Adressenhändlern, ihre Listen zu erweitern.

Und nicht zuletzt gibt es auch unseriöse Unternehmen mit der Bezeichnung Markt- und Sozialforschung, die Zahlen publizieren, welche auf undurchsichtige, jedenfalls nicht seriöse Weise zustande kommen. Einen besonders bösen Ruf hatte in den achtziger Jahren das Wickert-Institut, dem seriöse Umfrageforscher nachsagten, daß der Institutsleiter am

¹² Bei Gebühren von über 2.000 DM für zwei- bis dreitägige Seminare müssen die Themen der Seminare auf wirkliche Bedürfnisse gezielt sein.

Aufwand für Interviews spare. Über viele Jahre hinweg übernahmen die Medien solch windige Zahlen, wenn sie denn nur aktuell genug waren. Und Aktualität ist bei Verzicht auf seriöse Feldarbeit unschwer zu erreichen. Die ADM versucht dagegen zu halten durch Pressemitteilungen und nicht zuletzt durch „Standards zur Qualitätssicherung in der Markt- und Sozialforschung“. Einigermaßen einheitliche Mindeststandards und ein gehöriger Abstand gegenüber Wünschen, sich für Public-Relations-Zwecke instrumentalisieren zu lassen, bleiben jedoch Dauerprobleme.

5. Sozialforschung als eine gemeinsame Veranstaltung

Die Sozialforschung - auf der Grundlage privatwirtschaftlicher Institute oder gemeinnütziger Einheiten - hat in Deutschland gemeinsam mit mancherlei Hindernissen zu kämpfen. In der Vergangenheit wurde das nicht selten bei dem Versuch deutlich, verschiedene amtliche Register zur Grundlage von Stichprobenziehungen zu nutzen. In den siebziger und achtziger Jahren dürfte hier ein brauchbarer Standard durchgesetzt worden sein. Schwierigkeiten gab es auch für die gemeinnützigen Institute bei der Auslegung von Steuergesetzen. Wäre auf die Förderungsbeträge die Umsatzsteuer angewandt worden, so hätte es eine Anzahl von Forschungen nicht mehr geben können. Die ASI hat hier eine bedeutende Rolle gespielt.

Eine wichtige Grundlage für Forschungspläne sind die Volkszählungen. Hier gab es in den achtziger Jahren große Turbulenzen, wobei Sozialforscher durch Begleitforschungen zur Volkszählung zeigen konnten, daß die Bevölkerung insgesamt diesen amtlichen Erhebungen positiv gegenübersteht¹³. Die bei der Erhebung 1987 erfolgten Konfrontationen und die Politisierung der Volkszählung auch in der SPD hat dann dazu geführt, daß Bundeskanzler Kohl die Durchführung einer Volkszählung, die europaweit für 2001 vorgesehen ist, blockierte. Auch die neue Regierung Schröder hat diese Blockade nicht aufgehoben, wohl eingedenk der Proteste gegen Volkszählung in linken Kreisen. Wenn es überhaupt gelingen sollte, diesen deutschen Sonderweg bei der europäischen Volkszählung zu korrigieren, dann nur durch das gemeinschaftliche Auftreten der akademischen und privatwirtschaftlichen Institute.

¹³ Erwin K. Scheuch, Lorenz Gräf und Steffen Kühnel: „Volkszählung, Volkszählungsprotest und Bürgerverhalten“. Stuttgart 1989.

Eine solche Gemeinsamkeit hat sich insbesondere bisher bei der Entwicklung des Datenschutzes bewährt. Anfang der siebziger Jahre waren akademische Sozialforscher mitentscheidend dafür, daß überhaupt ein Verständnis für das Problem des Datenschutzes angesichts der Entwicklung des Informationswesens entstand. Auf einer Tagung im Oktober 1981 in Heidelberg thematisierten wir, daß durch die Vernetzung von Registern die Anlage von Dossiers über einzelne Bürger möglich werde. Eingeladen hatten erstmals gemeinsam ADM und ASI, was demonstrativ wirken sollte und wirkte. Starredner der Tagung, auf der immerhin auch Ex-Bundespräsident Scheel auftrat, war Karl W. Deutsch¹⁴.

In verschiedenen Ländern, so auch in der Bundesrepublik, wurde damals die Einführung eines allgemeinen „Personenkennzeichens“ vorbereitet, mit dem die verschiedenen Register mit Angaben über Einzelpersonen verknüpft werden sollten; in den USA geschieht dies vermittels der Sozialversicherungsnummer. Wir argumentierten, daß dieser Transport von Daten aus einem Register in ein anderes durch Datenschutz zu regeln sei. Es ging hier nicht so sehr um Vertraulichkeit für ein besonderes einzelnes Merkmal; das ist durch entsprechende Gesetze ohnehin geregelt. Vielmehr ging es um die Vernetzung von Daten. Dies wurde mit Juristen ausführlich besprochen, und das schließliche Datenschutzgesetz des Bundes (BDSG) wurde so zu einem Gesetz zur Regelung des Datentransportes. Dies steht im Gegensatz zu dem Ansatz im französischen Datenschutzgesetz, dessen Kern eine Aufzählung schützenswürdiger Einzelinformationen ist, was bei gesetzeskonformer Auslegung zu erheblichen Schwierigkeiten für die Sozialforschung führen könnte.

Wie im Rechtswesen verbreitet ergeben sich auch bei der Anwendung des BDSG durch untere Instanzen Probleme. Hier entscheiden nachgeordnete Stellen nach dem in Deutschland bei unteren Behörden üblichen Prinzip: „Wenn Du im Zweifel bist, sage Nein; denn dann machst Du keinen Fehler“. Bisher hat das gemeinsame Auftreten von ASI und ADM Schaden begrenzen können.

Schwierigkeiten gab es zunächst für Panelerhebungen, was angesichts der bereits erwähnten steigenden Bedeutung dieser Vorgehensweise eine nachhaltige Beeinträchtigung unseres Wissens über die Gesellschaft zur Folge haben konnte. Es gelang jedoch, mit den Datenschutzbeauftragten der Länder eine Konvention zu vereinbaren, die sowohl dem Datenschutz als auch den Bedürfnissen der Forschung Rechnung trägt. Als Sozialforscher sind wir ja an der konkreten Einzelperson nicht interessiert; sie ist für uns ein „Fall“ unter

¹⁴ Max Kaase, Werner Ott und Erwin K. Scheuch (Hg.): „Empirische Sozialforschung in der modernen Gesellschaft“. Frankfurt a. M. 1983.

Fällen. Entsprechend brauchen wir nicht das Wissen über die einzelne Person, die in verschiedenen Registern erfaßt ist, sondern lediglich die Gewißheit, daß wir es bei einer Panelerhebung im Zeitablauf mit der gleichen Person zu tun haben. Das kann erreicht werden durch eine „funktionale Trennung“: die Verteilung des Wissens über eine Person auf denjenigen, der die Identitäten qua Fallnummer kennt, dann dabei allerdings keinen Zugang zu den Antworten dieser Person hat; und umgekehrt.

Der Datenschutz ist eines der Probleme, die immer wieder in verschiedener Gestalt als Zweifel an der Zulässigkeit von Forschung auftreten. Über die Unterschiede in den rechtlichen Verfassungen und wirtschaftlichen Grundlagen hinweg ist die Gemeinsamkeit der Forschungsstellen, die auf unterschiedlicher Rechtsstellung alle Umfrageforschung als Informationsmittel benutzen, während der siebziger und achtziger Jahre durch Kongresse öffentlich deutlich gemacht worden, die von der ADM und der ASI gemeinsam betrieben wurden. Die Tagung des Jahres 1990: „40 Jahre Markt- und Sozialforschung in der Bundesrepublik“ war hier als besonderes Signal gedacht.

5. Ausblick

Mit der Vereinigung der beiden deutschen Teilstaaten ändert sich vieles für die Umfrageforschung. Nicht zuletzt stellt sich die Frage, was denn ein angemessener Sampling Frame ist. Wie selbstverständlich ist es in der Umfrageforschung üblich, die nationalen Grenzen auch als Definition des Kollektivs zu wählen, für das man Verallgemeinerungen begründen will. Aus pragmatischer Einsicht entschied man sich mit der Wiedervereinigung, bis auf weiteres die beiden deutschen Teilstaaten als separate Kollektive zu behandeln. Es ist Routine für westdeutsche Gebiete, eine Stichprobe des Umfangs 2.000 und für die ostdeutschen eine zweite des Umfangs 1.000 zu ziehen. Das soll erst anders werden, wenn die Unterschiede zwischen beiden Teilen des nun staatlich vereinigten Landes geringer geworden sind.

Aber ist es wirklich angemessen, die Bundesrepublik über die verschiedensten Themen hinweg als das Kollektiv zu behandeln, auf das Verallgemeinerungen auszurichten sind? Vielleicht gibt es Familienformen, die für Landesteile spezifisch sind? So argumentierte früher René König, daß es in Deutschland eine andere Familientradition in den Alpengebieten gebe als außerhalb derselben. Dies ist nicht der Ort, ein solches Grundsatzproblem weiter zu spezifizieren. Die staatliche Vereinigung Deutschlands im Jahre 1990 ist allerdings das Ereignis, das ein grundsätzliches Nachdenken hierüber notwendig macht.

Korrespondenzadresse

*Prof. Dr. Erwin K. Scheuch
Kölner Gesellschaft für Sozialforschung e.V.
Lillencronsraße 6
50931 Köln-Lindenthal*

SOZIALBERICHTERSTATTUNG IN DER BUNDESREPUBLIK DEUTSCHLAND¹

WOLFGANG ZAPF

Das mir gestellte Thema habe ich leicht verändert: „Sozialberichterstattung in der Bundesrepublik“ statt „Sozialindikatorenforschung“, um auch die Leistungen der ZUMA-Abteilungen Mikrodaten und ALLBUS und andere Aktivitäten in GESIS zu würdigen, wie es ja auch der Wissenschaftsrat in seiner Evaluation vor zwei Jahren getan hat. Sozialberichterstattung ist die aktive Form der in der GESIS-Satzung festgeschriebenen Aufgabe der gesellschaftlichen Dauerbeobachtung: Man tut nicht nur Gutes, sondern berichtet und streitet auch darüber. Sozialberichterstattung mißt den sozialstrukturellen Wandel in zentralen Lebensbereichen und bewertet ihn anhand von Zielen, wie z.B. den Verfassungsgrundsatz der Schaffung „gleichwertiger Lebensverhältnisse“ in allen Teilen der Bundesrepublik. Gemäß diesem Ziel, das in den letzten Monaten die Bewertung der deutschen Vereinigung beherrscht hat, haben sich die materiellen Lebensbedingungen in Ost- und Westdeutschland deutlich angeglichen, das subjektive Wohlbefinden hat sich leicht verbessert, aber in zentralen Werteinstellungen und Bewertungen von Institutionen (Freiheit versus Gleichheit, Vertrauen in die Demokratie, die soziale Sicherheit und die öffentliche Sicherheit) bleiben erhebliche Differenzen. Zu diesen Themen hat ZUMA jüngst wichtige Beiträge geleistet und u.a. im ISI (dem sehr gelungenen Informationsdienst Soziale Indikatoren) veröffentlicht (vgl. Habich u.a. 1999).

Im folgenden möchte ich – auf einer etwas allgemeineren Ebene – zwei Themen ansprechen: erstens die Entwicklung wichtiger Wohlfahrtsbegriffe, d.h. von Wohlfahrtszielen und Wohlfahrtsmaßen, auf der theoretischen Seite (I) und zweitens die Entwicklung von Instrumenten der Sozialberichterstattung auf der praktischen Seite (II). Dabei will ich einen Zeitraum von etwa 30 Jahren behandeln, aber nicht länger als 30 Minuten sprechen.

¹ Vortrag zum 25. Geburtstag des Zentrums für Umfragen, Methoden und Analysen (ZUMA), Mannheim, am 4.11.1999. Der Text beruht in Teilen auf einer längeren Fassung in Zapf 1999.

I. Wohlfahrtsbegriffe

Die theoretische Diskussion um einen erweiterten Wohlfahrtsbegriff konzentrierte sich um 1970 auf die Konzepte der *Lebensqualität* und des *qualitativen Wachstums*. Parallel zu dem in der Wirtschaftswissenschaft entwickelten Programm des qualitativen Wachstums wurde für das politische System die Steuerung einer *aktiven Gesellschaftspolitik*, gesamtgesellschaftlich eine *aktive Gesellschaft* (Etzioni 1968) entworfen, mit einer erhöhten Beteiligung der Betroffenen, die aus ihren bloß passiven Rollen als Konsumenten, Sparer und Wähler heraustreten.

Das nächste Thema war *privater Reichtum und öffentliche Armut* im Sinne der schon früher von K. Galbraith geäußerten Kritik am künstlich erzeugten Bedarf des privaten Konsums auf Kosten öffentlicher Einrichtungen. Diese Kritik wurde um 1970 vorgetragen in der Theorie *öffentlicher Güter* (bis hin zu meritorischen, d.h. staatlich auferlegten Gütern) und der staatlich vorlaufenden Entwicklung der *Infrastruktur*. Als Mechanismen der Unterversorgung wurden u.a. die kurzfristigen Zeitpräferenzen der demokratischen Politik, der „Trittbrettfahreffekt“ im Verhandlungssystem der Großgruppen und der Ressortegoismus der Bürokratie ermittelt.

Obwohl die Ölkrisen von 1973 und 1979 die schlimmsten Befürchtungen der Wachstumskritiker zu bestätigen schienen, ist die Entwicklung in mehreren Schüben in andere Verzweigungen gegangen, die natürlich auch die Diskussionen um Wohlfahrt und Lebensqualität beeinflusst haben. Die Ölkrisen erwiesen sich als politisch produziert, nicht als die Erschöpfung von Ressourcen, und sie konnten marktwirtschaftlich überwunden werden.

Das weitere Wachstum der OECD-Länder hat die Stagnationstheorien in den Hintergrund treten lassen. Mit der Wirtschaftsentwicklung wuchs erneut der private Wohlstand und die vielfältigen Prozesse, die als *Individualisierung* und *Pluralisierung* die industriegesellschaftliche Klassenstruktur verändert haben. Der Ausbau der materiellen Infrastruktur, des Bildungssystems, insgesamt des Wohlfahrtsstaates ließen die Kritiker an der öffentlichen Armut verstummen oder dieses Phänomen als ein amerikanisches Sonderproblem erscheinen. Ökologische Innovationen erlaubten die Vorstellung, die „ökologische Frage“ könne schrittweise ebenso erfolgreich geregelt werden wie früher schon die „soziale Frage“.

Nach wie vor gab es eine intensive Diskussion und eine umfangreiche Forschung zu Fragen der Lebensqualität. Aber in gewisser Weise hatte sich das Konzept selbst individualisiert. Lebensqualität wurde zunehmend als *individuelle Wohlfahrt* (well-being) konzipiert (vgl. Argyle 1996), die wesentlich nicht nur von guten *objektiven Lebensbedingungen*, sondern auch vom *subjektiven Wohlbefinden* bestimmt ist. „Lebensqualität muß

letztlich von den Betroffenen wahrgenommen werden“ (is in the eye of the beholder, A. Campbell 1972: 422).

Ich will allerdings kein allzu idyllisches Bild der Wohlfahrtsforschung der 1980er Jahre zeichnen. Die Vorstellung, nach den Problemen des *Lebensstandards* seien auch die Probleme der *Lebensqualität* ihrer Lösung nahe, und die postmoderne Vielfalt der *Lebensstile* nähme nun eine Schlüsselrolle in der Wohlfahrtsforschung wie auch in der Ungleichheitsforschung ein, hat sich nicht allgemein durchgesetzt, und zu Beginn der 1980er Jahre hat es eine heftige Auseinandersetzung um die „Krise des Wohlfahrtsstaats“, die „neue Armut“ und die „Sozialdemontage“ gegeben, die zum Teil dieselben Begriffe und Argumente verwendet hat, wie wir sie heute wieder hören, obwohl sich z.B. von 1980 bis 1990 die Sozialausgaben real (1985=100) um fast 100 Milliarden DM, das sind 15 Prozent erhöht haben.

Heute befinden wir uns erneut inmitten heftiger Kontroversen um die Wohlfahrt der Bürger und um die „soziale Gerechtigkeit“. Ganz offensichtlich ist der Zusammenhang, daß Verteilungsfragen um so wichtiger werden, je unsicherer Gewinne aus Wachstum sind. Diese Kontroversen oder Diskurse werden in mehreren Zusammenhängen geführt, so z.B. um die Themen weitergehende Modernisierung versus zweite Moderne, Ende des Wohlfahrtsstaates, Standort- und Innovationskrise.

Was sind die neuen, heute diskutierten Wohlfahrtskonzepte und Wohlfahrtsmaßstäbe? *Nachhaltige Entwicklung* wie auch *Human Development* werden als weltweite Programme auf der Ebene der Vereinten Nationen propagiert. Dabei ist nachhaltige Entwicklung eine modernisierte Version des qualitativen Wachstums, die von dem Problem ausgeht, daß wir national und international weiterhin Wachstum brauchen, daß dieses aber so gestaltet sein muß, daß die Lebensgrundlagen künftiger Generationen erhalten bleiben. Das Konzept kann auch auf subnationaler Ebene (z.B. für Firmen und Organisationen) angewendet werden. Gegenbegriffe wären Nullwachstum, Verzicht, Askese, Umverteilung.

Human Development hat durch die jährlichen Reports des United Nations Development Program eine große Verbreitung erhalten. Es ist gegenwärtig das anspruchsvollste und weitreichendste Wohlfahrtskonzept. Grundidee ist die Steigerung der Wahlmöglichkeiten des einzelnen, der Qualität des menschlichen Lebens. Das geht über die Entwicklung der Humanressourcen und wohlfahrtsstaatlichen Programme hinaus, die die Menschen eher als Empfänger denn als Gestalter von Leistungen betrachten. Theoretisch umfaßt Human Development deshalb neben *Produktivität*, *Gleichheit* und *Nachhaltigkeit* auch *Gestaltungsmöglichkeit* (empowerment). Empirisch werden (1997) alle 174 Nationen der Welt mit Hilfe eines Human Development Index vermessen. Die Human Development Reports und der HDI sind bei aller Kritik die Einlösung von zwei zentralen Forderungen der

Sozialberichterstattung, nämlich der internationale Vergleich und die Präsentation eines zusammenfassenden Indikators jenseits des Bruttosozialprodukts. Die Bundesrepublik lag übrigens 1997 auf Platz 14 hinter 13 anderen OECD-Ländern.

Auf der Mikroebene verfolgt das Konzept des *Personal Development* ein ähnliches Ziel (vgl. Lane 1996), nämlich Lebensbedingungen und subjektives Wohlbefinden mit Gestaltungs- und Wachstumsmöglichkeiten (Aneignungsqualitäten) zu verbinden. Aber die Mikroperspektive muß mit der Makroperspektive zusammengebracht werden, insofern *personal development* an eine lebenswerte Gesellschaft - das Konzept der *Liveability* (vgl. Veenhoven 1997) - gebunden ist. In diesem Konzept geht es um die wichtigen Perspektiven des Institutionenvertrauens, der Zukunftszuversicht und der wahrgenommenen Sicherheit, ohne die auch gute Lebensbedingungen nicht stabil in subjektives Wohlbefinden umgesetzt werden können. Dies mag – wie eingangs erwähnt – manche Probleme im gegenwärtigen Deutschland, vor allem in Ostdeutschland, erklären.

Das Thema *Exklusion* bezeichnet innerhalb der Europäischen Union die aktuellsten Wohlfahrtsprobleme und die moderne Form von Ungleichheit: nicht mehr nur Abstand, sondern Ausschluß vom gesellschaftlichen Leben, nicht mehr nur Ausschluß durch Armut und Arbeitslosigkeit (d.h. durch Ressourcendefizit), sondern Ausschluß auch durch Diskriminierung und Verweigerung von Zugangschancen (vgl. Silver 1994).

Exklusion zerstört die Solidarität einer Gesellschaft, die Fairnessgrundlagen von Wettbewerb, die politischen Grundrechte auf Teilnahme und Integration. *Solidarität* und *Integration* können um so besser bewahrt, ausgebaut oder wiedergewonnen werden, je größer das *soziale Kapital* einer Gesellschaft ist, d.h. der Bestand an geteilten Normen und Werten, personalen Hilfebeziehungen und sozialen Netzwerken. Während *soziales Kapital* in der ursprünglichen Fassung (bei Bourdieu) als ein prinzipiell knappes und ungleich verteiltes Gut behandelt wurde, nämlich als „soziale Beziehungen“ im engen Sinn von vorteilsstiftenden Kontakten, wird es in den neueren Ansätzen (Coleman, Putnam) als prinzipiell vermehrbar und verteilbar dargestellt. Soziales Kapital erhöht zugleich die Kompetenz, soziale Probleme besser zu lösen: „...Ziviles Engagement und enge Sozialkontakte produzieren positive Ergebnisse - bessere Schulen, schnellere ökonomische Entwicklung, niedrigere Kriminalität, effektivere Verwaltung“ (Putnam 1995: 65f.).

Auf die Lösung eines ganz herausragenden Problems, nämlich der finanziellen und institutionellen Engpässe des Wohlfahrtsstaats, d.h. des Systems der sozialen Sicherung, konzentrieren sich die Beiträge zum *Wohlfahrtspluralismus* (vgl. Evers/Olk 1996). Ihr Grundgedanke ist, daß soziale Sicherheit nicht mehr allein von Markt und Staat garantiert werden kann, sondern daß dazu auch die Ressourcen der Zivilgesellschaft (das soziale Kapital, s.o.) aktiviert werden müssen, also die großen gesellschaftlichen Organisationen

ebenso wie die lokalen Gemeinden, Nachbarschaften und Familien. Durch Kombination der Leistungen dieser verschiedenen Institutionen, durch ein „portfolio of economies“ (R. Rose), soll die Überlastung einzelner Institutionen verhindert werden. Das praktische Beispiel ist heute die auf mehreren Säulen beruhende Alterssicherung.

II. Instrumente der Sozialberichterstattung

In der Anfangsphase der Sozialberichterstattung um 1970 waren die folgenden Themen aktuell - und sie sind es zum Teil heute noch:

- Ausbau der amtlichen Sozialstatistik
- Ausbau der Umfrageforschung
- Analyse nationaler Zielsetzungen (national goals accounting)
- Sozialprognosen und gesamtgesellschaftliche Simulationssysteme
- Soziale Gesamtrechnungen

Demgegenüber war schon zu Beginn der 1970er Jahre bezweifelt worden, daß zwei andere Ziele mit der Sozialberichterstattung erreicht werden könnten, nämlich die wissenschaftliche *Ermittlung gesamtgesellschaftlicher Prioritäten* und die überzeugende *Evaluierung sozialpolitischer Interventionsprogramme*.

Die praktischsten Erfolge wurden seinerzeit bei der Entwicklung von Sozialindikatoren und Sozialberichten erzielt. *Sozialindikatoren* sind Statistiken, die sich durch mehrere Eigenschaften von üblichen Statistiken unterscheiden: sie sollen Leistungen messen, nicht Aufwendungen; sie sollen sich im wesentlichen auf die Wohlfahrt der Bürger beziehen, nicht auf die Aktivitäten von Behörden; sie sollen über Wandlungsprozesse informieren, das heißt in Form von Zeitreihen vorgelegt werden; sie sollen in einem theoretischen Kontext stehen, das heißt über ihre kausale Beziehung zum Indikaturniveau Klarheit herrschen. Von einigen Autoren wurde darüber hinaus die Entwicklung hoch-aggregierter Indikatoren propagiert, z.B. das Verhältnis von total learning force/total labour force.

Damals haben wir in Frankfurt und Mannheim das sogenannte SPES-Indikatoren-System entwickelt (vgl. Zapf 1978), das bis heute hier im ZUMA – inzwischen in einem international-vergleichenden EU-Projekt – fortgeführt wird und mehrere hundert Zeitreihen, im Idealfall von 1950-1999, umfaßt (vgl. Noll 1995).

Sozialberichte sind gesellschaftspolitische Analysen mit der klaren Fragestellung, ob sich die objektiven Lebensbedingungen und das subjektive Wohlbefinden, und über individuelle Dimensionen hinaus, ob sich die Qualität der Gesellschaft verbessert haben. Das klassische Vorbild ist der amerikanische Regierungsbericht *Toward a Social Report*

(1969), der von Mancur Olson organisiert worden ist. „Unsere Idealvorstellung ist, daß soziale Indikatoren messen und daß ein Sozialbericht bewertet, was die volkswirtschaftliche Gesamtrechnung ausläßt“ (Olson 1969: 86/87). Dieser ambitionierte Bericht wurde nicht wiederholt, aber in mehreren Ländern wurden regelmäßige Sozialberichte etabliert. Die dichteste Folge liefern die seit 1970 erscheinenden *Social Trends* aus Großbritannien; mehrfach erschienen sind seit 1973 die französischen *Données sociales*, seit 1974 die *Social and Cultural Reports* der Niederlande, in den USA (dreimal, dann aber eingestellt) die *Social Trends in the United States*. Viermal sind in der Bundesrepublik *Materialien zum Bericht zur Lage der Nation* erschienen.

Der Höhepunkt der Sozialberichterstattung schien Ende der 1970er Jahre überschritten zu sein, jedenfalls trat in den 1980er Jahren eine spürbare Flaute ein. Einige Autoren haben von der Veralltäglicung der Sozialberichterstattung in dem Sinn gesprochen, daß deren Ideen in den verschiedensten Einzeldisziplinen und Bindestrich-Soziologien umgesetzt worden sind, so daß es keiner besonderen „Sozialindikatorenbewegung“ mehr bedurfte: Auflösung durch Diffusion (vgl. Glatzer, in: J. Public Policy 1988). Von heute aus gesehen haben sich aber an anderen Stellen eine Reihe von vielversprechenden Entwicklungen ergeben, die dazu führten, daß die Sozialberichterstattung neuen Schwung erlangt und inzwischen beachtliche Erfolge erzielt. Heinz-Herbert Noll und ich haben in einem Artikel für das letzte ZUMA-Jubiläum 1994 einige dieser Entwicklungen zusammengefaßt. Unter den wichtigen Entwicklungen notieren wir zunächst die enorme Verbreitung von nationalen *Sozialberichten*, also Berichten auf Länderebene, die in der Regel von Ministerien, Planungsbehörden oder dem Statistischen Zentralamt herausgegeben werden. Die Ideen von Social Trends und Données Sociales haben sich also durchgesetzt. In der Bundesrepublik wird der Datenreport in zweijährigem Abstand vom Statistischen Bundesamt zusammen mit Sozialwissenschaftlern vom ZUMA und vom WZB herausgegeben, die mit Umfragedaten den zweiten Teil „Objektive Lebensbedingungen und subjektives Wohlbefinden“ bestreiten (vgl. Datenreport 1997).

Eine zweite Neubelebung ergibt sich auf der *supranationalen Ebene* internationaler Organisation, z.B. in den World Development Reports und in den Social Indicators of Development der World Bank, in den Publikationen von Eurostat und seit 1990, jährlich und viel beachtet, in den Human Development Reports des UNDP-Programmes. Auf der anderen Seite können wir vielfältige Aktivitäten auf *regionaler und lokaler Ebene* beobachten, z.B. statistische Kompendien für Regionen und Städte bzw. Sozialberichte und Entwicklungsberichte für diese subnationalen Einheiten.

Ein Bereich, in dem die Sozialberichterstattung in den letzten Jahren besonders deutliche Fortschritte gemacht hat, sind die auf *spezifische Bevölkerungsgruppen und spezifische*

Lebens- und Politikbereiche ausgerichteten Berichtssysteme. Wiederum H.-H. Noll vom ZUMA hat in einem Sammelband (1997) dieses Feld für die Bundesrepublik dargestellt.

In unserer Zwischenbilanz (Noll/Zapf 1994: 11-13) kamen wir zu einer Reihe von aktuellen Entwicklungstendenzen, die auch noch heute aktuell sind.

- Wichtig wird mittelfristig die *Sozialberichterstattung über die Transformation* derjenigen Länder bleiben, die einen Systemwechsel von Diktatur zu Demokratie und von Kommandowirtschaft zur Marktwirtschaft – zumeist mit großen Mühen – durchmachen.
- Zentral ist die Stärkung der *international vergleichenden Sozialberichterstattung*, wie sie schon früh in den 1970er Jahren von der OECD angestoßen wurde und wie sie heute sowohl amtlicherseits (z.B. von Eurostat) als auch von der Wissenschaft betrieben wird; ZUMA und GESIS spielen dabei eine führende Rolle. Das Eureporting-Projekt vom ZUMA und das Euromodul-Projekt des WZB (zusammen mit ZUMA) möchte ich hier ausdrücklich nennen.
- Methodisch sind die *Längsschnittanalysen* (Panelstudien, retrospektive Lebensverlaufsbefragungen, begleitende Entwicklungsstudien) Ressourcen, die die Sozialberichterstattung schon heute auf ein neues Niveau gebracht haben und deren Potential noch lange nicht ausgeschöpft ist. Neu hinzugekommen sind die deutlich verbesserten Beziehungen zur amtlichen Statistik, die ihre großen Datenbestände (Mikrozensus, Einkommens- und Verbrauchsstichprobe u.a.) benutzerfreundlich und kostengünstig zur Verfügung stellt. Die ZUMA-Abteilung Mikrodaten hat dazu einen wesentlichen Beitrag geleistet.
- Methodisch und technisch stehen nach wie vor die Probleme auf der Agenda, überzeugende *synthetische (Wohlfahrts-)Indikatoren* und *Verfahren einer sozialen Gesamtrechnung* zu entwickeln. Insider erzählen, daß Jacques Delors immer wieder gefordert habe: Gebt mir zwei, drei zusammenfassende Indikatoren, mit denen ich das Bruttosozialprodukt korrigieren kann.
- Noch immer ungeküßt ist die Idee einer *prospektiven Sozialberichterstattung*. Dies meint Versuche, die verstreuten prognostischen Verfahren wie Szenariumstechniken und Projektionsmodelle, systematisch auszubauen, um neben die Retrospektiv- und Status quo-Beschreibung auch kontrollierte Prognosen oder Vergleiche alternativer Zukunftsentwürfe stellen zu können.

Meine Schlußbemerkung gilt der Beobachtung, daß die Sozialberichterstattung – und das ist auch gut so – heute mitten in der gesellschaftspolitischen Auseinandersetzung steht: zwischen Bewahrung und Umbau unserer wohlhabenden modernen Gesellschaften. Auf der einen Seite wird gefordert, die Sozialberichterstattung habe auf die Gefahren hinzu-

weisen, die bei einem Abbau des Wohlfahrtsstaats drohen: Ungleichheit, Unterversorgung, Exklusion und Desintegration (so Joachim Vogel 1999, unser Freund von Statistics Sweden). Auf der anderen Seite sind *Ranking* und *Benchmarking* Konzepte, die zentrale Anliegen der Sozialberichterstattung praktisch werden lassen: den Leistungsvergleich und die Suche nach „der besten Praxis“ (vgl. Land 1996; Oliver 1996). Kritiker weisen diese Ansätze mit Gründen zurück, hier werde die in den 1970er Jahren (und natürlich auch im Sozialismus) gescheiterte Idee der gesellschaftlichen Steuerung und Planung wiederbelebt: als hänge es von den Machern und den Organisationen ab, was man erreichen könne und nicht von Sozialstruktur und langfristigen „Entwicklungspfaden“.

Ich halte es mit Richard Rose's „Lesson-Drawing“ (1993), im Untertitel: Learning across Time and Space. Man kann durch Vergleich manche Fehler vermeiden. Man kann manche wertvollen Ideen und institutionellen Regelungen übernehmen. Zahlreiche Organisationen, Institutionen, Verfahren sind *nicht* übertragbar. Zahlreiche Organisationen, Institutionen, Verfahren sind sehr unterschiedlich, erreichen aber ähnliche Ergebnisse. Diese Lektionen zu beherzigen, ist ein guter Ratschlag für die Sozialberichterstattung. Und übrigens auch für ZUMA, zu dessen 25. Geburtstag ich herzlich gratuliere.

Korrespondenzadresse

*Prof. Dr. Wolfgang Zapf
Wissenschaftszentrum Berlin für Sozialforschung
Reichpietschufer 50, 10785 Berlin
Tel.: 030/25491-384, Fax 030/25491-360*

Literatur

Argyle, M., 1996: Subjective Well-being. S. 18-45 in: Avner Offer (Hg.), In Pursuit of the Quality of Life. Oxford: Oxford University Press.

Campbell, A., 1972: Aspiration, Satisfaction and Fulfillment. S. 441-466 in: Campbell Angus/Philip Converse (Hg.), The Human Meaning of Social Change. New York: Russell Sage.

Datenreport, Zahlen und Fakten über die Bundesrepublik Deutschland, Statistisches Bundesamt (Hg.), 1994; 1997.

Etzioni, A., 1968: The Active Society. New York: The Free Press.

Evers, A./T. Olk (Hg.), 1996: Wohlfahrtspluralismus. Vom Wohlfahrtsstaat zur Wohlfahrts-gesellschaft. Opladen: Westdeutscher Verlag.

- Habich, R./Noll, H.-H./Zapf, W., 1999: Subjektives Wohlbefinden in Ostdeutschland nähert sich westdeutschem Niveau, Informationsdienst Soziale Indikatoren (ISI) 22: 1-6. ZUMA Mannheim.
- Human Development Report, 1990ff.: United Nations Human Development Programme. New York (jährlich seit 1990).
- Journal of Public Policy, 9(4), 1988: Symposium: Whatever Happened to Social Indicators?, darin Kommentar von Wolfgang Glatzer.
- Land, K. C., 1996: Social Indicators and the Quality of Life: Where do we stand in the Mid-1990s? In: SINET, February 1990: 5-8.
- Lane, R. E., 1996: Quality of life and quality of persons: A new role for government? S. 256-293 in: Avner Offer (Hg.), In Pursuit of the Quality of Life. Oxford: Oxford University Press.
- Noll, H.-H., 1995: The Digital Information System Social Indicators: A New Form of Presentation of the German System of Social Indicators. Statistical Journal of the United Nations, ECE 12: 369-378.
- Noll, H.-H., (Hg.), 1997: Sozialberichterstattung in Deutschland, Weinheim/München: Juventa Verlag.
- Noll, H.-H./W. Zapf, 1994: Social Indicators Research: Societal Monitoring and Social Reporting. S. 1-16 in: I. Borg/P. Mohler (Hg.), Trends and Perspectives in Empirical Social Research. Berlin/New York: W. de Gruyter.
- Oliver, N., 1996: Art. Benchmarking. S. 359-364 in: International Encyclopedia of Business and Management. Vol. 1, London: Routledge.
- Olson, M., 1969: Toward a Social Report, US Department of Health, Education and Welfare. Washington.
- Putnam, R. D., 1995: Bowling Alone: America's Declining Social Capital. Journal of Democracy, 6,1: 65-78.
- Rose, R., 1993: Lesson-Drawing in Public Policy. A Guide to Learning Across Time and Space, Chatham, NJ.
- Silver, H., 1994: Social Exclusion and Social Solidarity: Three Paradigms. International Labour Review, 133, No. 5-6: 531-578.
- Veenhoven, R., 1997: Lebenszufriedenheit der Bürger: ein Indikator für die 'Lebbarkeit' von Gesellschaften? S. 267-293 in: H.-H. Noll (Hg.), Sozialberichterstattung in Deutschland. Weinheim/München: Juventa Verlag.
- Vogel, J., 1997: The Future Direction of Social Indicators Research. Social Indicators Research, 42: 103-116.
- Zapf, W., 1978: Angewandte Sozialberichterstattung: Das SPES-Indikatorensystem. S. 689-716 in: Schriften des Vereins für Sozialpolitik. N.F. Band 98, Berlin.

Zapf, W., 1999: Gesellschaftliche Wohlfahrt und Sozialberichterstattung. Erfahrungen und Möglichkeiten. S. 46-60 in: J. Gerhards/R. Hitzler (Hg.), *Eigenwilligkeit und Rationalität sozialer Prozesse*. Festschrift zum 65. Geburtstag von Friedhelm Neidhardt. Westdeutscher Verlag: Opladen.

EMPIRISCHE SOZIALFORSCHUNG IN DEN NEUNZIGER JAHREN¹

KARL ULRICH MAYER

Liebe Mitarbeiterinnen und Mitarbeiter von ZUMA, meine sehr verehrten Damen und Herren!

1. Von Max Planck stammt der Satz, daß wissenschaftliche Theorien nicht widerlegt würden, sondern mit dem Tod ihres Vertreters ausstürben. Wendet man dieses Axiom auch auf die empirische Sozialforschung an, so bin ich denkbar ungeeignet, über die empirische Sozialforschung der neunziger Jahre zu sprechen, da meine eigene Generation Sozialforschung aus der Zeit stammt, in der ich sie gelernt habe, also meiner Zeit bei ZUMA und davor in den siebziger Jahren. In diesem Haus über die empirische Sozialforschung der neunziger Jahre zu sprechen, ist also mehrfach riskant. Zum einen heißt es natürlich Eulen nach Athen tragen. Hier sitzen die Experten und wir sind die Kunden, die, wenn sie klug sind, sich von ZUMA Rat erbitten, und dies lassen, wenn sie weniger klug sind. Es besteht auch die Gefahr, daß ich Ihnen nur meine Ignoranz vorführe oder aus meinem eigenen, zwangsläufig sehr eng umgrenzten Forschungserfahrungshorizont berichte. Bevor ich das wage, werde ich zunächst eine etwas objektivere Methode wählen. Da ich aus eigener Erfahrung nur bedingt über das mir gestellte Thema sprechen kann, werde ich das machen, was jeder ordentliche Sozialforscher mit dem ihm Unbekannten macht: er beforscht es.

2. Ich habe mir drei Quellen herausgesucht, um zu beantworten, ob und was in den neunziger Jahren neu und anders war:

- neuere Auflagen von Lehrbüchern der Methoden der empirischen Sozialforschung im Vergleich mit denen vor 25 Jahren;
- die ZUMA-Nachrichten aus den neunziger Jahren im Vergleich zu den Jahrgängen, in denen ich ihr Herausgeber war (1980-83); und schließlich

¹ Vortrag anlässlich des 25jährigen Bestehens von ZUMA, Mannheim, 4. November 1999.

- die Stichworte der gerade im Entstehen begriffenen und 2001 erscheinenden International Encyclopedia of the Social and Behavioral Sciences im Vergleich mit der letzten Auflage von 1968.

3. Wenn man sich neuere Lehrbücher der Methoden der empirischen Sozialforschung - also den Kanon - durchsieht, so ist man zunächst einmal überrascht wie wenig sich selbst im Vergleich mit den sechziger Jahren, zumal seit dem Ende der siebziger Jahre verändert hat. Man muß schon fast mit der Lupe suchen, um Unterschiede in der grundsätzlichen wissenschaftstheoretischen Position, im Hinblick auf Forschungsdesign, Skalierung, Stichproben, Datenerhebung und Datenanalyse zu finden. Ein liberaleres Kriterium als der Text ist das Register.

Schauen wir uns zunächst einmal Andreas Diekmanns 1995 zum ersten Mal erschienenen Lehrbuch „Empirische Sozialforschung“ an. Da solche Lehrbücher in der Regel viele Auflagen erfahren, sollte ein neuer Text am besten Veränderungen widerspiegeln. Sucht man in dem Register nach „neuen“ Themen, so fällt zweierlei unschwer auf.

Übersicht 1: „Neuere“ Themen der empirischen Sozialforschung aus dem Stichwortverzeichnis von Andreas Diekmann (1995): Empirische Sozialforschung

ALLBUS	Mittelschichtbias
Big Mac Index	Nash-Gleichgewicht
CATI	Outlierwerte
ego-zentrierte Netzwerke	Pfadanalyse
Ereignisdaten	Postmaterialismus
Evaluationsforschung	Psychophysik
experimentelle Spieltheorie	Recall
experimentelle Wirtschaftsforschung	Retrospektivfragen
Gefangenendilemma	RRD Random Digit Dialing
GESIS	Selektionsbias
Forschungsethik	Sozialindikatoren
ISSP	telephonische Befragung
Kohortenanalysen	Total-Design-Method
Kollektivgut	ZUMA

Das Erfreuliche ist, daß eine ganze Reihe dieser relativ „neuen“ Stichworte unmittelbar mit der Tätigkeit von ZUMA verknüpft ist. So z.B. die Psychophysik von Bernd Wegeners Statusskala, der ALLBUS und das ISSP, Telephonbefragungen und die *total-design-method* bei schriftlichen Befragungen. Aber etwas Spezifisches für die neunziger Jahre wird man schwerlich finden. Sucht man in dem Text genauer nach, so sind es vor allem die Abschnitte über Panel- und Ereignisanalyse, über Spieltheorie und Nutzenmessung,

die ihn von traditionelleren Lehrbüchern unterscheiden, aber eben doch nicht von den Themen der siebziger und achtziger Jahre.

Noch ein Versuch: Rainer Schnell/Paul B. Hill/Elke Esser (1999): Methoden der empirischen Sozialforschung (6. Auflage).

Dieses Buch ist umfangreicher und detaillierter und daher möglicherweise offener für neue Tendenzen. Finden wir hier Neuere oder Anderes? Das Register ist umfangreicher und es ist voll von „Denglisch“, aber es zeigt - bei aller Traditionalität des Textes - doch eine ganze Reihe von Spuren neuerer Entwicklungen. Sie fallen vor allem in drei Gruppen. Zum einen ignorieren sie nicht mehr völlig wissenschaftstheoretische und methodische Positionen, die jenseits eines selbst in Lakatos'scher und Kuhn'scher Manier geläuterten logischen Empirismus und Falsifikationismus liegen: *case studies*, Konstruktivismus, Feyerabend, narrative Interviews, *going native*, Ethnomethodologie und *oral history* stehen für diese Entwicklungslinien. Zum Zweiten gibt es die Einflüsse der analytischen Disziplinierung durch die Nutzen-Erwartungstheorie: siehe Brückenhypothese. Und drittens bilden sich neuere methodisch-statistische Entwicklungen in den Bereichen der Datenerhebung und Datenanalyse ab: EQS, CAPI, *fit indices*, Mehrebenenanalyse, Lebensverlaufsdaten und ähnliches. Aber auch hier verschieben sich eher nur Akzente. Selbst diese Themen sind nicht gerade brandneu, und fast keines war nicht schon geläufig, als ich vor zwanzig Jahren anfang, bei ZUMA zu arbeiten. Es scheint eher einige Zeit gedauert zu haben, bis sie schließlich in die Lehrbücher diffundierten. Alt bekanntes bewegt sich von der Peripherie ins Zentrum der Aufmerksamkeit. Aber langsam. Wer z.B. etwas über Längsschnittdesigns und Längsschnittanalyseverfahren erfahren möchte, würde kaum mehr lernen als ein Student in der Mitte der sechziger Jahre.

Es ist freilich verdienstvoll, daß es an vielen Orten knappe Hinweise auf neuere Entwicklungen gibt, selbst wenn sie den funktionellen Aufbau kaum verändern.

Vielleicht sind Lehrbücher doch nicht die geeignete Materialquelle. Sie erinnern insgesamt eher daran, was in Analogie zu Hans Magnus Enzensbergers „Museum der modernen Poesie“ als „Museum der modernen Sozialforschung“ bezeichnet werden könnte. Und in der Tat waren die neunziger Jahre sicher keine von jugendlichem und optimistischem Elan charakterisierte Hochzeit der empirischen Sozialforschung. Sie galt als Überbleibsel der vorreflektiven Moderne. Im Auftrieb waren die Kulturwissenschaften, die *cultural studies*, die europäische Ethnologie, die Autopoiesis der selbst-referentiellen Systeme, die Diskursanalyse, die Medientheorie und die Erinnerungskultur.

Übersicht 2: „Neuere“ Themen der empirischen Sozialforschung aus dem Stichwortverzeichnis von Rainer Schnell/Paul B. Hill/Elke Esser (1999): Methoden der empirischen Sozialforschung, 6. Aufl.

Attitude-Behavior-Kontroverse	Konstruktivismus
Audits	Lebensverlaufsdaten
Autopoiesis	Life History Calendar
Bounded Recall	Mehrebenenanalyse
Brückenhypothese	methodologischer Anarchismus
CAPI	Multiple Imputation Methods
Case Studies	Non-response
Cox-Regression	narrative Interviews
Discrete Choice Models	Oral History
EQS	Phänomenologie
Ethnomethodologie	Proportional-hazard
Fit-Indices	Repräsentativität
Frame of Reference Probing	RRT Randomized-Response Technik
Going Native	Selbstbeobachtung
IZ	Survivalanalyse
Instrumententheorie	Think-aloud Technik
	Urnenmodell

4. Aber stimmt dieses Bild von Stagnation und Rückzug eigentlich? Was gab es Neues im Zentrum der Sozialforschung? Um dies herauszufinden habe ich die Inhaltsverzeichnisse der letzten Jahrgänge der ZUMA-Nachrichten konsultiert. Hier ist - in stark geraffter Form - das Ergebnis:

Der vorherrschende Eindruck, den man von den in den ZUMA-Nachrichten abgehandelten Themen gewinnt, ist eher einer von Beständigkeit und Wiederholung. Eine Reflektion zentraler methodischer Debatten habe ich zwar nicht gefunden. Dennoch schälen sich eine Reihe ganz neuer Problemfelder heraus. Zunächst bei der Datenerhebung. Die gerade bei ZUMA initiierten und gepflegten Arbeiten zu den kognitiven Grundlagen von Norbert Schwarz und seinen Kooperationspartnern markieren eine Wasserscheide in dem Übergang von der Umfrageforschung als Kunsthandwerk zu einer theoretisch fundierten Methodik, in ähnlicher Weise wie dies schon Jahre vorher für die von Esser und anderen entfaltete *rational choice* Theorie der Befragung galt. „*Thinking About Questions*“ und „*Thinking About Answers*“ sind auch eine Spätfolge der Mannheimer Fakultätskonstellation im Hinblick auf die spezifische Kombination der Sozialpsychologie mit anderen sozialwissenschaftlichen Fächern.

Übersicht 3: Themen der neunziger Jahre in der empirischen Sozialforschung nach den Inhaltsverzeichnissen der ZUMA-NACHRICHTEN Nr. 32-44

Verknüpfung qualitativer und quantitativer Daten	Ausschöpfungsquoten
Beschäftigtenstichprobe	Ernährungsberichterstattung
Regionaldaten	Internetbasierte Umfragen: ZUMA-Online
Netzwerkanalyse	Research
kognitive Grundlagen von Umfragen	Imputation fehlender Einkommenswerte
Multi-attributive Discrete Choice Analyse	Delphimethode
Nonpubs	Methoden für Gesundheitserhebungen
Total-Design-Method	CASMIN Educational Classification
Mikrozensusdaten für die Sozialforschung	kognitive Pretestinterviews
formale Modelle kollektiver Aktionen	Simulation

Der Verlust der Naivität ist aber noch größer bei den Ergebnissen von Befragungen: Mittelschichtbias und dramatisch sinkende Ausschöpfungsquoten erzwingen die Erprobung alternativer Stichproben und Datenquellen - Telephon, Internet, Beschäftigungsregister, Mikrozensuszugang - und alternativer Feldverfahren - *Total-Design-Method* -, aber auch einen veränderten Umgang mit unvollständigen Daten - Imputationsverfahren und Korrekturschätzer für Selektivität. Fehlerbehaftete und unvollständige Umfragedaten machen nach einem Dornröschenschlaf auch Massendaten der amtlichen Statistik und Simulationsverfahren wieder attraktiv.

Die computergestützte Inhaltsanalyse - Stichwort TEXTPACK - ist zwar nicht neu und hat ZUMA die ganze Zeit begleitet, aber in keiner Zeit sind so viele neue Textdatenquellen so leicht zugänglich und recherchierbar geworden. Ich vermute, daß die Nachfrage nach TEXTPACK dies belegen könnte.

Auch im Bereich der Datenanalyse sind der Anfang der neunziger Jahre von ihrem Ende Lichtjahre entfernt. LISREL erhält sich durch die Muthensche Erweiterung auf kategoriale Daten über Wasser, *discrete choice* Modelle - Logit, Tobit, Probit - setzen sich aber als Standardverfahren durch. Ebenso bleibt die Verknüpfung quantitativer und qualitativer Daten mehr Programm als Realität. Die Lager sind eingegraben und spielen ihre ritualisierten Stellungskriege. Erst die Gründung separater Sektionen in der DGS erlaubt dann zögerliche Schritte aufeinander zu. Daß beide Lager in den Aufsichtsgremien von ZUMA vertreten sind, finde ich erfreulich.

Die Ende der siebziger Jahre noch esoterischen Methoden der Längsschnittanalyse - *Survival*- und Ereignisanalyse - werden ebenso zum Allgemeingut wie die dafür erforderlichen retrospektiven und prospektiven Längsschnittdaten. Gegen den Reduktionismus

der Übergangsratenmodelle wird mit *optimal matching* ein Verfahren der Sequenzanalyse propagiert, das verspricht, komplexe diachronische Verlaufsmuster angemessen zu analysieren. Freilich macht sich hier am Ende der Dekade wieder eine gewisse Ernüchterung bemerkbar.

5. Meine dritte Materialquelle ist die neue, im Entstehen begriffene International Encyclopedia of the Social and Behavioral Sciences. Dieses ehrgeizige Unternehmen ist auf 26 Bände angelegt und umspannt alle Sozialwissenschaften von der Genetik und den Neurowissenschaften bis zu Philosophie, Recht und Geschichte. Ich habe alle geplanten Artikel im Bereich von Methoden und Statistik verglichen mit den Stichwörtern aus dem gleichen Gesamtgebiet aus der 18-bändigen Ausgabe der International Encyclopedia of the Social Sciences von 1968. Zunächst fällt auf, daß es in der neuen Ausgabe sehr viel mehr methodische und methodologische Artikel gibt: 132 in der alten Version und 318 in der neuen Version.

Eine Reihe von Themen tauchte 1968 bereits auf. Sie werden aber 2001 nicht in einem, sondern in mehreren ausdifferenzierten Beiträgen abgehandelt. Dazu zählen u.a. Fallstudien, Massendaten der amtlichen Statistik, komparative Methoden, *experimental designs*, Bayessche Theorie, *goodness of fit*, interpretative Methoden, Längsschnittsdesigns und -analysen, multivariate Analysen und Beobachtungsmethoden. Auch das Feld der Umfrageforschung ist stärker vertreten und hat sich differenziert.

Was ist neu? Unter anderem Datenschutz und linguistische Methoden, Bayessche Entscheidungstheorie, Ereignis- und Sequenzanalysen, exploratorische und grafische Analyse, Metaanalyse, *missing data* Verfahren, *narratives*, *thick description* und Methodologie des Postmodernismus und der Reflexivität sowie *discrete choice* und hierarchische Modelle. Neu sind auch eine Reihe von Beiträgen über die Infrastruktur der Sozialwissenschaften, nicht zuletzt über die nationalen Umfrageforschungszentren und Datenarchive.

Übersicht 4: „Neue“ Themen der empirischen Sozialforschung auf der Grundlage eines Vergleichs zwischen der International Encyclopedia of the Social Sciences (1968) und der geplanten International Encyclopedia of the Social and Behavioral Sciences (2001)

Biographical Methods	Exploratory Data Analysis
Case Study	Goodness of Fit: Methods, Models and Diagnostics
Causal Inference and Statistical Fallacies	Graphical Models
Comparative Studies	Hierarchical Models: Random and Fixed Effects
Constructivism / Constructionism	Indicator: Methodology
Counterfactuals	Interpretative Methods
Decision Theory, Bayesian	Linear Hypothesis: Fallacies and Interpretive Problems, Simpson's Paradox
Emergent Properties	Longitudinal Data
Event History Analysis: Applications	Meta Analysis
Event Sequence Analysis Methods	Missing Statistical Data
Monte Carlo Methods and Bayesian Computation: Important Sampling	Robustness, in Statistics
Monte Carlo Methods and Bayesian Computation: Important Sampling	Sample Surveys: Cognitive Aspects of Survey Design
Multilevel Methods	Secondary Analysis
Multivariate Analysis: Discrete Variables, Logistic Regression	Selection Bias, Statistic of
Narrative: Methodology	Sequential Statistical Methods
Numerical Methods in Historical Linguistics	Survival Analysis: Cox Model
Observational Studies: Case / Control and Cohort Method	Thick Description
Pattern Matching / Modus Operandi Methods	Triangulation: Methodology
Postmodernism: Methodology	Visual Methods, in Statistics

6. Ich will nun auf dieser Grundlage den Versuch unternehmen, die Topoi und Trends der neunziger Jahre herauszuarbeiten.

Vorweg will ich aber nicht versäumen hervorzuheben, daß die Transformationsforschung Beginn wie Ende der neunziger Jahre der empirischen Sozialforschung prägt. Ich denke, daß sich die Bilanz der empirischen Sozialforschung auf diesem Gebiet im großen und ganzen sehen lassen kann und daß sich gerade das etablierte Arsenal an Methoden - Einstellungsforschung, Eliteforschung, Haushaltspanel und Kohortenstudien, Längsschnittstudien, Evaluationsforschung, Firmenfallstudien, Biographie- und Milieuforschung - trotz einiger anfänglicher Zweifel bewährt hat.

Vielleicht die bedeutendste Entwicklung bezieht sich weniger auf große Innovationen als auf den Ertrag längerfristiger, kumulativer Anstrengungen, die in den sechziger und sieb-

ziger Jahren begonnen wurden (und an denen viele hier im Raum in verdienstvoller Weise beteiligt waren). Ich meine damit den erstaunlichen Reichtum an nunmehr für Sekundäranalysen verfügbaren Aggregat- und vor allem an Mikrodatensammlungen für Regionen, Nationen und historische Zeitpunkte. Die Liste dieser Datenbestände ist fast unübersehbar geworden und stellt neue Probleme der Informationsaufbereitung und Informationsvermittlung. Ihnen allen sind ALLBUS und ISSP, die Wahlstudien, der *World Value Survey*, die TIMMS-Studien und PISA zu Schulleistungsvergleichen, der Eurobarometer, die Familien- und Jugendsurveys und die Wohlfahrtssurveys wohlbekannt. Dazu zählt aber auch CASMIN, die *Luxemburg Income Study*, der Alterssurvey, die Serie von Zeitbudgetstudien und viele andere. Hier konnte in den neunziger Jahren eine Ernte eingefahren werden, die noch längst nicht abgeschlossen ist. Was ansteht ist die Entwicklung einer weltweiten *virtual data library*, in der Datenarchive und lokale Datenbanken miteinander in einer Weise vernetzt sind, die es erlaubt, Variablen, Messinstrumente, Methodeninformationen und Daten auf einfache Weise zu recherchieren, abzurufen und zu verknüpfen. An der Harvard University hat sich gerade ein Center für Basic Research in the Social Sciences etabliert, das sich u.a. dieser Aufgabe widmet.

Mir selbst liegen die gewaltigen Fortschritte im Bereich der Panel- und Kohortenstudien besonders nahe. Nach dem Vorbild des PSID und des deutschen Sozio-Ökonomischen Panels sind nunmehr eine Vielzahl nationaler Haushaltspanelstudien im Gange, die häufig mit deutlich verbesserten Anteilen an retrospektiver Information das Problem der Linkszensur eingrenzen. Zu den neueren Längsschnittdaten für repräsentative Bevölkerungsstichproben gehören auch die HIS-Studienverlaufsstudien, der *Family and Fertility Survey*, die EUROSTAT *Labor Market Surveys* und Studien von INSEE. Daneben haben sich als Retrospektiv- und Prospektivstudien, Einzelkohortenuntersuchungen wie unsere *German Life History Study*, die *National Longitudinal Study of Youth*, die *US National Retirement Study* und die britischen Kohortenstudien durchgesetzt. Was hier noch fehlt, sind internationale Datenbanken, die bislang z.T. an nationalen Datenschutzbestimmungen scheitern. Eine große, noch prekäre Innovation ist für Deutschland der geplante Entwicklungsbuss für Fragestellungen der frühkindlichen und jugendlichen Entwicklung. Es wäre meines Erachtens für beide Seiten vorteilhaft, wenn sich ZUMA bei der methodischen Betreuung dieses Entwicklungsbusses stark engagieren würde. Repräsentative Längsschnittdaten und Längsschnittanalysen sind zum universalen Standard geworden. Dennoch haben sich einige der großen Erwartungen aus dem Beginn der achtziger Jahre nicht erfüllt. Nicht erfüllt hat sich z.B. die Erwartung, daß Verläufe von Zeitabhängigkeit gefunden würden, die sich inhaltlich-theoretisch ableiten lassen, wie z.B. die Diffusion von Infektionskrankheiten oder der Verlauf von Arbeitsplatzsuchprozessen.

Diese Art von Datensammlungen sind nicht nur für die Forschung, sondern auch für die Politikberatung von unschätzbarem Wert. Ein Problem in diesem Zusammenhang ist allerdings immer noch die für solche Zwecke häufig zu geringe Fallzahl. Denken Sie nur an die erbittert geführte Kontroverse über den Nutzen und Schaden der Kombination von Abitur und Lehre oder die Kontroversen über die Erfolge von Umschulungs- und Weiterbildungsmaßnahmen. Die von der Bundesregierung unterstützte Verdoppelung der Stichprobengröße des Sozio-Ökonomischen Panels ist hier ein großer Durchbruch.

Gleichzeitig ist freilich ein hoher Grad an Ernüchterung eingetreten, was die Qualität der Daten, die Eindeutigkeit und Reichweite der mit ihnen möglichen Aussagen und die Voraussetzungen herkömmlicher statistischer Schätz- und Prüfmodelle anbelangt. Ich will diese drei Themen hintereinander kurz beleuchten.

Schwindende Ausschöpfungsquoten beherrschten nicht nur als empirischer Befund die Szene und stellten die naive Anwendung der Inferenzstatistik in Frage. Die formale Einsicht in die Folgen von Selektivität als *sample selection bias* in der Korrelation zwischen Auswahlverzerrung und abhängigen Variablen bedrohen zunehmend auch das Zutrauen in kausale Befunde. Hinzukommt selbst bei ganz ordentlichen Stichproben das Problem fehlender Werte bei einzelnen Variablen, bei Retrospektivstudien das Problem von *recall bias*, bei Panelstudien die Probleme von Panelmortalität und nicht-kontinuierlicher Beteiligung. „Selektivität“ ist vermutlich das zentrale Thema der neunziger Jahre. Nach meiner Einschätzung sind die von den Meinungsforschungsinstituten ausgewiesenen Ausschöpfungsquoten selbst bei *random route* und Einwohnermeldeamtsstichproben häufig die Folge einer erheblichen Artistik. Aber selbst dann bewegen sie sich leicht um die 50 Prozent. Univariate Punkt- und Verteilungsschätzer werden so fast ausgeschlossen und sind auch durch subtile Gewichtungsmasnahmen nicht zu retten. Nun wissen wir zum Glück sowohl theoretisch abgeleitet wie empirisch, daß Repräsentativitätseinbußen bei hinreichender Heterogenität der unabhängigen Variablen für Kausalanalysen weniger gravierend sind und Kovarianzmatrizen eine hohe Robustheit bei Selektivität aufweisen. Leider ist damit das Problem nur verschoben. Woher wissen wir, ob Heterogenität vollständig ist? Wie kontrolliert man unbeobachtete Heterogenität? Selektivitätskorrekturen sind zwar inzwischen unverzichtbar, will man die *peer review* Hürde bei Zeitschriften nehmen, aber sie sind in der Regel wiederum höchst voraussetzungsreich.

Will man beim Sozio-Ökonomischen Panel kontinuierliche Verläufe über zehn Jahre messen, so landet man leicht bei 25 Prozent der Ausgangsstichprobe für einzelne Variablen. Dies beeinträchtigt den Nutzen von Panelstudien für echte Verlaufsanalysen beträchtlich. Was die Verzerrungen von Retrospektivdaten anbelangt, so hat sich zwar sowohl die psychologische Theorie - Stichwort *autobiographical memory* - als auch die

Forschungslage - Stichwort Mannheimer Scheidungsstudie - entscheidend verbessert und für bestimmte Variablen eher optimistisch gestimmt, aber wir sind noch weit entfernt von der endogenen Modellierung von *recall bias* in inhaltlichen Analysen.

Spätestens die Arbeiten von Lieberson und Freedman haben die kausalen Interpretationen des linearen Regressionsmodells und seiner Ableger totwund geschossen, obwohl die Forschungspraxis sich immer noch erstaunlich unberührt zeigt. Signifikanztests wurden zum *dirty word*. Gefragt waren nunmehr Maße für die Effektstärke, theoretisch fundierte Modellprüfung mit individuell angepasster Maximum-Likelihood-Schätzung und intelligere *goodness-of-fit*-Maße statt Tests auf statistische Unabhängigkeit. Die implizite Logik von Quasi-Experimenten mit *randomized controls* scheint kaum mehr zu verteidigen. Die Bayessche Statistik scheint dabei das Galton/Fisher/Yule Modell zu besiegen.

Martin Irle machte gegen Ende der siebziger Jahre auf dem Korridor von A4 eine beiläufige Bewertung, die sich als hellsichtiges Menetekel erwiesen hat: Umfragen seien Quasi-Experimente mit Selbstselektion der Probanden ohne Replikation.

Auch die international vergleichende quantitative, aber auch qualitative Forschung geriet in schwere Stürme. Das *small N*-Problem, das Missverhältnis der Anzahl von möglichen Kausalbedingungen und Fällen und die Autokorrelation der Zeitreihen im neuen Kleid der Pfadabhängigkeit stellte die Möglichkeit und Entscheidbarkeit kausaler Zurechnungen ernsthaft in Frage. Der makrosoziologische Induktivismus der *cross-national aggregate indicators* Tradition - vulgo OECD-Forschung - scheint endgültig obsolet, obwohl damit noch ganze Bibliotheken gefüllt werden.

Die genannten Probleme von Mikro- und Makroanalysen potenzieren sich noch, wenn man aus dem Elfenbeinturm tritt und aus sogenannten Kausalanalysen Politikempfehlungen ableiten will. Was ergibt sich aus den TIMMS-Studien für die Schulpraxis des deutschen Mathematikunterrichts? Welche Niedrigeinkommenspolitik kann man aus der OECD-Korrelation von Beschäftigungs- und Arbeitslosigkeitsquote ableiten?

In den neunziger Jahren wurde zwar die Diagnose dieser Malaisen weit vorangetrieben. Die Bilanz der Lösungen ist aber noch wenig imposant. Ich nenne stichwortartig nur einige davon:

- der Zugang zu Registerdaten ohne Selektivitätsprobleme;
- die systematische Erschließung von Daten zwischen der Ebene von Personen und der Ebene von Nationen: Haushalte, Netzwerke, Firmen und deren Verknüpfung;
- Triangulation: Verhaltensmessungen, Einstellungsdaten, kulturelle Indikatoren, Firmenpanel, qualitative Daten, Texte und historische Quellen, Beobachtungen;
- viel mehr explizite Theorie und die Explikation konkurrierender Erklärungen;

- detaillierte Verlaufsanalysen und Akteursmodelle als Annäherungen an kausale Mechanismen;
- eine Renaissance des Wertes deskriptiver Untersuchungen ohne Erklärungsanspruch zum Auffinden von Explananda;
- statistisch-formale Prüfung von Prämissen unbeobachteter Heterogenität und Selektivität sowie formale Korrekturen.

Überraschenderweise ist die wissenschaftstheoretische Differenz zwischen quantitativer und qualitativer Forschung und von beiden zur postmodernen Diskursanalyse eher kleiner geworden. Es geht um die Konstruktion von Geschichten und deren Plausibilisierung. Die Wahrheit scheint ungreifbarer als je.

Im Ergebnis gibt es wohl viel Anlaß zum Stolz auf das Erreichte - vor allem in bezug auf die Informationsgrundlagen, die noch nie so reichhaltig, so standardisiert waren und noch nie so viele Zeiträume und Nationen abdeckten und in bezug auf die statistische Professionalität.

Es gibt aber mindestens soviel Anlaß zur Bescheidenheit. Das Geschäft der empirischen Sozialforschung ist schwieriger und teurer geworden und erfordert ein Mehr an Forschungsorganisation. Die Unverzichtbarkeit und die gesteigerte und veränderte Bedeutung einer Einrichtung wie ZUMA ist offensichtlich. Sie muß das Erreichte absichern und vermitteln ohne in Konventionalität abzugleiten oder zu verharren. Sie muß sich radikal öffnen, die Malaise artikulieren und forcieren und die Bilanz der Lösungen verbessern.

Ich wünsche den Mitarbeiterinnen und Mitarbeitern und uns allen in diesem Sinne erfolgreiche weitere 25 Jahre.

Korrespondenzadresse

*Prof. Dr. Karl Ulrich Mayer
Max-Planck-Institut für Bildungsforschung
Lentzeallee 94
D-14195 Berlin
Tel. +49-30-824 06-261/262
E-mail: sekmayer@mpib-berlin.mpg.de*

Literatur

- Diekmann, A., 1995: Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Burghard König (Hrsg.), rowolts enzyklopädie. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Freedman, D. A., 1991: Statistical Models and Shoe Leather. S. 291-358 in: P. V. Marsden (Hrsg.), *Sociological Methodology*, Vol. 21. Washington, DC: American Sociological Association.
- Lieberson, S., 1985: Making It Count. The Improvement of Social Research and Theory. Berkeley et al.: University of California Press.
- Lieberson, S., 1985: Asking too much, expecting too little. *Sociological Perspectives* 31: 379-397.
- Lieberson, S., 1998: Causal Analysis and Comparative Research: What Can We Learn from Studies Based on a Small Number of Cases? S. 129-146 in: H.-P. Blossfeld/G. Prein (Hrsg.), *Rational Choice Theory and Large-Scale Data Analysis*. Boulder, CO: Westview Press.
- Marini, M.M./Singer, B., 1988: Causality in the Social Sciences. S. 347-409 in: Clifford C. Clogg (Hrsg.), *Sociological Methodology*, Vol. 18. Oxford: Basil Blackwell.
- Mayer, K. U., 1998: Causality, Comparisons and Bad Practices in Empirical Social Research. A Comment on Stanley Lieberson's Chapter. S. 146-157 in: H.-P. Blossfeld/G. Prein (Hrsg.), *Rational Choice Theory and Large-Scale Data Analysis*. Boulder, CO: Westview Press.
- Schnell, R./Hill, P.B./Esser, E., 1999: *Methoden der empirischen Sozialforschung*. (6., völlig überarbeitete und erweiterte Auflage). München [u.a.]: Oldenbourg.
- Sills, D. L., (Hrsg.), 1968: *International Encyclopedia of the Social Sciences* : Bd. 1-19. New York [u.a.]: Macmillan.
- Singer, B./Marini, M. M., 1987: Advancing Social Research: An Essay Based on Stanley Lieberson's Making it Count: The Improvement of Social Theory and Research. S. 373-391 in: Clifford C. Clogg (Hrsg.), *Sociological Methodology*, Vol. 17. Washington, DC: American Sociological Association.
- Smelser, N. J./P. B. Baltes (Hrsg.) 2001: *International Encyclopedia of the Social and Behavioral Sciences* : Bd. 1-26. Oxford: Elsevier.

DIE BEFRAGUNG VON ARBEITSMIGRANTEN: EINWOHNERMELDEAMT-STICHPROBE UND TELEFONISCHE ERHEBUNG?

NADIA GRANATO

Im folgenden Bericht wird der Frage nachgegangen, ob die Befragung spezieller Populationen wie der Arbeitsmigranten telefonisch durchgeführt werden kann, da es sich bei derartigen Zielgruppen empfiehlt, die Stichprobe aus einem Einwohnermelderegister zu ziehen. Zwei entscheidende Faktoren bei dieser Fragestellung sind die Telefondichte und die Identifizierungsquote. Die Ergebnisse einer Studie, die sich mit der Untersuchung dieser beiden Größen bei Migranten und speziell bei Türken befaßt, werden hier vorgestellt.

When conducting a survey with a special target population such as migrants, it is recommendable to draw the sample from a citizen register. The question this report deals with is whether this sampling method can successfully be combined with telephone interviewing. Two crucial factors in this context are the telephone coverage and the identification quota, an analysis of which for (Turkish) migrants is discussed in this report.

1. Hintergrund

Für die Befragung spezieller Bevölkerungsgruppen wie zum Beispiel der Migranten bietet es sich an, die Stichprobenziehung auf der Grundlage eines Einwohnermelderegisters durchzuführen. Da diese Register Angaben zur Nationalität der erfaßten Personen enthalten, ist eine Abgrenzung der interessierenden Auswahlgesamtheit möglich (Koch 1997: 106). Dies ist - neben der weitgehenden Vollständigkeit der Urliste - ein entscheidender Vorteil gegenüber anderen Methoden der Stichprobengewinnung, da auf diese Weise nur Elemente ausgewählt werden, die zur Zielpopulation gehören und somit kein Screening-Aufwand entsteht. Mit der Entscheidung für eine Einwohnermelderegister-Stichprobe ist aber die Frage, welche Erhebungsmethode verwendet werden soll, noch offen. Naheliegender wäre es, Face-to-face Interviews durchzuführen, weil die Adressen der Zielpersonen vorliegen. Persönlich-mündliche Befragungen sind aber mit vergleichsweise hohen organisatorischen und finanziellen Kosten verbunden. Geringere Kosten, weniger organisato-

rischer Aufwand und die hohe Qualität der erhobenen Daten sind aber nicht zuletzt die Gründe dafür, warum telefonische Befragungen in den letzten Jahren in der empirischen Sozialforschung zunehmend Verwendung gefunden haben (vgl. de Leeuw/van der Zouwen 1988; Hippler/Schwarz 1990: 441).

In diesem Beitrag soll also untersucht werden, ob es möglich ist, eine telefonische Befragung durchzuführen, wenn die Stichprobe aus einem Melderegister gezogen wurde. Die Frage, ob die Voraussetzungen dafür gegeben sind, daß die Vorteile, die eine telefonische gegenüber einer Face-to-face-Erhebung aufweist, auch bei der Befragung von Migranten, die zufällig aus einem Melderegister gezogen wurden, zum Tragen kommen können, entstand im Rahmen des Projektes 'Partizipation von Zuwanderern'¹. Ziel des Projektes ist es, die Frage zu beantworten, warum sich Zuwanderer unterschiedlichen ethnischen und deutschen Vereinigungen anschließen. In der ersten Projektphase soll zunächst die Partizipationsquote ermittelt werden, um das Stichprobendesign der Hauptstudie bestimmen zu können. Dazu wird eine Befragung unter den türkischen Einwohnern Mannheims (als größter Migrantengruppe) durchgeführt. Im Vorfeld dieser 'Randverteilungsstudie', deren Stichprobe aus dem Einwohnermelderegister der Stadt Mannheim gezogen wird, galt es in einer weiteren Studie die Frage zu klären, welche Erhebungsmethode - Face-to-face versus telefonisch - verwendet werden sollte. Die Ergebnisse dieser Vorstudie werden hier dargestellt². Dabei wird zum einen die grundsätzliche Bedeutung der Telefondichte für die Durchführung einer telefonischen Erhebung erläutert, zum anderen wird dargestellt, welche zusätzliche Größe infolge der beschriebenen Kombination von Einwohnermelderegister-Stichprobe und telefonischer Befragung zu berücksichtigen ist. Es handelt sich hierbei um die sogenannte Identifizierungsquote. Gemeint ist damit der Anteil der Stichprobenelemente, für die eine Telefonnummer ermittelt werden kann. In einem weiteren Schritt sollen dann die Telefondichte und die Identifizierungsquote empirisch untersucht werden, um festzustellen, ob die Durchführung einer telefonischen Erhebung bei Migranten möglich ist.

¹ Das Projekt 'Partizipation von Zuwanderern' ist am Mannheimer Zentrum für Europäische Sozialforschung angesiedelt. Es wird von der Deutschen Forschungsgemeinschaft finanziert und steht unter der Leitung von Claudia Diehl und Hartmut Esser.

² Da diese Vorstudie auf die speziellen Anforderungen des Forschungsprojektes ausgerichtet ist, stehen bei einigen Analysen die Türken in den alten Bundesländern beziehungsweise die türkischen Bewohner Mannheims im Mittelpunkt. Auch wird aufgrund der sehr spezifischen Fragestellung auf eine allgemeine Diskussion der Vor- und Nachteile verschiedener Erhebungsmethoden verzichtet.

2. Telefondichte und Identifizierungsquote

Unabhängig davon, welcher Auswahlrahmen der Stichprobenziehung zugrunde liegt, steht die Telefondichte bei telefonischen Erhebungen in unmittelbarem Zusammenhang mit der Qualität der erhobenen Daten. Eine hohe Datenqualität kann nur dann erreicht werden, wenn die Zielgruppe der Befragung ausreichend mit Telefonen versorgt ist. Denn nur wenn jedes Mitglied der zu befragenden Population über ein Telefon zu erreichen ist und damit eine Chance hat befragt zu werden, ist es möglich, eine repräsentative Auswahl zu treffen (Häder 1994: 3). Dies wäre bei einer Vollversorgung der Bevölkerung mit Telefonen der Fall. Doch selbst wenn nicht jedes Mitglied der interessierenden Population telefonisch erreichbar ist, das heißt die Telefondichte unter 100 Prozent liegt, kommt es nicht zwangsläufig zu Verzerrungen. Diese treten nur dann auf, wenn derjenige Teil der Grundgesamtheit, der nicht mit einem Telefon versorgt ist, sich systematisch vom Rest unterscheidet (Trewin/Lee 1988: 9). Einen Hinweis auf solche systematischen Unterschiede liefert die Untersuchung der Zusammenhänge zwischen dem Telefonbesitz und einigen sozioökonomischen Merkmalen. Verzerrungen aufgrund der Telefondichte können also nur dann vermieden werden, wenn entweder eine Vollversorgung der Grundgesamtheit vorliegt oder der Telefonbesitz nicht systematisch mit bestimmten sozioökonomischen Merkmalen variiert.

Will man also die Frage beantworten, ob telefonische Erhebungen auch bei ausländischen Befragten durchgeführt werden können, muß untersucht werden, wie hoch die Telefondichte in diesem Bevölkerungssegment ist und ob systematische Variationen vorliegen. Denn nur so kann sichergestellt werden, daß die Datenqualität nicht durch die gewählte Erhebungsmethode beeinträchtigt wird. Zwar liegt die Telefondichte von Privathaushalten in der Bundesrepublik schon seit 1988 bei über 90 Prozent, allerdings wird darauf verwiesen, daß neben den sozial Schwachen auch Ausländerhaushalte eine niedrigere Quote aufweisen (Frey et al. 1990: 15).

Speziell bei der Befragungen von Ausländern ist ferner zu berücksichtigen, daß diese nur circa neun Prozent der Bevölkerung in Deutschland ausmachen (Statistisches Bundesamt 1999). Sie stellen also eine 'seltene' Population dar. Hätte man bei der Stichprobenziehung nicht die Möglichkeit, die Zielpopulation abzugrenzen, würde bei der Durchführung einer Erhebung ein zusätzlicher Aufwand entstehen. Ungefähr die elffache Anzahl von Personen müßte 'gescreent' werden, um den gewünschten Stichprobenumfang in der Zielpopulation zu erreichen. Soll ein derartig aufwendiges Screening vermieden werden, ist es sinnvoll, die Stichprobe aus einem Einwohnermelderegister zu ziehen, weil hier eine Abgrenzung der Auswahlgesamtheit anhand der Nationalität möglich ist (Albers 1997: 123). Der Nachteil der Wahl dieser Stichprobengrundlage in Zusammenhang mit einer

telefonischen Erhebung liegt allerdings darin, daß zwar Name, Adresse, Geschlecht und Alter der Zielpersonen vorliegen, nicht aber ihre Telefonnummern. Diese müssen den Telefonverzeichnissen entnommen werden.

Seit einigen Jahren besteht für die Inhaber von Telefonanschlüssen aber die Möglichkeit, die Angaben zu Namen und Adresse im Telefonbuch nur unvollständig oder überhaupt nicht veröffentlichen zu lassen. Liegt die sogenannte Eintragsdichte, die den Anteil der im Telefonbuch eingetragenen Telefonanschlüsse mißt, unter 100 Prozent, so ist es nicht möglich, für alle Personen, die einen Telefonanschluß haben, die entsprechende Telefonnummer zu ermitteln. Dies bedeutet, daß neben der Telefondichte auch die Eintragsdichte für die Qualität der Ergebnisse einer telefonischen Erhebung, deren Stichprobe aus einem Einwohnermelderegister stammt, ausschlaggebend ist.

Für das frühere Bundesgebiet berichtet von der Heyde (1997: 200) auf Haushaltsebene eine Telefondichte von 97 Prozent. Aus den Telefonlisten ließen sich aber nur 77 Prozent der befragten Haushalte, die angaben ein Telefon zu besitzen, identifizieren. In der zitierten Untersuchung konnte die Eintragsdichte aber nur deshalb (nachträglich) ermittelt werden, weil die über Random-Route ausgewählten Haushalte in persönlich-mündlichen Interviews zum Telefonbesitz befragt wurden.

Für Stichproben, die aus Einwohnermelderegistern gezogen werden, läßt sich diese Eintragsdichte nicht ermitteln, weil derartige Register keine Informationen darüber enthalten, welche Personen über ein Telefon verfügen. Für solche Stichproben läßt sich nur eine Identifizierungsquote ermitteln, der Anteil an Personen nämlich, deren Rufnummer aus einem Telefonverzeichnis zu entnehmen ist. Die Notwendigkeit, die Identifizierungsquote zu ermitteln, ergibt sich als Folge der Kombination des Auswahlrahmens für die Stichprobenziehung (Einwohnermelderegister) und der Erhebungsmethode (telefonische Interviews).

Es ist also festzuhalten, daß bei telefonischen Erhebungen die Telefondichte für die Qualität der Ergebnisse ausschlaggebend ist. Weiterhin wurde verdeutlicht, daß es durchaus sinnvoll ist, bei der Befragung spezieller Populationen wie beispielsweise der Ausländer die Stichprobe aus einem Melderegister zu ziehen. Aus diesem Vorgehen resultiert, daß eine zweite Größe, die Identifizierungsquote, die ebenso wie die Telefondichte als Indikator für die Qualität der erhobenen Daten anzusehen ist, bei der Entscheidung, ob eine telefonische Erhebung durchgeführt werden kann, zu berücksichtigen ist. In den folgenden Abschnitten wird daher die empirische Verteilung dieser beiden Indikatoren bei Migranten untersucht.

3. Empirische Untersuchung der Telefondichte bei Ausländern und Deutschen

Nachdem im vorigen Abschnitt deutlich wurde, welche Bedeutung der Telefondichte bei der Wahl der Erhebungsmethode zukommt, wird hier deren empirische Verteilung bei Ausländern - und als Referenzgruppe auch bei Deutschen - untersucht. Berücksichtigung finden dabei auch einige sozioökonomische Merkmale, soweit diese in den herangezogenen Datenquellen ausgewiesen sind.

Bezüglich der Telefondichte in Deutschland sind nur wenige Angaben in den amtlichen Statistiken der Wirtschaftsrechnung privater Haushalte zu finden. So wird der Anteil von Haushalten mit Telefon für 1997 folgendermaßen dargestellt:

Tabelle 1: Telefondichte nach Haushaltstyp

Haushaltstyp	Telefondichte	
	Alte Bundesländer	Neue Bundesländer
2-Personen-Haushalte von Renteneempfängern mit niedrigem Einkommen	97,1%	95,1%
4-Personen-Haushalte von Arbeitern und Angestellten mit mittlerem Einkommen	99,7%	95,9%
4-Personen-Haushalte von Beamten und Angestellten mit höherem Einkommen	99,5%	96,7%

Quelle: Statistisches Jahrbuch 1998 für die Bundesrepublik Deutschland (S. 547-548).

Insgesamt zeigen sich also kaum Unterschiede zwischen den Haushaltstypen und selbst bei Haushalten mit niedrigem Einkommen kann die Telefondichte mit 97 Prozent im Westen und 95 Prozent im Osten als sehr hoch bezeichnet werden. Allerdings ist eine detaillierte Untersuchung der Unterschiede zwischen Haushalten mit und ohne Telefon genau dann besonders wichtig, wenn die Telefondichte hoch ist. Denn bei einer allgemein hohen Telefonversorgung ist davon auszugehen, daß es sich bei denjenigen ohne Telefon um eine spezielle Subgruppe der Bevölkerung handelt, die sich deutlich vom Rest unterscheidet (Trewin/Lee 1988: 24).

Mit Hilfe der aggregierten Daten der amtlichen Statistik läßt sich zwar festzustellen, daß die Telefondichte besonders in den alten Bundesländern hoch ist. Es ist aber keine Differenzierung nach Nationalität möglich. Für solche Analysen ist es notwendig, Individualdatensätze heranzuziehen. Zu den Individualdatensätzen, die für die Ermittlung nationalitätenspezifischer Telefondichten verwendet werden können, zählen das Europäische Haushaltspanel und das Sozio-oekonomische Panel. Zum Erhebungsprogramm des Euro-

päischen Haushaltspanels für Deutschland³ gehört auch die Frage, ob im Haushalt ein Telefon zur Verfügung steht. Eine Analyse dieser Daten zeigt, daß der Anteil der Haushalte mit Telefon 1995 im gesamten Bundesgebiet bei 94 Prozent lag, wobei dieser im Osten mit 79 Prozent noch deutlich niedriger als im früheren Bundesgebiet (98 Prozent) war. Da die im Rahmen des Projektes 'Partizipation von Zuwanderern' interessierende Migrantengruppe hauptsächlich in den alten Bundesländern angesiedelt ist, wird für die Ermittlung der Telefondichte von Ausländern im folgenden nur noch das frühere Bundesgebiet berücksichtigt.

Das Europäische Haushaltspanel erfaßt ausländische Haushalte nur insoweit, als eine Durchführung der Interviews auf deutsch möglich ist, was zu einer Untererfassung dieser Haushalte führt (Statistisches Bundesamt 1998: 2-1). Weiterhin wird die Nationalitätenvariable nur mit den Ausprägungen 'deutsch' und 'nicht deutsch' weitergegeben, so daß eine detaillierte Analyse der doch sehr heterogenen Gruppe der Migranten nach der jeweiligen Nationalität nicht möglich ist. Mit Hilfe dieser Daten kann also nur ein erster Hinweis auf die Höhe der Telefondichte bei Ausländern gewonnen werden.

Es zeigt sich, daß die Telefondichte bei den ausländischen Befragten mit knapp 96 Prozent nur leicht unter der der Deutschen liegt (Tabelle 2). Zwischen diesen beiden relativen Häufigkeiten besteht ein statistisch signifikanter Unterschied⁴. Aufgrund der sehr niedrigen Fallzahl bei den Ausländern ist eine weitere Differenzierung nach sozioökonomischen Merkmalen nicht sinnvoll.

Tabelle 2: Telefondichte bei ausländischen und deutschen Befragten

	Telefondichte	
Deutsche	98,6%	(n=6575)
Ausländer	95,5%	(n=291)
Insgesamt	98,5%	(n=6866)

Quelle: Europäisches Haushaltspanel, Welle 2 (1995), eigene Berechnungen.

Für eine empirische Untersuchung der sozioökonomischen Merkmale im Zusammenhang mit der Telefondichte bieten sich jedoch die Daten des Sozio-oekonomischen Panels an.

³ Das Europäische Haushaltspanel ist eine Längsschnitterhebung zur Abbildung der Lebens- und Arbeitsbedingungen in Europa, die im Auftrag der Europäischen Kommission durchgeführt wird. Die Stichprobe für Deutschland enthält circa 5000 Haushalte (URL: <http://www.statistikbund.de/download/panel/panel1.html>).

⁴ Das Signifikanzniveau des durchgeführten Tests wurde, wie auch bei allen folgenden Anwendungen, bei einem Wert von $\alpha=0,01$ festgelegt.

Zum einen wird erhoben, ob im Haushalt ein Telefon vorhanden ist, zum anderen ist hier eine Abgrenzung der interessierenden Migrantengruppen möglich. Im Vordergrund der Analyse stehen dabei die Migranten aus den ehemaligen Anwerbeländern Italien, Griechenland, Spanien, Ex-Jugoslawien und der Türkei⁵. Für die Analysen herangezogen wurde die Welle M (1996) des Sozio-oekonomischen Panels, da zu diesem Zeitpunkt die aktuellste Erhebung der Telefondichte erfolgte⁶.

Wie aus Tabelle 3 zu erkennen ist, unterscheidet sich die Telefondichte bei verschiedenen Migrantengruppen nur zum Teil deutlich von der bei den Deutschen. Während die Quote der griechischen Befragten fast so hoch ist wie die deutsche und kein signifikanter Unterschied zwischen diesen beiden Gruppen vorliegt, beträgt die Differenz zwischen den türkischen und deutschen Befragten sieben Prozentpunkte und ist statistisch signifikant. Bei Jugoslawen, Italienern und Spaniern sind die Unterschiede zu den Deutschen nicht so stark ausgeprägt wie bei den Türken, aber ebenfalls signifikant.

Tabelle 3: Telefondichte nach Nationalität in den alten Bundesländern

Nationalität	Telefondichte	
Türkei	91%	(n=759)
Ex-Jugoslawien	94%	(n=394)
Italien	94%	(n=352)
Spanien	94%	(n=124)
Griechenland	98%	(n=264)
Deutschland	98%	(n=6639)

Quelle: Sozio-oekonomisches Panel, 1996, eigene Berechnungen.

Insgesamt kann die Telefondichte aber selbst bei den Türken mit 91 Prozent als ziemlich hoch angesehen werden. Dennoch bleibt die Frage offen, ob sich die Gruppe derjenigen, die nicht über ein Telefon verfügen, durch bestimmte Merkmale auszeichnet. Dies gilt es, in einem weiteren Schritt zu klären. Zu den unterrepräsentierten Gruppen bei Telefoninterviews gehören nach bisherigen Erkenntnissen Personen mit niedriger Schulbildung, geringem Haushaltseinkommen, niedrigem beruflichen Status, Personen in kleinen Haushalten und Personen, die am oberen oder unteren Ende der Altersskala angesiedelt sind (Schulte 1997: 160; Jung 1990: 394). Als eine der möglichen Ursachen hierfür könnte eine geringe Telefondichte bei diesen Gruppen vermutet werden. Es stellt sich die Frage,

⁵ Auf die Einbeziehung portugiesischer Befragter mußte aufgrund der geringen Fallzahl verzichtet werden.

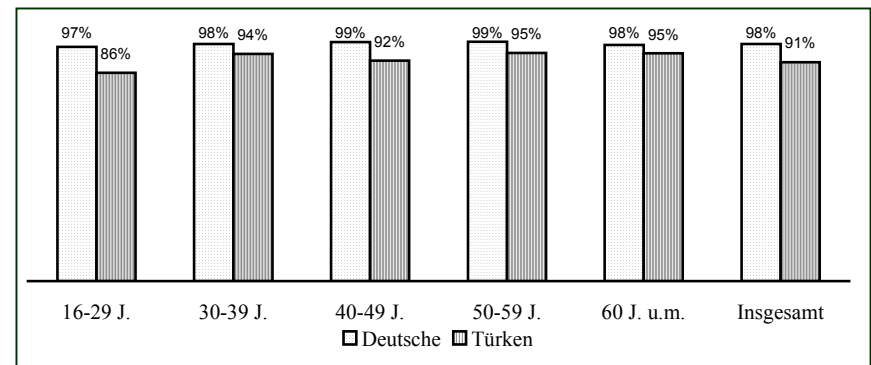
⁶ Es sei darauf hingewiesen, daß die Daten des Sozio-oekonomischen Panels aufgrund von Panelmortalität möglicherweise systematische Verzerrungen enthalten.

ob auch bei Ausländern Unterschiede in der Telefondichte anhand dieser Kriterien zu erkennen sind. Die Berücksichtigung dieser Merkmale dient also bei den folgenden Analysen dazu festzustellen, ob bei einer telefonischen Erhebung bereits im Vorfeld bestimmte Migranten aufgrund einer niedrigen Telefondichte unterrepräsentiert wären.

Bei der Darstellung der jetzt folgenden Verteilungen werden von den zuvor untersuchten Migrantengruppen nur noch die Türken berücksichtigt. Dies hat zwei Gründe: Zum einen stehen sie im Projekt 'Partizipation von Zuwanderern' im Mittelpunkt. Zum anderen erlaubt es die geringe Anzahl der Personen ohne Telefon bei den anderen Migrantengruppen nicht, detailliertere Analysen vorzunehmen.

Auf Personenebene werden für türkische und deutsche Befragte das Alter, die Schulbildung, die Erwerbsbeteiligung sowie die Stellung im Beruf der Befragten in die Analysen einbezogen. Die wichtigsten Merkmale auf Haushaltsebene sind sowohl die Haushaltsgröße als auch das Einkommen (vgl. Thornberry/Massey 1988).

Abbildung 1: Telefondichte bei Türken und Deutschen nach Altersgruppen



Quelle: Sozio-oekonomisches Panel, 1996, eigene Berechnungen.

Als eines der Merkmale, anhand dessen Verzerrungen bei telefonischen Befragungen festgestellt wurden, wird zunächst das *Alter* untersucht. Sowohl für türkische als auch für deutsche Befragte ist die Telefondichte in der niedrigsten Altersstufe (unter 30 Jahre) am geringsten (Abbildung 1). Mit 86 Prozent liegt sie fünf Prozentpunkte unter dem Durchschnitt bei den türkischen Befragten, bei den Deutschen beträgt diese Differenz allerdings nur einen Prozentpunkt. Dieser Unterschied wirkt sich besonders bei den Türken auf die Quote insgesamt aus, da in der Altersgruppe unter 30 Jahren 42 Prozent der befragten Türken zu finden sind, während es bei den Deutschen nur 22 Prozent sind.

Der Unterschied in der Telefondichte zwischen der jüngsten und den restlichen Altersgruppen ist sowohl bei den Türken als auch bei den Deutschen signifikant.

Bezüglich der *Schulbildung* ist weder bei Deutschen noch bei Türken ein signifikanter Unterschied in der Telefondichte zu erkennen. Von den Türken mit Hauptschulabschluß verfügen 90 Prozent (98 Prozent bei den Deutschen) über ein Telefon, bei denjenigen mit einem höherem Abschluß als dem Hauptschulabschluß liegt die Quote einen Prozentpunkt niedriger. Die Differenz zwischen diesen beiden Gruppen ist bei den deutschen Befragten genauso groß, allerdings zugunsten derjenigen mit höherem Schulabschluß. Der Befund, daß Personen mit niedriger Schulbildung bei telefonischen Erhebungen unterrepräsentiert sind (Schulte 1997), kann demzufolge nicht auf eine geringere Telefondichte in dieser Gruppe zurückgeführt werden.

Im Gegensatz zur Schulbildung wirkt sich die *Erwerbstätigkeit* bei den türkischen Befragten leicht auf die Telefondichte aus. Während die Quote bei den Nichterwerbstätigen bei 89 Prozent liegt, ist sie bei den Erwerbstätigen vier Prozentpunkte höher, allerdings muß dieser Unterschied als zufällig betrachtet werden. Für die deutschen Befragten zeigt sich kein nennenswerter Unterschied.

Unterscheidet man die Erwerbstätigen nach der *Stellung im Beruf*, so sind weder für Türken noch für Deutsche Unterschiede in der Telefondichte⁷ erkennbar.

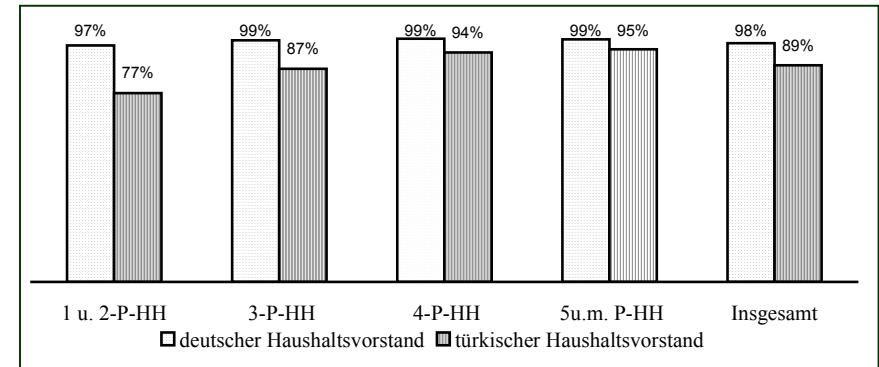
Faßt man die Ergebnisse der Untersuchung auf Personenebene zusammen, so zeigt sich im Hinblick auf Nationalität, Alter und Erwerbstätigkeit ein Unterschied in der Telefondichte. Bezüglich der Nationalität zeichnen sich besonders Türken durch eine niedrige Telefondichte aus, bei der Berücksichtigung des Alters trifft dies auf die jüngste Altersgruppe zu. Erwerbstätige weisen eine leicht höhere Telefonquote auf, wobei hier der Unterschied zu den Nichterwerbstätigen nicht signifikant ist. Bei der Unterscheidung nach Schulbildung oder der Stellung im Beruf konnten keine bedeutsamen Differenzen zwischen den Gruppen festgestellt werden.

Neben persönlichen Merkmalen spielen aber auch Haushaltsmerkmale bei der Beurteilung der Repräsentativität einer Stichprobe eine Rolle. Die Datenstruktur des Sozio-oekonomischen Panels erlaubt es, Analysen bezogen auf Haushalte durchzuführen. Zu den Merkmalen, anhand derer sich Haushalte mit und ohne Telefon unterscheiden, gehören die Haushaltsgröße, das Einkommen und der Erwerbsstatus des Haushaltsvorstands. So ist

⁷ Zwar fallen bei den türkischen Erwerbstätigen die Selbständigen mit einer Telefondichte von 100 Prozent auf, allerdings ist eine sinnvolle Interpretation dieser Quote aufgrund der geringen Fallzahl (n=18) nicht möglich.

bekannt, daß die Telefondichte in kleinen Haushalten, bei geringem Einkommen und erwerbslosen Haushaltsvorständen niedrig ist (Trewin/Lee 1988: 14).

Abbildung 2: Telefondichte nach Haushaltgröße für Haushalte mit deutschem bzw. türkischem Haushaltsvorstand



Quelle: Sozio-oekonomisches Panel, 1996, eigene Berechnungen.

Betrachtet man die *Haushalte* unter Berücksichtigung ihrer *Größe* (Abbildung 2), zeigt sich, daß bei den Haushalten mit türkischem Vorstand besonders die kleinen Haushalte⁸ eine niedrige Telefondichte aufweisen. Dies trifft vor allem auf die Ein- und Zweipersonenhaushalte⁹ zu. Überdurchschnittliche Quoten weisen hingegen die größeren Haushalte auf. Dies ist insofern von Bedeutung, als über 50 Prozent aller Haushalte mit türkischem Haushaltsvorstand in die beiden höchsten Kategorien fallen. Hierin kann eine Ursache dafür gesehen werden, daß die Telefondichte auf Haushaltsebene insgesamt mit 89 Prozent etwas niedriger ist als auf Personenebene.

Auch hinsichtlich des *Haushaltseinkommens* ergeben sich signifikante Unterschiede in der Verteilung der Telefondichte. Sowohl für Haushalte mit deutschem als auch mit türkischem Vorstand ist der Anteil derer, die über ein Telefon verfügen, bei einem Haushaltsnettoeinkommen bis zu 3000 DM niedriger als bei Haushalten mit höherem Einkommen. Während es sich aber bei den deutschen Haushalten nur um eine Differenz von fünf Pro-

⁸ Die Zusammenfassung von Ein- und Zweipersonenhaushalten ist mit der geringen Fallzahl bei den Haushalten mit türkischem Haushaltsvorstand in diesen beiden Kategorien zu begründen.

⁹ Auch hier zeigt sich beim Vergleich zwischen Ein- und Zweipersonen-Haushalten und dem Rest, daß der Unterschied in der Telefondichte signifikant ist.

zentpunkten handelt, sind es in der Gruppe der türkischen Haushalte immerhin 13 Prozentpunkte.

Die Analyse der verschiedenen Datenquellen zeigt ein einheitliches Bild. Die Telefondichte liegt bei Ausländern insgesamt etwas niedriger als bei Deutschen. Die niedrigste Quote unter den betrachteten Migrantengruppen weisen die Türken auf. Weiterhin wurde deutlich, daß unter den türkischen Befragten junge Leute und Nichterwerbstätige seltener über ein Telefon im Haushalt verfügen. Außerdem war bei kleinen Haushalten und niedrigem Einkommen der Anteil der Haushalte ohne Telefon vergleichsweise hoch. Es zeigt sich also, daß Personen, die zu diesen Gruppen gehören, bei einer telefonischen Erhebung aufgrund einer relativ geringen Telefondichte untererfaßt werden könnten.

4. Identifizierungsquote und Aktualität der Stichprobenadressen

Bislang wurde untersucht, ob die erste Voraussetzung für die Durchführung einer telefonischen Erhebung bei Migranten – eine ausreichende Telefondichte – gegeben ist. Diese ist selbst bei den türkischen Befragten, die die niedrigste Quote unter den betrachteten Migrantengruppen aufweisen, als hoch anzusehen. Die Berücksichtigung sozioökonomischer Merkmale bei den türkischen Befragten weist aber auf Unterschiede beim Telefonbesitz in der Zielpopulation hin. Unterschiede dieser Größenordnung sollten jedoch nicht in einer grundsätzlichen Ablehnung der Erhebungsmethode bei Migranten resultieren. Vielmehr gilt es jetzt zu überprüfen, inwieweit der zweite erwähnte Indikator - die Identifizierungsquote - für die Durchführung einer telefonischen Erhebung bei Migranten spricht.

Aufgrund der Tatsache, daß die Stichprobe aus einem Melderegister gezogen wird, stellt sich nämlich die Frage, wie hoch der Anteil an Personen ist, für die eine Telefonnummer gefunden werden kann. Im nachfolgenden Abschnitt wird also empirisch untersucht, ob die Quote der Personen, deren Rufnummer aus öffentlich zugänglichen Telefonlisten entnommen werden kann, deutlich unter der Telefondichte liegt und inwieweit sich Unterschiede für deutsche und türkische Befragte ergeben.

Wie eingangs erwähnt, empfiehlt es sich bei der Befragung von Migranten, für die Stichprobengewinnung auf Einwohnermelderegister zurückzugreifen. Da dort neben einigen anderen Merkmalen auch die Nationalität erfaßt wird, kann sichergestellt werden, daß derartige Stichproben nur Zielpersonen mit der gewünschten Nationalität enthalten. Für diese Zielpersonen gilt es, Telefonnummern zu ermitteln.

Dabei treten zwei Probleme auf. Zum einen haben nicht alle Personen ein Telefon, zum anderen kann nicht davon ausgegangen werden, daß alle Personen, die einen Anschluß

haben, auch identifiziert werden können. Identifiziert bedeutet in diesem Zusammenhang folgendes: Stichproben, die aus Einwohnermelderegistern stammen, enthalten bestimmte Angaben der Zielpersonen. Dies sind neben dem Namen, die Anschrift (Postleitzahl, Straße und Hausnummer), das Geschlecht und das Geburtsjahr, nicht aber die Rufnummer. Diese wird anhand des Nachnamens und der Anschrift aus einer Telefonliste ermittelt. Die Telefonnummer einer Zielperson gilt dann als identifiziert, wenn neben dem Nachnamen auch die Anschrift vollständig übereinstimmt.

Sind aber die Anschlüsse nicht oder nur unvollständig in den Telefonlisten erfaßt, was seit einiger Zeit auf Wunsch des Teilnehmers problemlos möglich ist, kann keine Identifizierung stattfinden. Dies gilt selbstverständlich auch für Personen, die kein Telefon haben.

Um die Quote der identifizierbaren Personen empirisch zu ermitteln, wurden zwei Stichproben, die aus dem Melderegister der Stadt Mannheim stammten, herangezogen. Auswahlgrundlage für die erste Stichprobe waren die türkischen Einwohner Mannheims ab 18 Jahren, bei der zweiten waren es die deutschen Einwohner ab 18 Jahren. Beide Stichproben umfassen jeweils 400 Fälle. Um Verzerrungen der ermittelten Identifizierungsquote aufgrund mangelnder Aktualität des Einwohnermelderegisters auszuschließen, wurden die Stichprobenadressen für bestimmte Stadtteile vor Ort überprüft. Beide Schritte wurden sowohl für die türkische als auch für die deutsche Stichprobe durchgeführt, um wiederum über eine Vergleichsgröße zu verfügen.

4.1 Ermittlung der Identifizierungsquote

Zur Ermittlung der Telefonnummern wurde das Online-Telefonbuch der Telekom (URL: <http://195.145.119.197/NSAPI/Anfrage>) verwendet, das sich durch große Aktualität auszeichnet. Ein Nachteil dieser Liste ist allerdings, daß Telefonnummern nur unter Angabe eines Namens gesucht werden können, nicht aber zum Beispiel nur unter Angabe einer Straße. Die Möglichkeit der Suche nur nach Straßennamen (das heißt ohne die Angabe eines Nachnamens) war besonders in Zusammenhang mit der türkischen Stichprobe von Bedeutung. Wie sich herausstellte, enthielten die Telefonlisten Schreibfehler oder Vor- und Nachnamen waren vertauscht. Mit solchen Fehlern behaftete Untersuchungseinheiten hätten nicht identifiziert werden können, wenn es nicht möglich gewesen wäre, Rufnummern straßenweise auflisten zu lassen. Für diese Art der Suche wurde ein Telefonbuch auf CD-ROM (DT-Info & Route 1999) herangezogen. Der Nachteil bei dieser Liste liegt allerdings in einer deutlich geringeren Aktualität¹⁰.

¹⁰ Während das Online-Telefonbuch der Telekom laufend aktualisiert wird, entspricht die CD-ROM DT-Info & Route '99 höchstens dem Stand zum Zeitpunkt der Veröffentlichung (März 1999).

Von den 400 Untersuchungseinheiten der deutschen Stichprobe konnten 250 Telefonnummern gefunden werden, was einer Identifizierungsquote von 62,5 Prozent entspricht. Bei der türkischen Stichprobe ergab sich eine Quote von nur 43,6 Prozent. Insgesamt stellte sich durch die straßenweise Auflistung der Rufnummern heraus, daß in zwölf Fällen Vor- und Nachnamen vertauscht waren und in acht Fällen Schreibfehler vorlagen. Diese 20 Fälle entsprechen immerhin fünf Prozent der türkischen Stichprobe, wobei davon auszugehen ist, daß nicht alle Fehler identifiziert werden konnten. Derartige Fehler tragen sicherlich auch zu der signifikant niedrigeren Identifizierungsquote in der türkischen Stichprobe bei.

Wie von der Heyde (1997: 202) zeigt, gehört das Alter zu den Merkmalen, entlang derer die Eintragsdichte variiert. Haushalte, in denen ältere Personen in mittleren bis großen Städten leben, lassen sich besonders gut identifizieren. Der Grund für die höhere Eintragsdichte wird darin vermutet, daß das Telefon zu einer Zeit angemeldet wurde, als die Eintragung noch nicht optional war. Auch Selbständige und Wohnungseigentümer zeichnen sich durch hohe Eintragungsquoten aus. Deutlich unterdurchschnittliche Quoten weisen junge Menschen und darunter vor allem junge Frauen auf. Hier werden die Furcht vor Belästigungen sowie die höhere Mobilität als Gründe für keine beziehungsweise unvollständige Eintragungen vermutet.

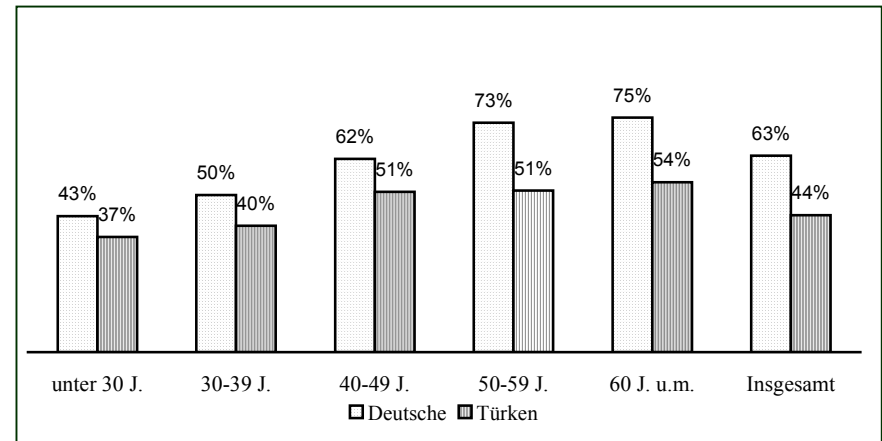
Es stellt sich nun die Frage, ob sich auch in den untersuchten Stichproben bei der Identifizierungsquote hinsichtlich des Alters Unterschiede aufzeigen lassen. Da die Stichproben aus dem Einwohnermelderegister der Stadt Mannheim Informationen zum Alter enthalten, wurde der Frage nachgegangen, ob hier ein Hinweis auf einen weiteren Grund für die niedrigere Identifizierungsquote in der türkischen Stichprobe zu finden ist, der von der Nationalität unabhängig ist¹¹. Die Analyse der Daten des Sozio-oekonomischen Panels hat bereits gezeigt, daß sich die türkischen Befragten durch eine stärkere Besetzung der niedrigen Altersgruppen auszeichnen.

Auch in den Einwohnermeldeamt-Stichproben unterscheiden sich Türken und Deutsche hinsichtlich des durchschnittlichen Alters deutlich. Mit einem Mittelwert von 38 Jahren liegt das Durchschnittsalter der Türken weit unter dem der Deutschen von 50 Jahren. Wie in Abbildung 3 zu erkennen, sind besonders in den beiden niedrigsten Altersgruppen die

¹¹ Die Überprüfung des Zusammenhangs zwischen Alter und Identifizierungsquote kann allerdings nur als Hinweis für die Stimmigkeit der 'Eintragungshypothese' gewertet werden. Zwar ist das Geburtsjahr der Zielperson bekannt, es gibt aber keine Angaben darüber, ob die betreffende Person einen eigenen Haushalt führt und es sich somit – sofern ein Telefon vorhanden – um einen Anschluß handelt, bei dem Einfluß auf die Eintragung in das Telefonbuch ausgeübt werden konnte.

Identifizierungsquoten¹² beider Stichproben unterdurchschnittlich. Berücksichtigt man jetzt noch die Tatsache, daß in der türkischen Stichprobe über 60 Prozent der Personen in diese Alterskategorien fallen (bei den Deutschen sind es nur 36 Prozent), so kann man dies als Hinweis auf eine der Ursachen der niedrigeren Trefferquote in der gesamten türkischen Stichprobe werten.

Abbildung 3: Identifizierungsquote nach Alter



Insgesamt zeigt sich, daß die Identifizierungsquote sowohl für deutsche als auch für türkische Zielpersonen deutlich unter der jeweiligen Telefondichte liegt. Zum einen könnte die Ursache für diese Differenz darin gesehen werden, daß von der Möglichkeit, die Telefonnummer gar nicht oder nur mit unvollständigen Angaben in Telefonlisten eintragen zu lassen, rege Gebrauch gemacht wird. Eine andere Ursache könnte aber sein, daß die Aktualität der Melderegister nicht der der Telefonlisten entspricht. Aus diesem Grund wird im nächsten Abschnitt die Überprüfung der Aktualität der Stichprobenadressen kurz dargestellt.

4.2 Aktualität der Stichprobenadressen

Um die Aktualität der Stichprobenadressen einschätzen zu können, wurde eine Überprüfung vor Ort vorgenommen, das heißt es wurde kontrolliert, ob unter der angegebenen

¹² Ein Vergleich der relativen Häufigkeit der gefundenen Telefonnummern zwischen den beiden jüngsten Altersgruppen und dem Rest zeigt sowohl für die türkische als auch für die deutsche Stichprobe signifikante Unterschiede.

Adresse ein Klingelschild mit dem entsprechenden Nachnamen zu finden war. Da eine Überprüfung der gesamten Stichprobe aufgrund der beschränkten Projektressourcen nicht zu realisieren war, wurden nur die im Innenstadtbereich gelegenen Straßen 'begangen'. Auf den Innenstadtbereich entfielen insgesamt 140 Adressen, 97 davon stammten aus der türkischen Stichprobe. Die Quote der Namen, die unter den angegebenen Adressen gefunden werden konnten, zeigt für die türkische Stichprobe mit 78 Prozent einen niedrigeren Wert als für die deutsche Stichprobe (93 Prozent). Allerdings liegen beide Werte so hoch, daß mangelnde Aktualität der Stichprobenadressen als Ursache für eine derartig niedrige Quote an gefundenen Telefonnummern ausgeschlossen werden kann¹³.

Die Ermittlung des Anteils der Namen, die unter der in der Stichprobe angegebenen Adresse gefunden werden konnten, ist aber unter einem weiteren Gesichtspunkt interessant. Anhand der gewonnenen Informationen ist es möglich, die Identifizierungsquote zu bereinigen. In der türkischen Stichprobe wurde immerhin für 51 Prozent der Zielpersonen, deren Adresse korrekt war, eine Telefonnummer gefunden. Unbereinigt lag die Identifizierungsquote im Innenstadtbereich bei nur 44 Prozent.

Faßt man die Ergebnisse der hier dargestellten Analysen zusammen, so läßt sich folgendes feststellen: Der Anteil an Zielpersonen, die unter der angegebenen Adresse gefunden werden können, liegt deutlich über dem Anteil an Zielpersonen, für die eine Telefonnummer ermittelt werden konnte. Dies zeigt, daß die niedrige Identifizierungsquote - besonders in der türkischen Stichprobe - weniger auf die mangelnde Aktualität der Melderegister zurückzuführen ist. Vielmehr weist die Differenz zwischen der für die türkische Bevölkerung ermittelten Telefondichte von ungefähr 90 Prozent und einer entsprechenden Identifizierungsquote von 44 Prozent auf die Unvollständigkeit von Telefonlisten hin. Sowohl für die Telefondichte als auch für die Identifizierungsquote wurde festgestellt, daß diese bei Personen unter 30 Jahren deutlich unterhalb des Durchschnittswertes lagen.

5. Fazit

In Anbetracht der Ergebnisse der dargestellten Untersuchung scheint es nicht sinnvoll, Erhebungen, deren Stichproben aus einem Einwohnermelderegister gezogen werden, nur telefonisch durchzuführen. Dies trifft nicht nur auf spezielle Populationen wie zum Bei-

¹³ Es sollte außerdem berücksichtigt werden, daß beide für die Begehung gewählten Wohnbezirke sich durch einen hohen Anteil von Mehrfamilienhäusern auszeichnen. Da in solchen Wohngebieten mit einer größeren Mobilität zu rechnen sein dürfte als in Stadtteilen mit überwiegendem Anteil an Ein- und Zweifamilien-Häusern, ist also anzunehmen, daß der Anteil der vor Ort auffindbaren Adressen in der Gesamtstichprobe etwas höher liegt.

spiel die Migranten zu, sondern auch auf deutsche Befragte. Welche Alternativen gibt es, eine telefonische Befragung durchzuführen, bei denen eine niedrige Identifizierungsquote kein Problem darstellt? Zum einen gibt es die Möglichkeit, die Stichprobe - wie bei telefonischen Erhebungen sonst üblich - anhand zufallsgenerierter Rufnummern zu gewinnen und dann ein Screening durchzuführen. Auf diese Art und Weise kann die Unvollständigkeit von Telefonlisten nicht zu den oben geschilderten Problemen führen. Allerdings sind die mit dem Screening verbundenen Kosten bei einer Entscheidung für eine telefonische Erhebung zu berücksichtigen. Darüber hinaus sind bei einem solchen Vorgehen möglichen Effekte aufgrund einer niedrigen Telefondichte nicht ausgeschlossen.

Zum anderen bietet sich in einem solchen Fall ein Methodenmix in der Form an, daß bei denjenigen Zielpersonen, für die keine Telefonnummer gefunden werden kann, die Befragung Face-to-face stattfindet (Groves 1990: 231). So können die Vorteile einer telefonischen Erhebung genutzt und gleichzeitig systematische Verzerrungen der Ergebnisse, die aus einer niedrigen Identifizierungsquote oder Telefondichte resultieren, vermieden werden (vgl. Groves 1989: 539; Anders 1990: 435).

Allerdings setzt die Entscheidung für einen Methodenmix voraus, daß zwischen beiden Erhebungsmethoden hinsichtlich der Teilnahmebereitschaft keine gravierenden Unterschiede bestehen. Zwar weisen die Ergebnisse vergleichender Untersuchungen darauf hin, daß diese bei telefonischen Befragungen durchaus höher ist als bei Face-to-face Interviews (Reuband/Blasius 1996: 304). Ungeklärt bleibt aber weiterhin die Frage, ob dies auch für Migranten zutrifft.

Welche der Alternativen zu wählen ist oder ob überhaupt telefonisch erhoben werden soll, kann nicht generell entschieden werden, es gilt, die Nutzen und Kosten im Einzelfall abzuwägen.

Literatur

Albers, I., 1997: Einwohnermelderegister-Stichproben in der Praxis. Ein Erfahrungsbericht. S. 117-126 in: S. Gabler/J. Hoffmeyer-Zlotnik (Hrsg.), Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.

Anders, M., 1990: Praxis der Telefonbefragung. S. 426-436 in: Forschungsgruppe Telekommunikation (Hrsg.), Telefon und Gesellschaft, Band 2. Berlin: Volker Spiess.

de Leeuw, E./van der Zouwen, J., 1988: Data Quality in Telephone and Face To Face Surveys: A Comparative Metaanalysis. S. 283-300 in: R. Groves et al., (Hrsg.), Telephone Survey Methodology. New York: John Wiley & Sons.

Frey, J./Kunz, G./Lüschen, G., 1990: Telefonumfragen in der Sozialforschung: Methoden, Techniken, Befragungspraxis. Opladen: Westdeutscher Verlag.

- Groves, R., 1989: Survey Errors and Survey Costs. New York: John Wiley & Sons.
- Groves, R., 1990: Theories And Methods Of Telephone Surveys. Annual Review of Sociology 16: 221-240.
- Häder, S., 1994: Auswahlverfahren bei Telefonumfragen. ZUMA-Arbeitsbericht 94/03.
- Hippler, H.-J./Schwarz, N., 1990: Die Telefonbefragung im Vergleich mit anderen Befragungsarten. S. 437-447 in: Forschungsgruppe Telekommunikation (Hrsg.), Telefon und Gesellschaft, Band 2. Berlin: Volker Spiess.
- Jung, M., 1990: Ausschöpfungsprobleme bei repräsentativen Telefonumfragen. S. 386-399 in: Forschungsgruppe Telekommunikation (Hrsg.), Telefon und Gesellschaft, Band 2. Berlin: Volker Spiess.
- Koch, A., 1997: ADM-Design und Einwohnermelderegister-Stichprobe. Stichprobenverfahren bei mündlichen Bevölkerungsumfragen. S. 99-116 in: S. Gabler/J. Hoffmeyer-Zlotnik (Hrsg.), Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.
- Reuband, K.-H./Blasius, J., 1996: Face-to-face-, telefonische und postalische Befragungen. Ausschöpfungsquoten und Antwortmuster in einer Großstadt-Studie. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48: 296-318.
- Schulte, W., 1997: Telefon- und Face-to-Face-Umfragen und ihre Stichproben. Allgemeine Bevölkerungsumfragen in Deutschland. S. 148-195 in: S. Gabler/J. Hoffmeyer-Zlotnik (Hrsg.), Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.
- Statistisches Bundesamt (Hrsg.), 1998: Statistisches Jahrbuch 1998 für die Bundesrepublik Deutschland. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt, 1998: Europäisches Haushaltspanel. Online im Internet. URL: <http://194.95.119.6/download/panel/panel1.htm>. (Stand 06.08.1999)
- Statistisches Bundesamt, 1999: Bevölkerung nach Geschlecht und Staatsangehörigkeit. Online im Internet. URL: <http://www.statistik-bund.de/basis/d/bevoe/bevoetab1.htm>. (Stand 06.08.1999).
- Thornberry, O./Massey, J., 1988: Trends In United States Telephone Coverage Across Time and Subgroups. S. 25-50 in: R. Groves et al., (Hrsg.), Telephone Survey Methodology. New York: John Wiley & Sons.
- Trewin, D./Lee, G., 1988: International Comparison of Telephone Coverage. S. 9-24 in: R. Groves et al., (Hrsg.), Telephone Survey Methodology. New York: John Wiley & Sons.
- von der Heyde, C., 1997: Random-Route und Telefon. Struktur von Telefonhaushalten. S. 196-206 in: S. Gabler/J. Hoffmeyer-Zlotnik (Hrsg.), Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.

MESSUNG VON INTERNEN UND EXTERNEN KONTROLLÜBERZEUGUNGEN IN ALLGEMEINEN BEVÖLKERUNGSUMFRAGEN

NINA JAKOBY UND RÜDIGER JACOB

Das Konzept der internen und externen Kontrollüberzeugungen wurde von J. B. Rotter 1966 in die Diskussion eingeführt und in der Folge von verschiedenen Autoren weiterentwickelt. Leider haben alle vorgelegten Operationalisierungen den gravierenden Nachteil, sehr zeitintensiv zu sein, da zur Messung 20 Items oder mehr verwendet werden. Der Einsatz in allgemeinen Bevölkerungsumfragen verbietet sich deshalb häufig schon aus Kostengründen. Problematisch ist zudem, daß diese Instrumente, die häufig ohnehin recht lange Befragungsdauer verlängern würden, was die Teilnahmebereitschaft beeinträchtigen kann. Da Kontrollüberzeugungen aber für eine Reihe von Merkmalen wichtige erklärende Variablen darstellen, haben wir zwei Kurzskalen mit je drei Items entwickelt. Die Items wurden 1995 und 1996 im Rahmen des ZUMA-Sozialwissenschaften-BUS getestet. Die Kurzskalen haben sich als hinreichend reliabel erwiesen (PCA und Cronbachs Alpha). Korrelationen zur Konstruktvalidierung fallen zwar nur mäßig stark aus, gehen aber in theoretisch erwartbare Richtungen, so daß wir die Instrumente auch als ausreichend valide einstufen.

The concept of an internal or external locus of control was introduced by J. B. Rotter in 1966 and modified by several other researchers in the interim. Unfortunately, all the available instruments consist of 20 items or more to measure the locus of control and are thus very time-consuming. Using them in general social surveys is often impossible not only because of high costs, but because the extra time needed can increase unit non-response – people are perhaps less willing to participate. Nonetheless, the locus of control is often an important predicting variable. We therefore developed two short scales with three items each. The items were tested in 1995 and 1996 as a part of the SOWIBUS omnibus survey (ZUMA-Sozialwissenschaften-BUS). The scales proved sufficiently reliable (PCA, Cronbach's Alpha). Correlations to prove construct validity are only moderate but produce theoretical expected directions. We therefore consider the instruments to be sufficiently valid.

1. Zum Konzept der Kontrollüberzeugungen

Das Konzept der internen und externen Kontrollüberzeugungen (IEC) wurde von J. B. Rotter 1966 in die lerntheoretische Persönlichkeitspsychologie eingeführt, um das zielgerichtete Verhalten von Personen nicht ausschließlich als Funktion eines spezifischen Bedürfnisses zu erklären. Nach Rotters Sozialer Lerntheorie bilden Personen generalisierte Kontrollerwartungen über ihre Fähigkeiten aus, die Ereignisse in ihrem Leben selbst bestimmen zu können. Dabei bildet die Erwartung der Konsequenzen des eigenen Verhaltens eine zentrale unabhängige Variable zur Erklärung von Handlungen bzw. Entscheidungen. Interne Kontrollerwartung als Persönlichkeitsmerkmal liegt vor, wenn Personen erwarten, daß sie durch ihr eigenes Verhalten den Verlauf ihres Lebens bestimmen und steuern können. Das Vertrauen in eigene „Fähigkeiten“ und „Anstrengungen“ ist vorhanden und wird für die kausale Erklärung von Umweltereignissen herangezogen. Personen mit einer externen Kontrollerwartung glauben dagegen, daß ihr eigenes Verhalten die Ereignisse in ihrem Leben kaum oder gar nicht beeinflussen kann. Dementsprechend dominieren Erklärungskonzepte wie „Zufall“, „Glück“, und „die sozialen Verhältnisse“ die Erwartungen, und Ereignisse, die die eigene Person betreffen, werden auf externe Ursachen in der Umwelt zurückgeführt (Rotter 1966).

In allgemeinen Bevölkerungsumfragen stellen die Persönlichkeitskonstrukte der internen und externen Kontrollüberzeugung wichtige erklärende Variablen dar, wenn mehr oder weniger starke Zusammenhänge zwischen dem persönlichen Verhalten und Ereignissen in der Umwelt vermutet werden, z. B. im medizinisch-soziologischen Forschungsbereich für das Verhalten unter Unsicherheit oder präventives Verhalten (vgl. hierzu Jacob 1995).

2. Messung von Kontrollüberzeugungen

Zur Messung von Kontrollüberzeugungen werden in den meisten Forschungsarbeiten Fragebogenverfahren angewandt. Unterschieden wird dabei zwischen eindimensionalen Modellen von Kontrollüberzeugungen mit den Endpunkten „internal vs. external“ und multidimensionalen Modellen von Kontrollüberzeugungen. Multidimensionalität liegt dann vor, wenn inhaltlich homogene Items sich auf einen speziellen Verhaltens- oder Situationsbereich des Lebens beziehen (zu dieser Bereichsspezifität von Kontrollüberzeugungen vgl. Krampen 1982: 98 f.).

Der von Rotter 1966 veröffentlichte Fragebogen zur Messung interner und externer Kontrollüberzeugung basiert auf Vorarbeiten von James (1957) und Phares (1957). Der IEC-Fragebogen besteht aus 29 Items, von denen sechs sogenannte Füller-Items sind, die nicht ausgewertet werden. Der Antwortmodus ist die forced-choice-Methode mit jeweils einer

internen und externen Antwortalternative. Die Test-Retest-Korrelation des IEC-Fragebogens lag nach unabhängigen Untersuchungen mit mindestens einmonatigem Abstand und einer Stichprobe von mindestens 50 mehrere Male über .70. Häufiger Kritikpunkt ist jedoch die Korrelation des IEC-Fragebogens mit dem Maß sozialer Erwünschtheit (gemessen mit der Marlowne-Crowne-Social-Desirability-Scale), obwohl sie selten einen Wert über .35 erreicht (vgl. Mielke 1982a: 24 ff.). Allerdings stellt die Präferenzierung sozial erwünschter Antworten bei allen Fragebögen zur Erfassung interner und externer Kontrollüberzeugung ein Problem dar. Persönlichkeitsmerkmale, die zusammenfassend als „interne Kontrollüberzeugungen“ bezeichnet und durch Fragen nach der Verantwortungsbereitschaft, der Durchsetzungsfähigkeit oder der Leistungsorientierung indiziert werden, gelten in Wettbewerbsgesellschaften typischerweise als wünschens- und erstrebenswert. Dagegen kann das klare Bekenntnis zu Eigenschaften, die auf externe Kontrollüberzeugungen schließen lassen - Zurechnung von Erfolg auf Glück oder Schicksal anstatt auf individuelle Leistung, Delegation von Verantwortung an andere, Fatalismus - deklassierend und stigmatisierend wirken.

Rotter sieht seinen Fragebogen als ein eindimensionales Meßinstrument für interne und externe Kontrollüberzeugungen an, da sich bereits in den ersten Faktoranalysen ein Hauptfaktor ergeben hat, der den größten Teil der Varianz erklärt. Der IEC-Fragebogen will das locus of control-Konstrukt auf einer generellen Ebene erfassen, unabhängig von spezifischen Lebensbereichen. Unberücksichtigt blieben jedoch eine Reihe kleinerer Faktoren, die neben dem Generalfaktor jeweils nur wenige Items enthalten. Es wurde keine Notwendigkeit gesehen, aufgrund dieser Tatsache die Bildung von Subskalen vorzuschlagen. Die Items des IEC-Fragebogens beziehen sich auf unterschiedliche Situationen. Auch der Bedeutungsinhalt der Externalität bezieht sich auf unkontrollierbare Faktoren wie das Schicksal, Glück oder den Zufall, auf eine allgemeine Schwierigkeit und Komplexität der Welt oder mächtige andere Personen, die die Eindimensionalität des Fragebogens anzweifeln lassen. Eine spätere Faktoranalyse anderer Autoren (Gurin, Gurin, Lao und Beattie 1969) belegt eine Mehrdimensionalität des IEC-Fragebogens. Sie umschreiben die Faktoren zum einen mit „persönlicher Kontrolle“, und zum anderen mit „Kontroll-Ideologie“ bzw. „politische Kontrolle“. Dieses Ergebnis wurde durch spätere Analysen mehrfach bestätigt (vgl. Rost-Schaude 1982: 158 f.).

Im weiteren Verlauf sind eine Reihe von Meßinstrumenten im englisch- und deutschsprachigen Raum zur Erfassung der internen und externen Kontrollüberzeugung entwickelt worden, die zum Teil auf die Items des IEC-Fragebogens zurückgreifen, aber auch neue Items beinhalten.

Aus der Kritik am IEC-Fragebogen ist u.a. der ANS-IE (Adult Nowicki-Strickland-I-E-scale) für Erwachsene von Nowicki & Duke (1974) veröffentlicht worden, dem bekanntesten Instrument nach dem von Rotter entwickelten Fragebogen. Kritikpunkte an Rotters Fragebogen waren u.a. neben dem bereits erwähnten potentiellen Einfluß von sozialer Erwünschtheit infolge des forced-choice-Antwortmodus die mangelnde thematische Breite und die von Rotter behauptete Eindimensionalität des Fragebogens.

Bereits 1973 entwickelten Nowicki & Strickland einen Fragebogen für Kinder von 8 bis 18 Jahren, den sogenannten CNS-IE. Beide Fragebögen sind von Schneewind (1976) ins Deutsche übersetzt worden. Die deutschsprachige Form des ANS-IE wurde um 14 Items gekürzt und enthält 26 Items einschließlich sechs Füller-Items (vgl. Mielke 1982a: 33). Eine verkürzte Fassung des IEC-Fragebogens wurde 1974 von Salomone unter dem Namen SAIES veröffentlicht. Für diesen wurden 15 Items mit den höchsten Item-Gesamtpunktwerten ausgewählt und die 15 forced-choice-Items in 30 Items mit den Antwortvorgaben „richtig-falsch“ umgewandelt. Außerdem wurden die sechs Füller-Items weggelassen. Der entscheidende Vorteil des SAIES liegt in seiner Kürze und den geänderten Antwortvorgaben (vgl. Mielke 1982a: 29 f.).

Zur Erhebung von internen und externen Kontrollüberzeugungen im deutschsprachigen Raum wird zum einen die deutsche Übersetzung des IEC Fragebogens von Rost-Schaude (1975) verwendet. Dieses Instrument ist im ZUMA-Informationssystem (Glöckner-Rist/Schmidt 1999) dokumentiert und so der breiteren Fachöffentlichkeit gut zugänglich (vgl. dazu Rost-Schaude u.a. 1999).

Einen anderen - multidimensionalen - Ansatz zur Messung von Kontrollüberzeugungen stellt der Fragebogen von Levenson (1972) dar. Levenson schlägt eine differenzierte Erfassung des Persönlichkeitsmerkmals „interne und externe Kontrollüberzeugung“ vor. Er entwickelt einen Fragebogen mit drei Teilskalen zur persönlichen Kontrolle, der Kontrolle durch wichtige Andere und der Kontrolle durch Zufall (IPC-Fragebogen, vgl. dazu Mielke 1979: 6). Die Externalität wird mit diesem Meßinstrument in die Faktoren „Glaube an Zufall“ und „Glaube an mächtige Andere“ nochmals untergliedert. Es zeigte sich, daß Personen mit externer Kontrollüberzeugung weniger einheitlich antworten. Die größere Variabilität ihres Antwortverhaltens veranlaßte Levenson zu dieser weiteren Differenzierung der Externalität. Personen können nun danach unterschieden werden, ob sie glauben, daß mächtige andere Personen Kontrolle über ihre Verhaltensergebnisse haben oder ob sie auftretende Ereignisse für Zufälle halten und sich somit ihr ganzes Leben als unberechenbar und ungeordnet darstellt (vgl. Mielke 1982: 131f.).

Die drei Skalen des IPC-Fragebogens (I für „internal“, P für „powerful others“, C für „chance“) bestehen somit jeweils aus acht Items, die mit einer 6-Punkte-Likert-Skala

beantwortet werden müssen, wobei Levenson teilweise Items von Rotter übernommen hat, aber auch Items neu formulierte. Die Test-Retest-Korrelationen, gemessen im einwöchigen Abstand, lagen bei .64 für die I-Skala, .74 für die P-Skala und .78 für die C-Skala und sind somit als recht gut zu beurteilen. Die Korrelation mit dem Marlowe-Crowne-Fragebogen zur sozialen Erwünschtheit und den einzelnen Items lagen zwischen 0 und .19 (vgl. Mielke 1982: 132ff.). Der IPC-Fragebogen liegt in der deutschen Übersetzung von Mielke (1979) und Krampen (1979) vor (zu weiteren deutschsprachigen Instrumenten vgl. Krampen 1982: 113ff.). Die Variante von Krampen ist ebenfalls im ZUMA-Informationssystem dokumentiert (vgl. dazu Krampen 1999).

Leider haben all diese Instrumente den gravierenden Nachteil, sehr zeitintensiv zu sein, da zur Messung interner und externer Kontrollüberzeugung 20 Items oder mehr verwendet werden. Der Einsatz in allgemeinen Bevölkerungsumfragen verbietet sich deshalb häufig schon aus Kostengründen. Problematisch ist zudem, daß diese Instrumente, die häufig ohnehin recht lange Befragungsdauer verlängern würden, was die Teilnahmebereitschaft beeinträchtigen kann. Da Kontrollüberzeugungen aber – wie schon einleitend erwähnt – für eine Reihe von Merkmalen wichtige erklärende Variablen darstellen, haben wir zwei Kurzskalen zur Messung von allgemeinen, bereichsunspezifischen internen und externen Kontrollüberzeugungen entwickelt, die in allgemeinen Bevölkerungsumfragen verwendet werden können.

3. Kurzskalen zur Messung von Kontrollüberzeugungen

Zur Minimierung von Effekten sozialer Erwünschtheit infolge eines forced-choice-Antwortmodus, wie er bei dem Rotter-Instrument kritisiert wurde, sollten je drei Statements zu internen bzw. zu externen Kontrollüberzeugungen auf einer 5-stufigen Likert-Skala beurteilt werden. Dies schien uns auch deshalb vorteilhafter zu sein, weil auch andere Items der jeweiligen Befragungen entsprechend skaliert wurden und die Befragten deshalb mit diesem Antwortmodus vertraut waren (vgl. Eirnbter u.a. 1997; Jacob u.a. 1997).

Die Items wurden erstmals im Sommer 1995 im Rahmen einer ZUMA-Sozialwissenschaften-BUS-Befragung getestet und im Herbst 1996 – wiederum bei einer ZUMA-Sozialwissenschaften-BUS-Befragung – repliziert.¹

¹ Die Befragung 1995 fand statt im Rahmen des Forschungsprojektes „AIDS und die gesellschaftlichen Folgen“ und ist näher dokumentiert in Jacob u. a. 1997. Die Befragung 1996 wurde während des Forschungsprojektes „Krankheit und Gesellschaft“ durchgeführt. Näheres dazu findet sich bei Eirnbter u.a. 1997.

Es ergaben sich folgende prozentuale Verteilungen²:

Tabelle 3.1 Items zu internen und externen Kontrollüberzeugungen (in %)

Item 1. Wert: 1995/2. Wert: 1996	stimme sehr zu	stimme eher zu	teils-teils	stimme eher nicht zu	stimme überhaupt nicht zu
Ich übernehme gern Verantwortung	28,0/21,3	34,9/39,8	28,5/28,4	6,8/8,8	1,8/1,7
Es hat sich für mich als gut erwiesen, selbst Entscheidungen zu treffen, anstatt mich auf das Schicksal zu verlassen	39,2/37,2	41,7/42,3	16,4/16,6	2,3/3,5	0,4/0,5
Bei Problemen und Widerständen finde ich in der Regel Mittel und Wege, um mich durchzusetzen	26,4/25,7	40,4/45,7	29,0/24,4	3,5/3,7	0,6/0,5
Erfolg ist oft weniger von Leistung, sondern vielmehr von Glück abhängig	8,3/9,2	24,6/26,5	40,3/34,8	20,8/21,7	6,0/7,8
Ich habe häufig das Gefühl, daß ich wenig Einfluß darauf habe, was mit mir geschieht	4,5/6,4	15,6/22,0	30,1/27,0	32,9/29,2	17,0/15,3
Bei wichtigen Entscheidungen orientiere ich mich oft an dem Verhalten von anderen	4,5/5,6	18,5/20,8	37,5/32,0	29,1/27,7	10,3/13,9

Die interne Konsistenz wurde mit Hauptkomponentenanalysen (Kaiser-Kriterium, Varimax-Rotation) und der Berechnung von Cronbachs Alpha geprüft. Für die Datensätze aus beiden Umfragen wurden jeweils zwei Faktoren extrahiert, auf denen die Faktorladungen der Items in 1995 und 1996 sehr ähnlich sind.

Die durch beide Faktoren erklärte Varianz betrug 1995 59 Prozent und 1996 58 Prozent. Für die Skalierung wurde der jeweilige Skalenwert (zwischen 1 und 5) jedes beantworteten Statements für jeden Befragten addiert, durch die Gesamtzahl der Antworten geteilt und gerundet. Die so konstruierten Skalen haben damit die gleiche Dimension wie die zugrundeliegenden Statements und variieren ebenfalls zwischen den Werten 1 und 5.

² Verwendet wurden jeweils die ungewichteten Datensätze. N 1995= 2988, N 1996 = 3132.

Tabelle 3.2 Rotierte Faktorladungen der Items für Interne Kontrollüberzeugungen

Item	1995	1996
1. Wert: Faktorladungen auf Faktor 1 2. Wert: Faktorladungen auf Faktor 2		
Ich übernehme gern Verantwortung	.814/-.107	.751/-.086
Es hat sich für mich als gut erwiesen, selbst Entscheidungen zu treffen, anstatt mich auf das Schicksal zu verlassen	.778/-.111	.716/-.139
Bei Problemen und Widerständen finde ich in der Regel Mittel und Wege, um mich durchzusetzen	.761/-.105	.772/-.102

Erklärte Varianz Faktor 1: 1995: 38,4 Prozent; 1996: 37,3 Prozent.

Tabelle 3.3 Faktorladungen der Items für Externe Kontrollüberzeugungen

Item	1995	1996
1. Wert: Faktorladungen auf Faktor 2 2. Wert: Faktorladungen auf Faktor 1		
Erfolg ist oft weniger von Leistung, sondern vielmehr von Glück abhängig	.748/.048	.778/-.007
Ich habe häufig das Gefühl, daß ich wenig Einfluß darauf habe, was mit mir geschieht	.720/-.243	.787/-.154
Bei wichtigen Entscheidungen orientiere ich mich oft an dem Verhalten von anderen	.717/-.152	.692/-.184

Erklärte Varianz Faktor 2: 1995: 20,9 Prozent; 1996: 20,7 Prozent.

Für beide Skalen (Interne und Externe Kontrollüberzeugung) haben sich auch akzeptable Alpha-Koeffizienten ergeben. Den Anforderungen nach Reliabilität und der internen Konsistenz der Skalen wird demnach mit zufriedenstellender Qualität entsprochen. Lediglich der Alpha-Koeffizient der Skala „Externe Kontrollüberzeugung“ (1995) ist im Vergleich zu den anderen Werten mit .581 schwächer ausgefallen. Allerdings führte weder in diesem noch in den anderen Fällen das Weglassen einzelner Items zu einer Erhöhung der Alpha-Koeffizienten.

Tabelle 3.4 Skala „Interne Kontrollüberzeugungen“

	1995	1996
stimme sehr zu	24,7	19,9
stimme eher zu	49,8	55,9
teils-teils	23,1	21,3
stimme eher nicht zu	2,1	2,8
stimme überhaupt nicht zu	0,3	-
N	2975	3130

Tabelle 3.5 Skala „Externe Kontrollüberzeugungen“

	1995	1996
stimme sehr zu	1,7	2,4
stimme eher zu	16,3	21,8
teils-teils	47,4	41,5
stimme eher nicht zu	30,4	28,8
stimme überhaupt nicht zu	4,1	5,6
N	2979	3132

Tabelle 3.6 Alpha-Werte der Skalen

Skala	1995	1996
interne Kontrollüberzeugung	.706	.623
externe Kontrollüberzeugung	.581	.641

Der vergleichsweise niedrigere Alpha-Koeffizient für die Skala „Externe Kontrollüberzeugungen“ des 95er Datensatzes wie auch die in den Tabellen 3.4 und 3.5 ausgewiesenen Verteilungen der Skalen zeigen, daß trotz der Vermeidung eines forced-choice-Antwortmodus bei Fragen nach Kontrollüberzeugungen ein Meßproblem besteht, welches unserer Ansicht nach mit dem Social-Desirability-Response-Set und Aspekten der Selbstrepräsentation zusammenhängt. Eine Tendenz zu sozial wünschenswertem Antwortverhalten scheint sich auch bei Ratings nicht vermeiden zu lassen und führt zu asymmetrischen Verteilungen: Während die Befragten Indikatoren zu internen Kontrollüberzeugungen in beiden Befragungen mit großer Mehrheit zustimmen, ergibt sich bei der Skala zur Mes-

sung externer Kontrollüberzeugungen eine sehr auffällige Präferenz für die Mittelkategorie „teils-teils“.³

Dies hat eine nur mittelstarke Interkorrelation der Skala „Interne Kontrollüberzeugungen“ mit der Skala „Externe Kontrollüberzeugungen“ (Korrelationskoeffizient Gamma) zur Folge. Die Interkorrelation beträgt für den Datensatz aus 1995 -.38 und für 1996 -.36.

Tabelle 3.7 Korrelationen (Gamma)

	Interne Kontrollüberzeugung	Externe Kontrollüberzeugung
Bildungsstatus (1995)	-.253	.269
Bildungsstatus (1996)	-.266	.219
AIDS als Risiko (1995)	.231	-.196
AIDS als Gefahr (1995)	-.188	.341
Krankheiten als Gefahr (1996)	-.112	.346
Große Bedrohung durch Infektionskrankheiten (1996)	-.191	.405

Codierung: Bildung: 1 = kein Abschluß, 4 = Fachhochschulreife/Abitur; übrige Konstrukte: 1 = stimme sehr zu, 5 = stimme überhaupt nicht zu.

Korrelationen der Skalen mit inhaltlich ähnlichen Meßinstrumenten und der Frage nach dem höchsten allgemeinbildenden Schulabschluß liefern zwar ebenfalls nur mäßige bis mittelstarke Korrelationen,⁴ gehen aber in die erwartete Richtung. 1995 wurden die Skalen mit ebenfalls Likert-basierten Indikatoren zur Interpretation von AIDS als vermeidbarem Risiko oder als schicksalhafter Gefahr korreliert, 1996 mit Items zur Interpretation von Krankheiten generell als schicksalhafte, nicht zu vermeidende Gefahren und der perzipierten Bedrohlichkeit von Infektionskrankheiten. Personen mit höherem Bildungsstatus zeigen eine stärkere Präferenz für Items interner Kontrollüberzeugung bei ausgeprägter Ablehnung der Statements zur Messung externer Kontrollüberzeugung. Personen mit niedrigem Bildungsstatus verhalten sich genau umgekehrt. In analoger Weise präferie-

-
- 3** Anzumerken bleibt hier allerdings, daß dies eine Interpretation der Ergebnisse ist, die wir nicht explizit getestet haben, weil entsprechende Instrumente zur Analyse sozial-wünschenswerten Antwortverhaltens in den jeweiligen Surveys nicht verwendet werden konnten, da jeweils inhaltliche und weniger methodentestende Aspekte bei der Fragebogenkonstruktion im Vordergrund standen.
- 4** Dabei sollte berücksichtigt werden, daß Korrelationen über Werten von .3 in sozialwissenschaftlichen Survey-Untersuchungen aufgrund von grundsätzlichen Operationalisierungsproblemen und der Schwierigkeit bzw. Unmöglichkeit der Konstruktion von isolierten Systemen, mit denen man einfache Kausalzusammenhänge konstruieren kann, eher selten sind.

ren Personen, die AIDS als individuell vermeidbares Risiko ansehen, interne Kontrollüberzeugungen während Befragte, die in AIDS eine allgegenwärtige schicksalhafte Gefahr sehen (1995) oder die Bedrohung durch Infektionskrankheiten als hoch einschätzen (1996) ihr Leben auch eher als fremdbestimmt ansehen und Indikatoren zu externen Kontrollüberzeugungen vergleichsweise häufig zustimmen. Zumindest hinsichtlich der Richtung der Zusammenhänge können diese Korrelationen unserer Ansicht nach mithin als zufriedenstellende Konstruktvalidierung eingestuft werden.

4. Schlußbemerkung

Die Kurzskalen haben sich als hinreichend reliabel erwiesen. Korrelationen zur Konstruktvalidierung fallen zwar nur mäßig stark aus, gehen aber in theoretisch erwartbare Richtungen, so daß wir die Instrumente auch als ausreichend valide einstufen. Die Skalen können damit in allgemeinen Bevölkerungsumfragen durchaus eingesetzt werden. Allerdings weist die starke Präferenz für die Mittelkategorie „teils-teils“ bei Items zu externen Kontrollüberzeugungen darauf hin, daß hier Antworten in Richtung sozialer Erwünschtheit abgegeben wurden. Einschränkend ist deshalb festzuhalten, daß das Problem der Antwortverzerrung durch sozial wünschenswertes Antwortverhalten auch durch die hier vorgestellten Skalen nicht zufriedenstellend gelöst werden konnte. Dies könnte darauf zurückzuführen sein, daß die verwendeten Items zu externen Kontrollüberzeugungen unzureichend operationalisiert sind. Wünschenswert wäre hier eine Anschlußuntersuchung, in der auch andere Operationalisierungen getestet und die Skalen mit Instrumenten zur Messung sozial wünschwertem Antwortverhaltens korreliert werden könnten.

Korrespondenzadresse

*Dr. Rüdiger Jacob
Universität Trier
FB IV Soziologie
54286 Trier*

Literatur

Eirmbter, W.H./Hahn, A./Hennes, C./Jacob, R./Lettke, F., 1997: Krankheit und Gesellschaft. Zur Bedeutung von Krankheitsbildern und Gesundheitsvorstellungen für die Prävention. Forschungsbericht an die DFG, Trier.

Glöckner-Rist, A./Schmidt P. (Hrsg.), 1999: ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente, Version 3.00. Mannheim: ZUMA.

- Gurin, P./Gurin, G./Lao, R./Beattie, C., 1969: Internal-external control in the motivational dynamics of negro youth. *Journal of Social Issues* 25: 25-53.
- Jacob, R., 1995: Krankheitsbilder und Deutungsmuster. Wissen über Krankheit und dessen Bedeutung für die Praxis. Opladen: Westdeutscher Verlag.
- Jacob, R./Eirmbter, W. H./Hahn, A./Hennes, C./Lettke, F., 1997: AIDS-Vorstellungen in Deutschland. Stabilität und Wandel. Berlin: Edition Sigma.
- James, W. H., 1957: Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation, Ohio State University.
- Krampen, G., 1982: Differentialpsychologie der Kontrollüberzeugungen. Göttingen: Hogrefe.
- Krampen, G., 1999: Interne/Externe Kontrolle, in: Glöckner-Rist, A./Schmidt P. (Hrsg.): ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente, Version 3.00. Mannheim: ZUMA.
- Levenson, H., 1972: Distinction within the concept of internal-external control. Development of a new scale. S. 261-262 in: Proceedings of the 80th Annual Conception of the APA 7.
- Mielke, R., 1979: Entwicklung einer deutschen Form des Fragebogens zur Erfassung interner vs. externer Kontrolle von Levenson (IPC). Bielefelder Arbeiten zur Sozialpsychologie, Nr. 49.
- Mielke, R., 1982a: Locus of control - Ein Überblick über den Forschungsgegenstand. S. 15-42 in: Mielke, R., (Hrsg.), Interne/externe Kontrollüberzeugung: theoretische und empirische Arbeiten zum Locus of control-Konstrukt. Bern: Hans Huber.
- Mielke, R., 1982: Das Problem der Mehrdimensionalität bei der Erfassung interner/externer Kontrollüberzeugung. S. 119-138 in: Mielke, R., (Hrsg.), Interne/externe Kontrollüberzeugung: theoretische und empirische Arbeiten zum Locus of control-Konstrukt. Bern: Hans Huber.
- Nowicke, S./Duke, M. P., 1974: A locus of control scale for noncollege as well as college adults. *Journal of Personality* 38: 136-137.
- Phares, E. J., 1957: Expectancy changes in skill and chance situations. *Journal of Abnormal Social Psychology* 54: 339-342.
- Rost-Schaude, E., 1982: Untersuchungen zu einer deutschen Form des IEC-Fragebogens von Rotter. S. 156-177 in: Mielke, R., (Hrsg.), Interne/externe Kontrollüberzeugung: theoretische und empirische Arbeiten zum Locus of control-Konstrukt, Bern: Hans Huber.
- Rost-Schaude, E./Kumpf, M./Frey D., 1999: Interne und externe Kontrolle, in: Glöckner-Rist, A./Schmidt P., (Hrsg.), ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente, Version 3.00. Mannheim: ZUMA.
- Rotter, J. B., 1966: Generalized expectancies for internal vs. external control of reinforcement. *Psychological Monographs: General and Applied* 80.

Schneewind, K. A., 1974: Entwicklung eines Fragebogens zur Erfassung internaler versus externaler Kontrollüberzeugungen bei Erwachsenen (Loc-E). Arbeitsbericht 15 aus dem EKB Projekt der Universität Trier.

THEMATIK ODER INCENTIVES? ZUR ERHÖHUNG DER RÜCKLAUFQUOTEN BEI POSTALISCHEN BEFRAGUNGEN

ROLF PORST

Im Rahmen einer postalischen Befragung sollte der Einfluß aktueller und interessanter Einstiegsfragen auf die Teilnahmebereitschaft überprüft werden. Zugleich wurde die Wirkung von incentives auf die Ausschöpfung bei postalischen Befragungen untersucht. Bei einem insgesamt schwachen Rücklauf, von dem nicht auszuschließen ist, daß er durch die Befragungszeit über den Jahreswechsel 1998/99 bewirkt worden ist, zeigt sich, daß sowohl ein incentive in Form von Briefmarken wie auch „interessante“ Einstiegsfragen die Ausschöpfungsquote bei schriftlichen Befragungen steigern können.

A mail survey was used to investigate whether questions about interesting current issues at the beginning of a questionnaire could enhance response rates. The study also investigated the impact of incentives on response rates. The low overall response rate obtained may be related to the time of mailings – over the Christmas and New Year period, 1998/1999. The findings nevertheless indicate that both interesting introductory questions and the use of incentives can increase response rates.

1. Einleitung

Im Rahmen einer postalischen Befragung, die zur Jahreswende 1998/99 von ZUMA in vier deutschen Städten durchgeführt worden ist („Vier-Städte-Befragung“), sollte der Einfluß *aktueller* und *interessanter Einstiegsfragen* auf die Teilnahmebereitschaft überprüft werden. Zugleich wurde ein weiteres Mal die Wirkung von *incentives* auf die Ausschöpfung bei postalischen Befragungen untersucht.

2. Zur Ausschöpfung bei postalischen Befragungen

Als ein wesentliches Problem postalischer Befragungen benennt Hippler (1988: 244) „die außerordentlich große Schwankungsbreite der Rücksenderate, die zwischen zehn und 90

Prozent liegen kann“. Hippler selbst berichtet von einer postalischen Umfrage bei der erwachsenen Wohnbevölkerung Mannheims, bei der er unter Anwendung der „total design method“ (vgl. Dillman 1978) eine Ausschöpfung von 78 Prozent erzielen konnte (Hippler 1985).

Verglichen mit persönlich-mündlichen Befragungen werden von postalischen Befragungen insgesamt aber eher niedrigere Ausschöpfungsquoten berichtet; Yu/Cooper (1983) berechnen in einem Übersichtsartikel eine gewichtete durchschnittliche Ausschöpfungsquote von 47,3 Prozent bei postalischen Befragungen, von 72,3 Prozent bei telefonischen und von 81,7 Prozent bei persönlich-mündlichen Befragungen¹. Solche Ergebnisse wurden früher unter anderem mit einem Phänomen zu erklären versucht, das mit dem Schlagwort „Bildungsbarriere“ umschrieben worden ist: die geringe Rücksendebereitschaft durch unwillige, vor allem aber unfähige Zielpersonen. In den letzten Jahren hat sich diese Denkweise allerdings etwas verändert, und man beschäftigt sich nun auch zunehmend mit technischen Unzulänglichkeiten in der Vorgehensweise bei schriftlichen Befragungen.

Wenngleich es doch sehr lange dauerte, bis sich Überlegungen dieser Art auf die Studiendesigns postalischer Befragungen auswirken sollten, hatte Dillman (1978) dies bereits sehr früh erkannt und versucht, die Rücklaufquoten bei postalischen Befragungen durch vorwiegend technische Lösungsvorschläge zu erhöhen. Sein Ansatz, „total design method“ genannt, basiert auf der Idee, der potentiellen Befragungsperson deutlich zu machen, daß die Vorteile (der Nutzen), die aus einer Teilnahme an der Befragung resultieren, die damit verbundenen Nachteile (Aufwand und Kosten, wie z.B. die dafür zu investierende Zeit) überwiegen. Im Mittelpunkt steht dabei die Kommunikation zwischen Forscher und Zielperson durch Inhalte und Form des Befragungsmaterials: wenn es gelingt, die Zielperson davon zu überzeugen, daß ihr die Teilnahme an der Befragung „mehr nutzt als schadet“, wird sie den Fragebogen ausfüllen und zurückschicken – und damit zu einer Erhöhung der Ausschöpfungsquote beitragen.

Da die Kommunikation mit den zu befragenden Personen nur über das Instrument selbst und damit in Zusammenhang stehende Maßnahmen geleistet werden kann, hat sich eine relativ umfangreiche Forschung mit der Frage beschäftigt, wie Aspekte des Befragungs-

¹ Auch wenn diese Zahlen auf deutliche Unterschiede zwischen den Befragungsmodi hinweisen, ist doch grundsätzlich eine gewisse Skepsis angebracht, wenn man Ausschöpfungsquoten in derart globaler Darstellung zur Kenntnis erhält. Weil die Berechnung von Ausschöpfungsraten oft nicht nachvollziehbar dargestellt ist und relevante Sachverhalte in unterschiedlichen Quellen oft unterschiedlich behandelt werden, kann man nicht in jedem Falle ausschließen, daß „Birken mit Äpfeln verglichen“ werden (zur Problematik der Ausschöpfungsberechnung siehe Porst 1996: 4-6).

materials so zu gestalten sind, daß die Motivation der angeschriebenen Personen zur Teilnahme an der Befragung und dadurch die Ausschöpfungsrate erhöht wird.

Als außerordentlich wirksame Strategie zur Erhöhung der Ausschöpfungsraten bei postalischen Befragungen erweisen sich – ganz im Sinne der Dillman'schen total design method – wiederholte Kontakte (Nachfaßaktionen) mit den zu befragenden Personen, wobei mit jedem weiteren Kontakt der Stil der Ansprache (zunehmende Pointierung der negativen Folgen individueller Nicht-Teilnahme für das Erreichen des Studienziels) oder/und die Form (bis hin zu eingeschriebenen Briefen) dramatisiert werden. Jede zusätzliche Kontaktierung führt zu einer erheblichen Verbesserung der Rücklaufquote. Als „Faustregel“ gilt, daß pro Nachfaßaktion ein Anteil an den bisher noch ausstehenden Personen realisiert wird, der so groß ist wie der Anteil der Teilnehmer an der ersten Aussendung (vgl. Hippler 1988: 245f.).

Den unbestrittenen Nutzen von Mehrfachkontakten versucht man sich auch beim Einsatz von *Vorankündigungen* dienstbar zu machen. Anstatt direkt mit der Tür ins Haus (mit dem Fragebogen in den Briefkasten) zu fallen, baut man auf die positive Wirkung einer Vorabankündigung. Leider besteht in der einschlägigen Literatur keine Übereinstimmung darüber, ob eine Vorabankündigung wirklich die gewünschte Wirkung zeitigt. Während auf der einen Seite gezeigt werden konnte, daß eine (schriftliche) Vorabankündigung den Rücklauf bei mail surveys signifikant erhöht (Linsky 1975; Fox et al. 1988; Conant/Smart/Walker 1990; Yammarino et al. 1991²), werden auch Studien berichtet, bei denen durch eine Vorabankündigung keine nennenswerte Steigerung des Rücklaufes zu verzeichnen war (Jones/Lang 1980; Tedin/Hofstetter 1982).

Aber selbst dort, wo eine Vorabankündigung zu einer erhöhten Rücklaufquote geführt hatte, findet sich Gegensätzliches hinsichtlich der Vorgehensweise: Während Autoren wie Linsky (1975), Fox et al. (1988) oder Conant/Smart/Walker (1990) die besondere Wirkung einer *schriftlichen* Vorabankündigung betonen, die in jedem Falle effektiver sei als eine telefonische oder gar keine, kommen andere Studien (Allen/Schewe/Wijk 1980; Dillman/Frey 1974; Hansen 1980; Kerin/Peterson 1977; Walker/Burdick 1977) genau zu gegenteiligem Schluß: Gerade die *telefonische* Vorabankündigung habe sich als besonders effektiv erwiesen (vgl. Rucker et al. 1984: 894f.).

Da aber wiederholte Kontakte egal mit welchem Medium – sei es nun durch mehrere Nachfaßaktionen in Einklang mit der total design-Methode oder durch Vorabankündigung der Befragung – aufwendig sowie zeit- und kostenintensiv sind, sucht man nach Maß-

² Die in ihren Meta-Analysen allerdings zeigen konnten, daß dieser Zusammenhang für Studien, die vor 1976 durchgeführt worden waren, nicht nachweisbar war (Yammarino et al. 1991: 628-629).

nahmen, die auch dann wirksam werden können, wenn nur ein einziger Kontakt (nämlich das Zuschicken des Fragebogens) mit den zu befragenden Personen möglich ist.

Eine Vielzahl solcher singulärer Maßnahmen ist – mehr oder weniger systematisch – untersucht worden. So wissen wir unter anderem, daß die Rücklaufquoten bei postalischen Befragungen erhöht werden können durch (vgl. Hippler 1988; Lamnek/Trepl 1991)

- a) kurze Fragebogen, die nicht mehr als zwölf Seiten umfassen sollten,
- b) offizielle Briefbogen von Universitäten oder universitären Einrichtungen,
- c) Rückrufnummern im Anschreiben, unter denen die angeschriebenen Personen Fragen stellen und zusätzliche Auskünfte zur Studie erhalten können,
- d) frankierte Rückumschläge,
- e) Ankündigungen in der örtlichen Presse bei lokalen Umfragen.

Neben diesen eher technischen Aspekten wirken sich bei der postalischen Befragung vor allem das *Thema* selbst und der *Einsatz kleinerer Geschenke* (incentives) positiv auf die Teilnahmebereitschaft aus.

3. Incentives und Thematik

3.1 Incentives

Die postalische Befragung ist diejenige Art der Datenerhebung, bei der verstärkt versucht wird, durch kleinere Anreize und Geschenke (incentives) die Teilnahmebereitschaft und damit die Rücklaufquote zu erhöhen. Im allgemeinen ist die Wirkung finanzieller Anreize und die Wirkung von Geschenken auf die Erhöhung der Teilnahmebereitschaft bei postalischen Umfragen unumstritten (Armstrong 1975; Friedman/San Augustine 1979; Goodstadt et al. 1977; Heberlein/Baumgartner 1978; James/Bolstein 1990; Linsky 1975; Paolillo/Lorenzi 1984; Schewe/Cournoyer 1976; Wotruba 1966; Yu/Cooper 1983).³ Allerdings kann man davon ausgehen, daß materielle Anreize nur dann ausschöpfungsfördernd wirken, wenn sie bereits mit dem ersten Anschreiben übermittelt werden (Church 1993; Fox et al. 1988; James/Bolstein 1990; Kanuk/Berenson 1975; Linsky 1975).

Zumeist wird – dies gilt vor allem für Umfragen in den Vereinigten Staaten – ein Zusammenhang zwischen der Höhe des monetären Incentives und der Teilnahmebereitschaft postuliert und auch nachgewiesen (dokumentiert z.B. von Church 1993); im Gegensatz dazu steht eine Erfahrung, die im deutschen Kontext gemacht worden ist, nämlich daß

³ Singer (1998) konnte den gleichen Effekt, das Ansteigen der Ausschöpfungsrate durch incentives, auch für persönlich-mündliche und telefonische Befragungen aufzeigen.

weniger der materielle Wert der Geschenke als deren symbolische Bedeutung von Interesse zu sein scheint: „Generell kann davon ausgegangen werden, daß der symbolische Akt eines Entschädigungsversuchs für die Befragtenbemühungen das Entscheidende zu sein scheint, weniger die tatsächliche Höhe der Belohnung“ (Hippler 1988: 245).

In einer Untersuchung von Blumberg, Fuller und Hare (1974) war die Rücksendequote bei Personen, die mit dem Fragebogen einen 1-Dollar-Schein erhielten, um zehn Prozent höher als bei Personen, denen diese Belohnung nicht zugekommen war; bei Hackler und Bourgett (1973) führte der 1-Dollar-Schein gar zu einer Differenz von 32 Prozent. Daß selbst noch geringere Beträge die Rücklaufquote erhöhen konnten, spricht deutlich für den symbolischen Charakter von incentives: Huck und Gleason (1974) erhöhten durch Beigabe von 25 Cent die Rücklaufquote um 27 Prozent, Wiseman (1972) erreichte mit nur zehn Cent immerhin noch ganze acht Prozent Erhöhung (vgl. Paolillo/Lorenzi 1984).

Paolillo und Lorenzi (1984) teilten 400 Zielpersonen in vier gleich große Gruppen ein. Die erste Gruppe erhielt mit der Zusendung des Fragebogens keinerlei Geschenke, die Personen in der zweiten Gruppe erhielten einen Dollar, in der dritten Gruppe zwei Dollar, und die Personen in der vierten Gruppe erhielten ein Los für eine Lotterie. Die Rücksendefrist betrug sechs Wochen; am Ende dieser Zeit waren folgende Rücklaufquoten erreicht (siehe Tabelle 1):

Tabelle 1

Gruppe	Rücklaufquote
1. kein incentive	36%
2. \$ 1.00	65%
3. \$ 2.00	41%
4. Lotterielos	33%

Wie man sieht, hatte die „1-Dollar-Gruppe“ den eindeutig besten Rücklauf zu verzeichnen, die „2-Dollar-Gruppe“ folgte erst deutlich abgeschlagen auf Platz 2. Es verdichtet sich der Eindruck, als ob es tatsächlich gar nicht die Höhe des Betrages sei, die teilnahmemotivierend wirkt. Sind es vielleicht sogar die kleineren finanziellen Anreize, die sich als ausschöpfungserhöhend erweisen?

Zu ganz anderen Ergebnissen führte eine postalische Befragung, die 1986 in den Vereinigten Staaten durchgeführt wurde (vgl. James/Bolstein 1990: 346ff.): Die Stichprobe enthielt 850 Teilnehmer aus Virginia/USA, die in fünf gleich große Gruppen eingeteilt wurden. Die erste Gruppe erhielt als Kontrollgruppe keinerlei finanzielles incentive zugesandt; bei den anderen vier Gruppen wurden Beträge von 25 Cent, 50 Cent, 1 Dollar und

2 Dollar mit den Fragebogen verschickt. Es gab insgesamt vier Nachfassaktionen, die incentives wurden aber nur mit dem Erstversand verteilt.

Beim Erstversand verhielt sich der Rücklauf proportional zu dem Wert des jeweiligen Anreizes (siehe Tabelle 2).

Tabelle 2

	kein Geld	\$ 0,25	\$ 0,50	\$ 1,00	\$ 2,00
nach dem Erstversand	54,2%	62,7%	63,1%	72,8%	77,6%

Nach dem vierten Versand war eine Gesamtrücklaufquote von 91 Prozent erreicht; die Verhältnisse der Quoten in Abhängigkeit von der Höhe des incentives hatten sich aber verändert (siehe Tabelle 3).

Tabelle 3

	kein Geld	\$ 0,25	\$ 0,50	\$ 1,00	\$ 2,00
nach dem 4. Versand	88,1%	86,4%	86,5%	92,9%	95,5%

Hier wird also die höchste Ausschöpfung in den Gruppen mit den höchsten incentives erreicht, die „2-Dollar-Gruppe“ lag noch deutlich vor der „1-Dollar-Gruppe“. Die beiden Gruppen mit den Cent-Beträgen waren dagegen ganz ans Ende der Ausschöpfungsrangliste gerutscht, lagen sogar noch hinter der Kontrollgruppe. Vielleicht fühlen sich Personen durch den geringen Geldbetrag bei gleichzeitigem massiven Nachfassen doch etwas „auf den Arm genommen“.

In einem Artikel von Allan Church (1993) wird die Wirkung von „nonmonetary incentives“ erwähnt, denn außer finanziellen Anreizen gibt es natürlich eine Vielzahl anderer Möglichkeiten, den Teilnehmern einer Erhebung kleinere Geschenke und Aufmerksamkeiten zukommen zu lassen: „The nonmonetary incentives were more interesting in their diversity, however, with such items as entry in a lottery, donations to charity, coffee, books, pens, key rings, golf balls, tie clips, stamps and even a turkey“ (Church 1993: 68). Die Hypothese, daß die Rücklaufquote bei monetary incentives größer sei als bei nonmonetary incentives, konnte nicht aufrechterhalten werden. Neben der Feststellung, daß „the relative timing of the incentive is more important than the nature or form of what is included“ (ebd.: 73) wurde deutlich, daß der Rücklauf der Fragebögen dann höher ist, wenn die Anreize gleich mit dem Fragebogen verschickt werden und nicht erst später, nach Rücksendung des Bogens (vgl. auch Gajraj/Faria/Dickinson 1990; Brennan/Hoek/Astridge 1991).

Ausgehend von dieser Untersuchung könnte also gesagt werden, daß die Rücklaufquote proportional mit dem Wert des (finanziellen) Anreizes steigt; und weiter: „the findings of the meta-analysis have demonstrated that incentives do indeed have substantial positive effects on mail survey return rates. The results of the analysis of variance indicated a significant overall effect for the use of any incentive in increasing mail survey responses....” (Church 1993: 72).

3.2 Thematik

Bei postalischen Befragungen ist mehr als bei anderen Befragungsmethoden das Thema selbst von zentraler Bedeutung für die Teilnahmebereitschaft und damit die Rücklaufquote. Beschäftigt sich der Fragebogen mit interessanten, aktuellen Themen, werden die Fragebogen schneller und vollständiger zurückgesandt. Bei Umfragen mit hohem Aufmerksamkeitswert (also einer interessanten oder aktuellen Thematik) waren bei einem systematischen Vergleich „im Durchschnitt 77% Rücklauf zu verzeichnen..., mit mittlerem Aufmerksamkeitswert 66% und mit sehr geringem Aufmerksamkeitswert lediglich 42%“ (Hippler 1988: 244).

Während die Variation von incentives oder ihrer Höhe in der Forschungspraxis (die Verfügbarkeit entsprechender Mittel natürlich vorausgesetzt) mehr oder weniger im Belieben des Forschers steht, hat dieser nun aber bei der Auswahl seines Forschungsthemas im Normalfalle weniger Entscheidungsfreiheit. Sozialwissenschaftliche unterscheiden sich von anderen Befragungen nicht selten dadurch, daß weniger die Aktualität oder die Interessantheit für die Befragungspersonen den Fragebogen bestimmen, sondern Fragestellungen, die angemessen sein sollen zur Überprüfung postulierter Zusammenhänge im Rahmen allgemeiner sozialwissenschaftlicher Theorien.

Neben diese grundsätzliche Problematik tritt das Problem, daß oft per Augenschein gar nicht erkennbar ist, was eigentlich den Aufmerksamkeitsgrad eines Themas tatsächlich ausmacht (vgl. Thoma/Zimmermann 1996: 144); diese Frage kann nur empirisch geklärt werden.

Die Studie, deren Ergebnisse im folgenden präsentiert werden, beginnt denn auch mit der Auswahl eines aktuellen und interessanten Themas. Um allerdings der Realität des Forschungsprozesses, der wie gesagt nicht immer frei ist in der Auswahl der Themen, gerecht zu werden, wurden nur die Einstiegsfragen eines Fragebogens mit ansonsten „üblichen“ sozialwissenschaftlichen Fragen variiert.

4. Die Umfragen

4.1 Der Vorläufer

Ende 1996 führte ZUMA eine postalische Befragung durch, bei der 1.296 Personen ab 18 Jahren in Deutschland zum Thema „Einstellung zu Staat und Regierung“ befragt worden sind (Harkness et al. 1998). Das Untersuchungsdesign der Studie basierte auf der total design-Methode: Vierzehn Tage nach dem Versand des Fragebogens erhielten alle angeschriebenen Personen ein kurzes Dank- bzw. Erinnerungsschreiben, weitere vierzehn Tage danach wurden die vollständigen Befragungsunterlagen und ein Erinnerungsschreiben an diejenigen Zielpersonen versandt, die bisher in keiner Weise reagiert hatten, also weder den ausgefüllten Fragebogen zurückgesandt noch sich offensiv als teilnahmeunfähig bzw. teilnahmeunwillig erwiesen hatten.

Um die Wirkung eines Incentives auf die Teilnahmebereitschaft zu prüfen, hatte die Hälfte der Befragungsteilnehmer direkt mit den Fragebogen Sonderbriefmarken als Belohnung für ihre Teilnahme erhalten, während die andere Hälfte in ihren Befragungsunterlagen kein Incentive vorgefunden hatte. Tatsächlich zeigte sich, daß der Rücklauf mit Incentive mit 35,6 Prozent signifikant höher lag als der Rücklauf ohne Incentive mit 29,3 Prozent.

Da allerdings der Rücklauf insgesamt mit 32,5 Prozent doch unerfreulich niedrig war, begannen Spekulationen über die Gründe dafür. Genannt wurden unter anderem die Tatsache, daß die Befragung über die Weihnachtsfeiertage und den Jahreswechsel durchgeführt worden war (5. Dezember 1996 bis 3. Februar 1997), aber auch Zweifel an der Interessanztheit der Thematik für die angeschriebenen Personen.

Um hier Klarheit zu schaffen wurde entschieden, den gleichen Fragebogen unter weitestgehend identischen Bedingungen (insbesondere was die Terminierung des Feldes, aber auch den Einsatz der Incentives anging) Ende 1998 noch einmal ins Feld zu schicken mit einer Variation der Einstiegsfragen: anstelle der ersten Fragen des „alten“ Fragebogens sollte in einer Splitvariante die Teilnahmebereitschaft durch die Einschaltung „aktueller“ und „interessanter“ Fragen beeinflusst werden.

4.2 Die Pilotstudie „Auswahl interessanter Themen“

Da die Entscheidung, was denn nun wirklich ein interessantes oder aktuelles Thema sein könnte, nur empirisch zu lösen schien, wurden vier Themenbereiche ausgewählt und dazu jeweils ein kleiner Fragebogen mit je sechs Fragen entwickelt. Zur Wahl standen die Themenbereiche

- Deutsche Fußball-Nationalmannschaft,
- Jahr 2000,
- Sexueller Mißbrauch an Kindern,
- Tierversuche.

Diese Themen waren so gewählt gewesen, daß sie aktuelle öffentliche Diskussionen aufgriffen (Tierversuche, sexueller Mißbrauch) oder im Sinne einer Alltagsinteressantheit für Personen attraktiv sein sollten (Jahr 2000, Fußball).

Die Fragebogen wurden 40 per Quote ausgewählten Personen zur Beurteilung vorgelegt. Die befragten Personen hatten die Möglichkeit, die Fragebogen und die darin enthaltenen Fragen genau zu studieren, bevor sie eine Rangordnung der Themen unter dem Gesichtspunkt vornehmen sollten, welches sie am ehesten zur Beantwortung und zum Zurücksenden des Fragebogens motivieren würde.

Das Ergebnis war mehr als eindeutig: 30 von 40 Befragungspersonen entschieden sich für das Thema „Sexueller Mißbrauch von Kindern“. Weit abgeschlagen die anderen Bereiche: Tierversuche mit sechs Nennungen, Jahr 2000 mit drei und Fußball mit nur einer Nennung.

Der Fragebogen der Vorbildstudie „Einstellung zu Staat und Regierung“ wurde dann in einer Splitvariante dahingehend variiert, daß die ersten sechs Fragen gestrichen und durch die sechs Fragen zum Thema „Sexueller Mißbrauch“ ersetzt wurden (diese Fragen finden sich im Anhang); die andere Hälfte der Befragungspersonen sollte den Fragebogen „Einstellung zu Staat und Regierung“ unverändert erhalten.

4.3 Die „Vier Städte-Befragung“

Abgesehen von dieser Modifikation sollte in der neuen Umfrage („Vier-Städte-Befragung“) die Vorgehensweise der alten weitestgehend repliziert werden, was mit (finanziell bedingten) Abstrichen (beim dritten Versand wurde nur noch eine Stichprobe derjenigen angeschrieben, die bisher nicht reagiert hatten) denn auch gelang.

Neben der Variation der Einstiegsfragen sollte auch in der neuen Umfrage die Wirkung des incentives (wie im Vorläufer sollten den Befragungsunterlagen beim Erstversand Briefmarken im Wert von vier Standardbriefen beigelegt werden) überprüft werden; dadurch ergab sich – bei 800 Adressen, die schriftlich kontaktiert werden sollten – folgendes 2x2-Design (siehe Tabelle 4).

Insgesamt sollten also 800 Personen befragt werden; da aus Kostengründen (Stichprobenziehung!) eine bundesweite Befragung ausschied, beschränkte sich die Umfrage auf die Städte Berlin, Ludwigshafen am Rhein, Lübeck und Meißen. Die Personenadressen wur-

den nach einem von ZUMA vorgegebenen Ziehungsverfahren durch die jeweiligen Einwohnermeldeämter ermittelt und zur Verfügung gestellt. Mit Ausnahme der Berliner Adressen sind wunschgemäß auch Alter und Geschlecht der ausgewählten Personen zur Verfügung übermittelt worden.

Tabelle 4

	Kein Incentive	Incentive
Einstiegsfragen „Staat und Gesellschaft“	Version 1 (N = 200)	Version 3 (N = 200)
Einstiegsfragen „Kindesmißbrauch“	Version 2 (N = 200)	Version 4 (N = 200)

Der Erstversand erfolgte am 4. Dezember 1998; das Erinnerungsschreiben wurde am 18. Dezember 1998 verschickt.

Am 12. Januar 1999 erfolgte dann der dritte Versand, wobei aus Kostengründen nur noch eine Stichprobe derjenigen kontaktiert worden ist, die bisher nicht reagiert hatten (zum Feldverlauf und zur Rücklaufstatistik siehe Klein/Schneid 1999).

5. Ergebnisse

Aufgrund der kostenbedingten Reduzierung der Stichprobe für den dritten Versand von 556 noch offenen auf 200 Adressen ist die Berechnung der Gesamtausschöpfung nicht möglich. Prozentuierte man – sozusagen als „Notnagel“ – den gesamten Rücklauf der Studie auf die Nettostichprobe von 772 Fällen aus den ersten beiden Versandaktionen (N = 800 minus 23 neutrale Ausfälle aus Welle 1 minus 5 neutrale Ausfälle aus Welle 2 = 772), so erhielt man eine Gesamtausschöpfung von 22,7 Prozent. Dieser Wert unterschätzt die – ohne Reduzierung der Fallzahl in Welle 3 – zu erwartende Gesamtausschöpfung.

Will man die Ausschöpfungsergebnisse unserer Studie mit denen der Vorbildstudie vergleichen, bleibt somit nur der Vergleich der Ausschöpfungen aus den ersten beiden Wellen (siehe Tabelle 5).

Damit liegt die Ausschöpfungsrates unserer Studie noch (gering) unter derjenigen der Vorbildstudie – zumindest was den Rücklauf in der ersten und zweiten Welle angeht. Diese Werte können nur als absolut schlechtes Feldergebnis gewertet werden. Auch wenn bei einer telefonischen Nachbefragung der Nichtteilnehmer (Schneid/Stiegler 1999) die Weihnachtszeit nur selten explizit als Grund für die Nichtteilnahme angegeben worden ist, kann nicht ausgeräumt werden, daß sich der spezielle Zeitraum der Befragung doch nega-

tiv auf die Teilnahmebereitschaft ausgewirkt hat; immerhin fast ein Drittel der telefonisch befragten Nichtteilnehmer hatte ganz allgemein „Zeitgründe“ für die Nichtteilnahme angegeben. Um die Frage des Einflusses der Weihnachtszeit definitiv beantworten zu können, wäre eine erneute Replikation der Befragung in einem Zeitraum erforderlich, in den weder Feiertage noch Schulferien bzw. Ferien der großen Wirtschaftsunternehmen fallen.

Tabelle 5

	Vorbildstudie (N _{netto} = 1.122)	Vier-Städte-Befragung (N _{netto} = 772)
Ausschöpfung nach der ersten Feldphase gesamt	152 (13,6%)	101 (13,1%)
Ausschöpfung nach der zweiten Feldphase gesamt	265 (23,6%)	156 (20,2%)

Wesentlich wichtiger als die Gesamtausschöpfung erscheint der Vergleich der unterschiedlichen Versionen des Fragebogens. Hier hatte sich – jetzt wieder nach allen drei Versandaktionen – folgendes Ergebnis (siehe Tabelle 6) herausgestellt (absolute Zahlen).

Tabelle 6

	Kein Incentive	Incentive
Einstiegsfragen „Staat und Gesellschaft“	Version 1 25	Version 3 51
Einstiegsfragen „Kindesmißbrauch“	Version 2 41	Version 4 58

Beide Versionen mit incentive haben einen deutlich höheren Rücklauf als die vergleichbaren Versionen ohne incentive; allerdings ist der Zusammenhang zwischen den Variablen statistisch nicht signifikant.

Betrachten wir die Wirkung von incentive bzw. von Einstiegsfragen jeweils getrennt, so stellen wir fest, daß *sowohl die incentives wie auch die Einstiegsfragen ausschöpfungssteigernd* wirken.

Erwartungsgemäß zeigt das incentive vor allem in der *ersten Feldphase* Wirkung: Nach Abschluß der ersten Feldphase lag die Ausschöpfung bei den Versionen mit incentive bei 17,9 bzw. 17,4 Prozent, während sie bei den Versionen ohne incentives mit 4,6 bzw. 12,0 Prozent deutlich niedriger war.

Demgegenüber wirkte sich der „aktuelle“ Vorspann erst in der *zweiten Feldphase* aus: Gegenüber der ersten Feldphase stieg die Ausschöpfung bei den Varianten mit Vorspann um 17,7 Prozent an, bei den Varianten ohne Vorspann nur um 10,9 Prozent (vgl. Klein/Schneid 1999: 8).

Der Zusammenhang zwischen incentive und Rücklauf ist allerdings nur in der ersten Welle – und dazu noch sehr schwach – statistisch signifikant ($p = 0,033$, $C = 0,207$), ebenso der Zusammenhang zwischen dem aktuellen Einstieg und dem Rücklauf ($p = 0,078$, $C = 0,231$).

6. Fazit

Bei einem insgesamt schwachen Rücklauf, von dem nicht auszuschließen ist, daß er durch die Befragungszeit über die Weihnachtsfeiertage bewirkt worden ist, zeigt sich, daß sowohl ein incentive in Form von Briefmarken wie auch „interessante“ Einstiegsfragen die Ausschöpfungsquote bei schriftlichen Befragungen steigern können. Übereinstimmend mit vorliegenden Forschungsergebnissen wird deutlich, daß das incentive aber nur beim Erstversand Wirkung zeigt, danach an Bedeutung verliert. Die Bedeutung interessanter Einstiegsfragen zeigt sich beim Erstversand zwar auch, allerdings schwächer als die Bedeutung des incentives; erst beim Rücklauf als Folge der Nachfabaktionen wirken sich die Einstiegsfragen ausschöpfungserhöhend aus.

Während mit der Bedeutung von incentives beim Erstversand ein bereits bekanntes Phänomen der Methodologie von mail surveys bestätigt werden konnte, bleibt die Frage nach der Bedeutung interessanter Einstiegsfragen unbeantwortet – zumindest konnte die mit solchen Fragen verbundene Erwartung an eine deutliche Ausschöpfungserhöhung bei postalischen Befragungen nicht bestätigt werden.

Literatur

- Allen, C.T./Schewe, C.D./Wijk, G., 1980: More on Self-Perception Theory's Foot Technique in the Pre-Call/Mail Survey Setting. *Journal of Marketing Research* 17: 498-502.
- Armstrong, J.S., 1975: Monetary Incentives in Mail Surveys. *Public Opinion Quarterly* 39/1: 111-116.
- Blumberg, H.H./Fuller, C./Hare, A.P., 1974: Response Rates in Postal Surveys. *Public Opinion Quarterly* 8/1: 113-123.
- Brennan, M./Hoek, J./Astridge, C., 1991: The Effects of Monetary Incentives on the Response Rate and Cost-Effectiveness of a Mail Survey. *Journal of the Market Research Society* 33/3: 229-241.
- Church, A. H., 1993: Estimating the Effects of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly* 57: 62-79.
- Conant, J. S./Smart, D.T./Walker, B.J., 1990: Mail Survey Facilitation Techniques: An Assessment and Proposal Regarding Reporting Practices. *Journal of the Market Research Society* 32/4: 569-582.
- Dillman, D.A., 1978: *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- Dillman, D.A./Frey, J.H., 1974: Contribution of Personalization to Mail Questionnaire Response as an Element of a Previously Tested Method. *Journal of Applied Psychology* 59: 297-301.
- Fox, R.J./Crask, M.R./Kim, J., 1988: Mail Survey Response Rate – A Meta-Analysis of Selected Techniques for Inducing Response. *Public Opinion Quarterly* 52: 467-491.
- Friedman, H.M./San Augustine, A.J., 1979: The Effects of a Monetary Incentive and the Ethnicity of the Sponsor's Signature on the Rate and Quality of Response to a Mail Survey. *Journal of the Academy of Marketing Science* 7: 95-101.
- Gajraj, A.M./Faria, A.J./Dickinson, J.R., 1990: A Comparison of the Effect of Promised and Provided Lotteries, Monetary and Gift Incentives on Mail Survey Response Rates, Speed and Costs. *Journal of the Market Research Society* 32/1: 141-162.
- Goodstadt, M.S./Chung, L./Kronitz, R./Cook, G., 1977: Mail Survey Response Rates: Their Manipulation and Impact. *Journal of Marketing Research* 14: 391-395.
- Hackler, J.C./Bourgette, P., 1973: Dollars, Dissonance and Survey Returns. *Public Opinion Quarterly* 37/2: 276-281.
- Hansen, R.A., 1980: A Self-Perception Interpretation of the Effect of Monetary and Non-Monetary Incentives on Mail Survey Respondent Behavior. *Journal of Marketing Research* 17: 77-83.
- Harkness, J./Mohler, P./Schneid, M./Christoph, B., 1998: Incentives in Two German Mail Surveys 1996/97 & 1997. S. 201-218 in: A. Koch/R. Porst (Hrsg.), *Nonresponse in Survey Research*. ZUMA-Nachrichten Spezial 4.

- Heberlein, Th. A./Baumgartner, R., 1978: Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. *American Sociological Review* 43: 447-462.
- Hippler, H.J., 1985: Schriftliche Umfragen bei repräsentativen Bevölkerungsstichproben oder: Wie erreicht man 78%? S. 71-74 in: M. Kaase/M. Küchler (Hrsg.), *Herausforderungen der empirischen Sozialforschung*. Mannheim: ZUMA.
- Hippler, H.J., 1988: Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven. *planung und analyse* 6: 244-248.
- Huck, S.W./Gleason, E.M., 1974: Using Monetary Inducements to Increase Response Rate from Mailed Surveys. *Journal of Applied Psychology* 59: 222-225.
- James, J.M./Bolstein, R., 1990: The effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys. *Public Opinion Quarterly* 54: 346-361.
- Jones, W.H./Lang, J.R., 1980: Sample Composition Bias and Response Bias in a Mail Survey: A Comparison of Inducement Methods. *Journal of Marketing Research* 17: 69-76.
- Kanuk, L./Berenson, C., 1975: Mail Surveys and Response Rates: A Literature Review. *Journal of Marketing Research* 12: 440-453.
- Kerin, R.A./Peterson, R.A., 1977: Personalization, Respondent Anonymity and Response Distortion in Mail Surveys. *Journal of Applied Psychology* 62: 86-89.
- Klein, S./Schneid, M., 1999: Feldbericht. Schriftliche Umfrage: Projekt „Mail-Survey“. ZUMA-Technischer Bericht 99/15. Mannheim: ZUMA.
- Lamnek, S./Treppl, R., 1991: Die Nichtteilnahme an sozialwissenschaftlichen Befragungen. *planung und analyse* 18/6: 205-211.
- Linsky, A.S., 1975: Stimulating Responses to Mailed Questionnaires: A Review. *Public Opinion Quarterly* 39: 82-101.
- Paolillo, J./Lorenzi, P., 1984: Monetary Incentives and Mail Questionnaire Response Rates. *Journal of Advertising* 13: 46-48.
- Porst, R., 1996: Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Die Sicht der Institute. ZUMA-Arbeitsbericht 96/07. Mannheim: ZUMA.
- Rucker, M./Hughes, R./Thompson, R./Harrison, A./Vanderlip, N., 1984: Personalization of Mail Surveys: Too Much of a Good Thing? *Educational and Psychological Measurement* 44/4: 893-905.
- Schewe, C.D./Courmoyer, N.D., 1976: Prepaid vs. Promised Incentives to Questionnaire Response: Further Evidence. *Public Opinion Quarterly* 40: 105-107.
- Schneid, M./Stiegler, A., 1999: Feldbericht Telefonumfrage Projekt „Mailsurvey Non-Response“. ZUMA-Technischer Bericht 99/05. Mannheim: ZUMA.
- Tedin, K.L./Hofstetter, R.C., 1982: The Effect of Cost and Importance Factors on the Return Rate for Single and Multiple Mailings. *Public Opinion Quarterly* 46/1: 122-128.

Thoma, M./Zimmermann, M., 1996: Zum Einfluß der Befragungstechnik auf den Rücklauf bei schriftlichen Umfragen. Experimentelle Befunde zur 'Total-Design-Methode'. ZUMA-Nachrichten 39: 141-157.

Walker, B.J./Burdick, R.K., 1977: Advance Correspondence and Error in Mail Surveys. Journal of Marketing Research 14: 379-382.

Wiseman, F., 1972: Methodological Bias in Public Opinion Surveys. Public Opinion Quarterly 36/1: 105-108.

Wotruba, T.R., 1966: Monetary Inducements and Mail Questionnaire Response. Journal of Marketing Research 3: 398-400.

Yammarino, F.J./Skinner, S.J./Childers, T.L., 1991: Understanding Mail Survey Response Behavior: A Meta-Analysis. Public Opinion Quarterly 55: 613-639.

Yu, J./Cooper, H., 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. Journal of Marketing Research 20: 36-44.

Anhang

Bevor Sie mit den Fragen zum eigentlichen Thema unserer Umfrage „Einstellung zu Staat und Regierung“ beginnen, bitten wir Sie, uns einige Fragen zum Thema „*Sexueller Mißbrauch an Kindern*“ zu beantworten.

1. *In den letzten Wochen und Monaten wurde in den Medien viel über sexuell mißbrauchte Kinder berichtet. Glauben Sie, daß heutzutage mehr Kinder als früher mißbraucht werden, oder glauben Sie, daß durch Presse und Fernsehen nur mehr Fälle bekannt werden? Kindesmißbrauch...*

- kommt heutzutage öfter vor als früher
- kommt nicht öfter vor als früher, wird aber mehr berichtet
- kommt seltener vor als früher, wird von der Presse aufgebauscht
- kann ich nicht beurteilen

2. *Wie sollte der Staat auf das Problem « Kindesmißbrauch » reagieren? Welche der folgenden Aussagen kommt Ihrer Ansicht am nächsten?*

- wir brauchen neue und schärfere Gesetze gegen Kindesmißbrauch
- die bestehenden Gesetze sind ausreichend, sie müßten nur richtig angewandt werden
- die bestehenden Gesetze und ihre Anwendung sind ausreichend
- habe darüber noch nicht nachgedacht

3. *Der schlimmste Fall ist der sexuelle Mißbrauch mit Todesfolge. Welche Strafe sollte Ihrer Meinung nach auf dieses Vergehen stehen ?*
- Gefängnisstrafe unter 15 Jahren
- Gefängnisstrafe von 15 Jahren
- Gefängnisstrafe von 15 Jahren mit anschließender Zwangsverwahrung
- Lebenslängliche Gefängnisstrafe mit Chance auf Begnadigung
- Lebenslängliche Freiheitsstrafe ohne Chance auf Begnadigung
- in diesem Falle sollte man auch die Todesstrafe nicht ausschließen
4. *Welche Strafe wäre Ihrer Meinung nach angemessen, wenn das Kind den Mißbrauch überlebt ?*
- Gefängnisstrafe unter 5 Jahren
- Gefängnisstrafe zwischen 5 und 10 Jahren
- Gefängnisstrafe zwischen 10 und 15 Jahren
- Gefängnisstrafe von 15 Jahren mit anschließender Zwangsverwahrung
- Lebenslängliche Gefängnisstrafe mit Chance auf Begnadigung
- Lebenslängliche Freiheitsstrafe ohne Chance auf Begnadigung
- in diesem Falle sollte man auch die Todesstrafe nicht ausschließen
5. *Wenden wir uns kurz dem Täter zu. Welche der folgenden Aussagen kommt Ihrer Meinung am nächsten ? Die meisten Täter ...*
- sind krank und brauchen ärztliche oder psychiatrische Behandlung
- sind zwar krank, aber dennoch für ihre Tat voll verantwortlich
- sind schlicht kriminell und sollten entsprechend behandelt werden
6. *Eine letzte Frage zu Wiederholungstätern, also zu Personen, die bereits wegen sexuellem Mißbrauch bestraft worden sind und dann wieder rückfällig werden. Welcher der beiden Aussagen stimmen Sie am ehesten zu?*
- Täter, die nicht dauerhaft inhaftiert werden können, sollten zwangsweise sterilisiert werden, dann gibt es keine Wiederholungstäter
- in einem Rechtsstaat ist eine Zwangssterilisation nicht möglich, wir müssen deshalb mit diesem Risiko leben

INTERAKTIONSEFFEKTE IN STRUKTURGLEICHUNGSMODELLEN MIT DER THEORIE DES GEPLANTEN VERHALTENS: MULTIPLE GRUPPENVERGLEICHE UND PRODUKTTERME MIT LATENTEN VARIABLEN¹

JOST REINECKE

Dieser Beitrag thematisiert Modellierungsmöglichkeiten von Interaktionseffekten in Strukturgleichungsmodellen. Wert×Erwartungsprodukte zu den Konstrukten Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle, formuliert in der Theorie des geplanten Verhaltens, werden hierzu herangezogen. Anhand einer repräsentativen Stichprobe von Jugendlichen und jungen Erwachsenen kann mit multiplen Gruppenvergleichen und latenten Produktmodellen gezeigt werden, daß für das Wert×Erwartungsprodukt der wahrgenommenen Verhaltenskontrolle ein signifikanter Interaktionseffekt vorliegt. Der Stellenwert unterschiedlicher Schätzverfahren (ML, GLS und WLS) wird in bezug auf die latenten Produktmodelle diskutiert.

The article discusses strategies of modeling interaction effects in structural equations. Expectancy×value products of the constructs attitude, subjective norm and perceived behavioral control of the Theory of Planned Behavior are considered. Multiple group comparisons as well as latent product models with data from a representative sample of adolescents and young adults show a significant interaction effect of the expectancy×value product of perceived behavioral control. The usefulness of different estimation procedures (ML, GLS and WLS) is discussed in relation to the latent product models.

¹ Die Untersuchungen sind mit finanzieller Unterstützung des Deutschen Akademischen Austauschdienstes (DAAD) und des American Council of Learned Societies (ACLS) entstanden. Ich danke Fan Yang Jonsson (University of Uppsala) und Icek Ajzen (University of Massachusetts) und den Reviewern für hilfreiche Kommentare.

1. Einführung

Lineare Strukturgleichungsmodelle sind in den Sozial- und Wirtschaftswissenschaften populär geworden, um kausale Beziehungen zwischen Variablen zu überprüfen. Die Trennung zwischen einem Strukturmodell mit Beziehungen der latenten Variablen untereinander und einem Meßmodell mit Beziehungen zwischen latenten und manifesten Variablen führt zur Kontrolle unsystematischer Meßfehler und damit zur Korrektur der Koeffizienten, die die kausale Beziehung zwischen den latenten Variablen anzeigen (correction for attenuation). Die Beziehungen in Strukturgleichungsmodellen werden in der Regel als lineare Beziehungen angenommen. Verschiedene EDV-Programme mit einer breiten Palette von Schätzverfahren werden seit Jahren angewendet und laufend weiterentwickelt. Hierzu zählen LISREL (vgl. Jöreskog/Sörbom 1996a), EQS (Bentler 1995) und AMOS (Arbuckle/Wothke 1999). Die Erweiterung dieser Modelle auf nicht-lineare Beziehungen ist in den letzten Jahren in der methodologischen Literatur auf großes Interesse gestoßen. Die Möglichkeit, nicht-lineare Restriktionen in LISREL ab der Programmversion 8 definieren zu können, ist für diese Entwicklung mitverantwortlich.

Die Pionierarbeit zur Modellierung nicht-linearer Terme in Strukturgleichungsmodellen ist von Kenny und Judd (1984) geleistet worden. Sie schlugen vor, Produkte von manifesten Variablen als Indikatoren für latente Produktterme zu verwenden und so ein nicht-lineares Meßmodell zu definieren. Hayduk (1987) diskutiert eine LISREL-Modellierung des Kenny-Judd-Modells mit einer großen Anzahl von Phantomvariablen, um die nicht-linearen Restriktionen spezifizieren zu können. Für die Anwendung erwies sich dieser Aufwand aber als wenig praktikabel. Erst die Implementationsmöglichkeit nicht-linearer Restriktionen in LISREL, Version 8, hat die Handhabung nicht-linearer Modelle bedeutend erleichtert und die Spezifikation der Phantomvariablen erübrigt. Auf der anderen Seite darf nicht übersehen werden, daß die statistische Theorie für nicht-lineare Modellierungen noch nicht voll entwickelt ist. Simulationsstudien zeigen, wie robust bestimmte Schätzverfahren wie Maximum Likelihood (ML), General Least Squares (GLS) und Weighted Least Squares (WLS) bei Modellen mit Interaktionsvariablen sind, wie die Goodness-of-fit Maße zu bewerten und wie Standardfehler der Parameter zu interpretieren sind. Baumgartner und Bagozzis (1995) Simulationsstudie zeigt, daß ML und WLS zuverlässige Parameterschätzungen aufweisen. Die auf der üblichen Multinormalverteilungsannahme basierende χ^2 -Statistik und die Standardfehler sind bei ML-Schätzern allerdings verzerrt. WLS-Schätzungen sind bei ausreichender Stichprobengröße verwendbar. Hier zeigen die Ergebnisse der Autoren, daß die auf der asymptotischen statistischen Theorie basierende χ^2 -Statistik und die Standardfehler korrekte Ergebnisse liefern (vgl. Baumgartner/Bagozzi 1995). Jöreskog/Yang (1996) und Yang Jonsson (1997) vergleichen ebenfalls Simulationsergebnisse für ML- und WLS-Schätzungen. Dabei variieren sie

neben der Stichprobengröße auch die Anzahl der manifesten Produktterme des Kenny-Judd-Modells. Ihre Ergebnisse lassen sich folgendermaßen zusammenfassen: Für ML-Parameter ergeben sich bei einer Stichprobengröße von mehr als 400 Personen keine Schätzprobleme. Die χ^2 -Statistik widerlegt aufgrund der Verletzung der Multinormalverteilungsannahme das Modell zu oft. Die statistischen Voraussetzungen sind beim WLS-Schätzer besser erfüllt und die χ^2 -Statistik zeigt eine bessere Übereinstimmung zwischen Modell und Daten an (Yang Jonsson 1997: 118f.).

Baumgartner/Bagozzi (1995) ergänzen ihre Simulationen mit einem empirischen Beispiel. Sie testen ein Wert×Erwartungsmodell abgeleitet aus der Theorie des überlegten Handelns von Ajzen/Fishbein (1980). Das untersuchte Verhalten bezieht sich auf den Gebrauch von Coupons (Rabattmarken) für Konsumgüter. Konform zu den Hypothesen der Theorie zeigen die Resultate, daß die Einstellungen zum Couponegebrauch sowohl durch die Haupteffekte der Wert- und Erwartungskomponente (beliefs and evaluations) als auch durch den Interaktionseffekt zwischen Werten und Erwartungen erklärt werden können. Für ML und GLS ist der Interaktionseffekt gerade noch signifikant, bei WLS kann die Signifikanz des Interaktionseffektes allerdings deutlich nachgewiesen werden (Baumgartner/Bagozzi 1995: 208).

Yang Jonsson (1997: 131f.) reanalysiert die Daten von Baumgartner und Bagozzi (1995) durch einen multiplen Gruppenvergleich, durch eine Regressionsanalyse mit der Faktorkovarianzmatrix und durch ein latentes Interaktionsmodell. Die verschiedenen Verfahren bestätigen den von Bagozzi und Baumgartner gefundenen Interaktionseffekt (Yang Jonsson 1997: 146).

Die Ergebnisse und Erkenntnisse der Studien von Baumgartner/Bagozzi (1995) und Yang Jonsson (1997) sind für die Analyse von Interaktionseffekten in Strukturgleichungsmodellen aus der Anwenderperspektive wichtig und richtungsweisend, da sie Simulationsdaten und empirische Daten für die Interaktionsmodelle verwenden und die Ergebnisse mit verschiedenen Schätzverfahren vergleichend diskutieren. Allerdings sind eine Reihe von Einschränkungen bezüglich der empirischen Daten und Modelle zu verzeichnen:

1. Das untersuchte Wert×Erwartungsmodell basiert auf einer Fallzahl (N=253), die insbesondere für WLS-Schätzungen und die Verwendung der asymptotischen Varianz/Kovarianzmatrix zu klein ist.
2. Die Autoren analysieren nur einen Teil der Theorie überlegten Handelns, nämlich das Verhältnis zwischen Verhaltensüberzeugungen (behavioral beliefs) und Einstellungen zum Verhalten (attitudes toward the behavior). Die Beziehung zwischen normativen Überzeugungen (normative beliefs) und subjektiven Normen (subjective norm) sind nicht berücksichtigt worden (vgl. auch Ajzen/Fishbein 1980: 49).

Diese Arbeit versucht, die genannten Einschränkungen zu überwinden. Eine entsprechende Stichprobengröße wird die Evidenz von Interaktionseffekten besser untermauern können. Außerdem werden sich die Analysen nicht nur auf Teile der Theorie des überlegten Handelns beziehen sondern explizit auch die Erweiterung dieser Theorie zur Theorie des geplanten Verhaltens (Ajzen 1988, 1991) berücksichtigen. Diese Erweiterung führt in Ergänzung zu Einstellungen und Normen ein weiteres Konzept, die wahrgenommene Verhaltenskontrolle, zur Erklärung von Verhaltensintentionen und Verhalten ein. Dieses dritte Konzept erfaßt den Grad der Kontrollmöglichkeiten, die der Handelnde zur Umsetzung seiner Verhaltensabsichten wahrnimmt.

Der folgende Abschnitt gibt eine kurze Einführung in das Konzept der Theorie des geplanten Verhaltens und ihrer Beziehung zu Wert×Erwartungsmodellen (Abschnitt 2). In Abschnitt 3 werden die Untersuchung, die Daten und die Meßinstrumente vorgestellt. Abschnitt 4 zeigt zwei statistische Verfahren zur Analyse von Interaktionseffekten: der multiple Gruppenvergleich zum Nachweis, ob überhaupt ein bedeutsamer Interaktionseffekt existiert und latente Produktmodelle zur Analyse der Einflußstärke solcher Interaktionseffekte. Die Ergebnisse, ermittelt durch drei Schätzverfahren (ML, GLS und WLS) werden anschließend miteinander verglichen. In Abschnitt 5 werden die Resultate zusammengefaßt, die verbliebenen Defizite erörtert und Hinweise für die Behandlung von Interaktionseffekten in zukünftigen Untersuchungen gegeben.

2 Theorie, Modell und nicht-lineare Strukturgleichungen

Die Theorie des geplanten Verhaltens (Ajzen 1988, 1991) baut auf Banduras Arbeiten zur Selbstwirksamkeitstheorie (Bandura 1977, 1982) auf und überwindet eine wesentliche Annahme der Theorie des überlegten Handelns (Ajzen/Fishbein 1980): Jegliches zu erklärende Verhalten steht immer unter der vollständigen, willentlichen Kontrolle der Person, die entsprechende Verhaltensabsichten äußert. Mit dem Konzept der wahrgenommenen Verhaltenskontrolle braucht diese Annahme nicht weiter vorausgesetzt zu werden, da das Ausmaß der Fähigkeiten, das zu zeigende Verhalten in die Tat umzusetzen, als Prädiktor der Verhaltensintention und des Verhaltens postuliert wird. Die Einstellung und die subjektive Norm sind, wie in der Theorie überlegten Handelns, direkte Prädiktoren der Verhaltensabsicht (vgl. Ajzen 1988: 133).

Eine anwendungsbezogene Diskussion der Theorie innerhalb der Sozialpsychologie (vgl. hierzu die Übersicht Frey et al. 1993) und der verhaltenstheoretisch orientierten Soziologie (vgl. hierzu Reinecke 1997) sowie der Politikwissenschaft (vgl. hierzu Kunz 1997) soll an dieser Stelle nicht vorgenommen werden.

Die Konstrukte Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle werden jeweils durch aufsummierte Produkte aus Wert- und einer Erwartungsvariablen bestimmt. Die Relevanz (saliency) der entsprechenden Wert- und Erwartungsvariablen wird in Vorstudien geprüft (vgl. Ajzen 1991). Für die Einstellung sind dies die Erwartungen, daß ein zu zeigendes Verhalten zu einem Ergebnis führt (b_i) und die Bewertungen des Ergebnisses (e_i):

$$(1) \quad \textit{Attitude (ATT)} = \sum_{i=1}^n b_i e_i$$

Das Produkt für die subjektive Norm besteht aus den Erwartungen über die Ausführung des zu zeigenden Verhaltens durch Mitglieder einer Bezugsgruppe (n_j) und den Motivationen, mit den wahrgenommenen Erwartungen übereinzustimmen (m_j):

$$(2) \quad \textit{Subjective Norm (SN)} = \sum_{j=1}^g n_j m_j$$

Wahrgenommene Verhaltenskontrolle wird bestimmt durch das Produkt aus Kontrollüberzeugungen (c_k) und der Wahrscheinlichkeit, die Fähigkeit aufzubringen, das Verhalten auszuführen (p_k):

$$(3) \quad \textit{Perceived Behavioral Control (PBC)} = \sum_{k=1}^l c_k p_k$$

Wenn die Wert×Erwartungsmodelle in den Gleichungen 1, 2 und 3 gültig sind, dann müssen die Überzeugungsmessungen bedeutsame Interaktionseffekte auf die jeweiligen Konstrukte (Einstellungen, subjektive Norm und Verhaltenskontrolle) haben. Die linearen Beziehungen zwischen Einstellung, subjektiver Norm, wahrgenommener Verhaltenskontrolle, Verhaltensintention und Verhalten werden hier nicht weiter untersucht, sondern ausschließlich das Verhältnis der Produktvariablen zu den Konstrukten Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle.

Gleichungen 1, 2 und 3 machen deutlich, daß die Verwendung von Produktvariablen in Strukturgleichungen die Berücksichtigung nicht-linearer Komponenten in den Modellen zur Folge hat. Wenn die Vorteile des Strukturgleichungsansatzes (Berücksichtigung eines Meßmodells) auch bei nicht-linearen Komponenten ausgenutzt werden soll, dann muß zwischen manifesten und latenten Produktvariablen unterschieden werden. Kenny und Judd (1984) schlagen ein Verfahren zur Handhabung von Interaktionseffekten vor, daß der latenten Interaktionsvariablen explizit manifeste Produktvariablen als Indikatoren

zuordnet. Die Beziehungen zwischen abhängigen und unabhängigen latenten Variablen werden in der LISREL-Terminologie folgendermaßen spezifiziert.²

$$(4) \quad \eta_4 = \alpha + \beta_{41}\eta_1 + \beta_{42}\eta_2 + \beta_{43}\eta_3 + \zeta_4$$

η_1 und η_2 sind die zwei unabhängigen latente Variablen. η_3 ist die latente Interaktionsvariable, die aus dem Produkt von η_1 und η_2 besteht. η_4 ist die abhängige latente Variable. α ist der Vektor für den „intercept term“, β der Vektor der zu schätzenden Strukturgleichungskoeffizienten, der sich aus den Haupteffekten β_{41} und β_{42} sowie dem Interaktionseffekt β_{43} zusammensetzt. ζ repräsentiert die Residualgröße von η_4 (vgl. Kenny/Judd 1984: 201).³

Abbildung 1 zeigt die graphische Umsetzung des Modells von Kenny und Judd auf einen Teil der Theorie des geplanten Verhaltens: die Beziehung zwischen Kontrollüberzeugungen (c_k), Fähigkeiten (p_k) und wahrgenommener Verhaltenskontrolle (PBC). Wie in Gleichung 3 formuliert, werden die Kontrollüberzeugungen mit den Fähigkeiten über k Items multipliziert. Die Kontrollüberzeugungen und die Fähigkeiten repräsentieren in diesem Beispiel mit je drei Messungen (y_1 bis y_6) die latenten Variablen η_1 und η_2 . Dies ermöglicht drei Produktvariablen ($y_1 * y_3$, $y_2 * y_4$ und $y_3 * y_6$) als Messungen der latenten Produktvariable η_3 .⁴ η_4 ist die abhängige Variable PBC, die durch sogenannte „globale“ Indikatoren (in diesem Beispiel die Variablen y_{10} und y_{11}) gemessen wird (vgl. Ajzen 1991: 191). Eine Umsetzung dieses Interaktionsmodells auf die Konstrukte Einstellung (ATT) und subjektive Norm (SN) würde in der gleichen Weise erfolgen.

Die korrespondierende Meßgleichung zur Strukturgleichung 4 wird gegeben durch:

$$(5) \quad y = \tau_y + \lambda_y \eta + \varepsilon$$

wobei $y' = (y_1, y_2, \dots, y_p)$ der Vektor der gemessenen Variablen ist, τ_y ist der Vektor der „intercept terms“, λ_y die Matrix der Faktorenladungen und ε der Vektor der Fehlervarian-

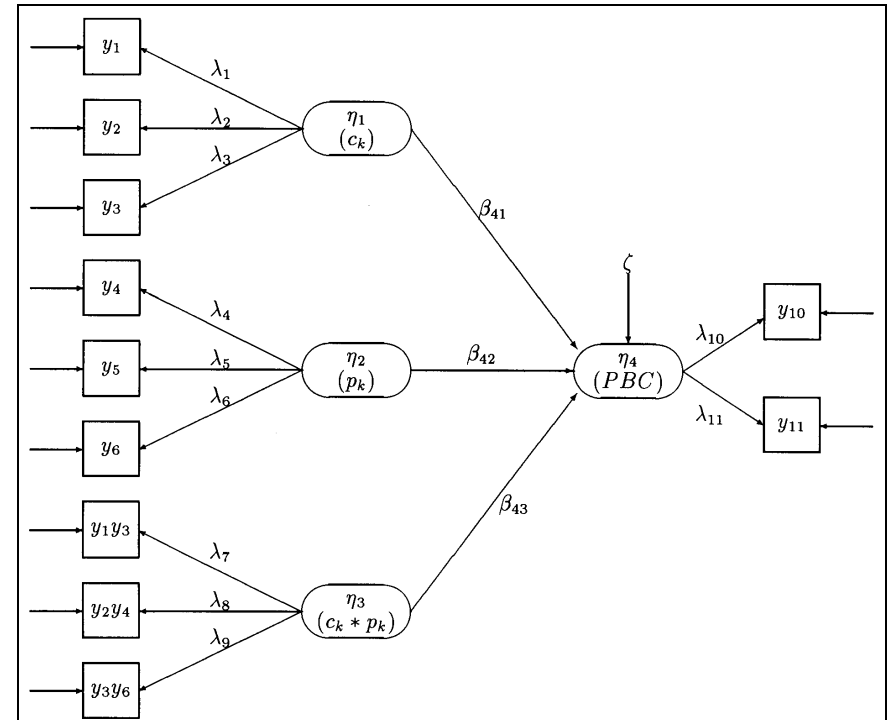
² Der Einfachheit halber ist das generalisierte LISREL-Modell verwendet worden: ξ , x und γ werden durch η , y and β ersetzt.

³ Yang Jonsson (1997: 7f.) weist darauf hin, daß Kenny und Judd keine „intercept terms“ der gemessenen und latenten Variablen berücksichtigen.

⁴ Nach Gleichung 4 ist jede Produktkombination der Indikatoren von η_1 und η_2 möglich. Die Bildung von Wert×Erwartungsprodukten nach der Theorie des geplanten Verhaltens erfordert aber ausschließlich die Produktbildung von inhaltlich korrespondierenden Wert- und Erwartungsmessungen (vgl. Gleichungen 1, 2 und 3). Die Reanalyse des Modells von Bagozzi/Baumgartner (1995) in Jöreskog/Yang (1996: 78) basiert auf einer der Theorie nicht entsprechenden Modellkonzeption, da alle Produktkombinationen als Indikatoren der latenten Produktvariablen verwendet werden. Das Modell ist aber in Yang Jonsson (1997: 135) durch die Beschränkung auf inhaltlich korrespondierende Produkte korrigiert worden.

zen der gemessenen Variablen. Die notwendigen nicht-linearen Restriktionen in den Gleichungen 4 und 5 werden in Abschnitt 4.2 diskutiert.

Abbildung 1: Graphische Umsetzung des Kenny-Judd-Modells auf ein Konstrukt (PBC) der Theorie des geplanten Verhaltens



3. Meßinstrumente und Daten

Das vorliegende Datenmaterial zur Untersuchung von Interaktionseffekten in Strukturgleichungsmodellen ist einer repräsentativen Längsschnittuntersuchung entnommen, die 1996 im Auftrag der Bundeszentrale für gesundheitliche Aufklärung (BZgA) durchgeführt wurde (vgl. Hübner et al. 1998). Die Themen dieser Untersuchung beziehen sich auf den Informationsstand zu Sexualität, Schwangerschaft, Verhütung und Partnerschaft sowie auf die Popularität von Verhütungsmethoden. Die Grundgesamtheit der Erhebung

besteht aus allen Jugendlichen und jungen Erwachsenen deutscher Staatsbürgerschaft im Alter von 16 bis einschließlich 24 Jahren. Insgesamt wurden bei der Erstbefragung 1.223 Interviews (615 Jungen/Männer und 608 Mädchen/Frauen) realisiert. Verhalten im Sinne der Theorie des geplanten Verhaltens bezieht sich in dieser Untersuchung auf die Verwendung der Antibabypille und des Kondoms zur Verhütung einer Schwangerschaft. In der Erstbefragung sind für beide Verhütungsmethoden alle Konstrukte der Theorie einschließlich der zugehörigen Wert- und Erwartungskonzepte operationalisiert worden. Ausschließlich Messungen, die sich auf die Verwendung des Kondoms beziehen, werden für die folgenden Analysen berücksichtigt. Jedes Item wurde auf einer 7-Punkt-Skala gemessen (zu den Itemformulierungen, siehe den Anhang). Die Konstrukte Verhaltensintention und Verhalten wurden mit je einem Item gemessen, die aber in den Interaktionsmodellen nicht berücksichtigt werden.

4. Methoden

4.1 Multiple Gruppenvergleiche zur Feststellung von Interaktionseffekten

Bevor die Effektstärke der Produktterme mit Hilfe latenter Interaktionsmodelle ermittelt wird, sollen zunächst multiple Gruppenvergleiche zeigen, ob die zugrundeliegenden Meßmodelle valide sind und ob bedeutsame und signifikante Interaktionseffekte zu erwarten sind. Wenn die Meßmodelle valide sind, aber keine Evidenz für Interaktionseffekte vorliegen, lassen sich die Beziehungen zwischen den Konstrukten der Theorie des geplanten Verhaltens durch ein lineares Modell beschreiben.

Für den multiplen Gruppenvergleich wird das Interaktionsmodell in Gleichung 4 folgendermaßen umformuliert:

$$(6) \quad \eta_4 = (\alpha + \beta_{42}\eta_2) + (\beta_{41} + \beta_{43}\eta_2)\eta_1 + \zeta_4$$

Die alternative Umformulierung der Gleichung 4 lautet:

$$(7) \quad \eta_4 = (\alpha + \beta_{41}\eta_1) + (\beta_{42} + \beta_{43}\eta_1)\eta_2 + \zeta_4$$

Gleichung 6 zeigt die Regression des unabhängigen Konstrukts η_1 auf η_4 in Abhängigkeit von η_2 . Existiert ein Interaktionseffekt, dann ist β_{43} größer als β_{41} für Personen, die über dem Mittelwert von η_2 liegen und kleiner für Personen, die unter dem Mittelwert von η_2 liegen. Werden die Indikatoren von η_2 zu einem Index aufsummiert, dann lassen sich mit Hilfe dieses Index zwei Gruppen bilden, für die die Variation des Regressionseffektes (η_1 auf η_4) geprüft werden kann.

Gleichung 7 zeigt die Regression des unabhängigen Konstrukts η_2 auf η_4 in Abhängigkeit von η_1 . Existiert ein Interaktionseffekt, dann ist β_{43} größer als β_{42} für Personen, die über dem Mittelwert von η_1 liegen und kleiner für Personen, die unter dem Mittelwert von η_1 liegen. Werden die Indikatoren von η_1 zu einem weiteren Index aufsummiert, dann lassen sich mit Hilfe dieses zweiten Index wiederum zwei Gruppen bilden, für die die Variation des Regressionseffektes (η_2 auf η_4) geprüft werden kann.

Für jedes Konstrukt der Theorie des geplanten Verhaltens (Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle) werden durch Medianisierung der beschriebenen Indizes je zwei Gruppen gebildet. Dies führt gemäß den Gleichungen 6 und 7 zu zwei multiplen Gruppenvergleichen pro Konstrukt: Beispielsweise wird der Zusammenhang zwischen den Kontrollerwartungen (control beliefs) und der wahrgenommenen Verhaltenskontrolle (perceived behavioral control) für die Gruppe mit niedrigen Kontrollbewertungen (low perceived facilitations) und für die Gruppe mit hohen Kontrollbewertungen (high perceived facilitations) simultan geprüft. Auf der anderen Seite kann der Zusammenhang zwischen den Kontrollbewertungen (perceived facilitations) und der wahrgenommenen Verhaltenskontrolle (perceived behavioral control) für die Gruppe mit niedrigen Kontrollerwartungen (low control beliefs) und für die Gruppe mit hohen Kontrollerwartungen (high control beliefs) simultan geprüft werden.

Tabelle 1: Mittelwerte, Standardabweichungen und Gruppengröße für die Gruppierungsvariablen

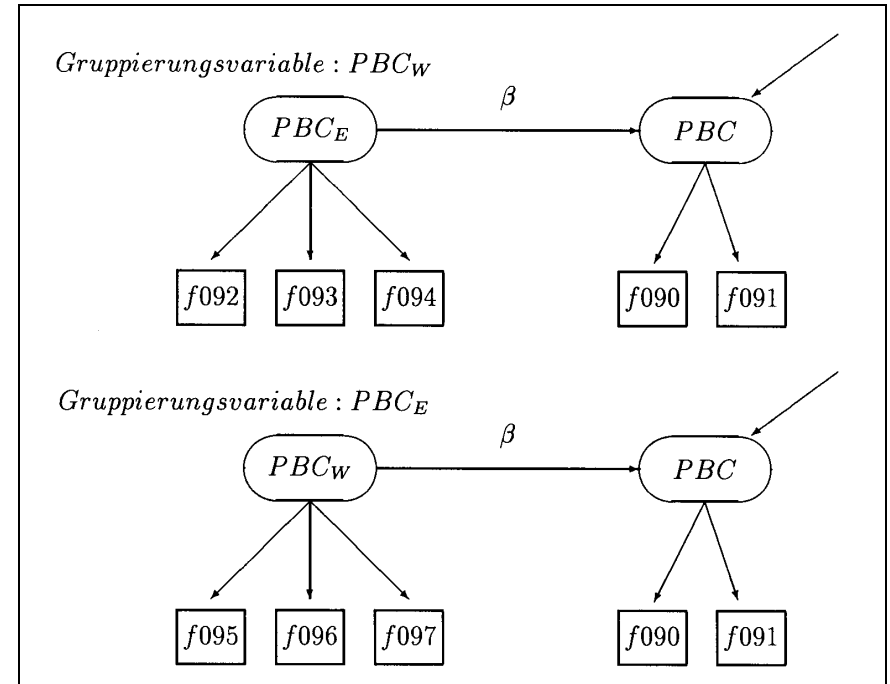
Gruppierungsvariable	\bar{x}	s	Gruppengröße	
			niedrig	hoch
ATT_W	12.14	2.90	536	586
ATT_E	11.78	2.45	601	521
SN_W	13.84	4.64	379	332
SN_E	12.91	5.85	360	351
PBC_W	9.16	4.60	569	397
PBC_E	8.46	4.45	540	426

Die Gruppierungsvariablen (Indizes) werden durch einfache Addition der jeweiligen Indikatoren gebildet, z. B. $PBC_W = f095 + f096 + f097$ und $PBC_E = f092 + f093 + f094$.⁵ Tabelle 1 informiert über Mittelwerte und Standardabweichungen der Gruppierungsvariablen und die einzelnen Gruppengrößen nach der Medianisierung. Die Einstellungs- und Normvariablen haben jeweils höhere Mittelwerte als die beiden Gruppierungsvariablen.

⁵ Die Suffixe **W** und **E** verdeutlichen die Unterscheidung nach **W**ert- und **E**rwartungskonstrukten.

riablen der Verhaltenskontrolle (PBC_W und PBC_E). Auffällig ist auch, daß die Streuungen bei den Gruppierungsvariablen der Einstellung (ATT_W und ATT_E) deutlich niedriger sind.

Abbildung 2: Subgruppenspezifische Regressionsmodelle für wahrgenommene Verhaltenskontrolle



Der obere Teil der Abbildung 2 zeigt den ersten multiplen Gruppenvergleich für die wahrgenommene Verhaltenskontrolle. Die Items $f092$, $f093$ und $f094$ sind die Messungen für die Erwartungsvariable (PBC_E). Die latente Variable PBC wird repräsentiert durch die Items $f090$ und $f091$ (zu den Itemformulierungen, siehe den Anhang). Der untere Teil der Abbildung 2 zeigt den zweiten multiplen Gruppenvergleich für die wahrgenommene Verhaltenskontrolle. Die Items $f095$, $f096$ und $f097$ sind die Messungen für die Bewertungsvariable (PBC_W). Die latente Variable PBC entspricht der Variablen im oberen Teil der Abbildung.

Ein unterschiedlicher Regressionskoeffizient (β) zwischen den Gruppen mit niedrigen und hohen Kontrollbewertungen (niedrige versus hohe PBC_W) bzw. mit niedrigen und

hohen Kontrollerwartungen (niedrige versus hohe PBC_E) würde jeweils eine nichtlineare Beziehung zwischen PBC_E und PBC bzw. PBC_W und PBC indizieren.

Tabelle 2: χ^2 -Werte und Freiheitsgrade der hierarchischen multiplen Gruppenvergleiche

Modelle	Einstellung		Subjektive Norm		Verhaltenskontrolle	
	ATT_E	ATT_W	SN_E	SN_W	PBC_E	PBC_W
Invarianzmodell	282.002 (19)	107.503 (19)	233.334 (19)	62.025 (19)	173.576 (19)	126.667 (19)
Meßfehler freigesetzt	182.307 (16)	48.064 (16)	107.710 (14)	30.709 (14)	61.726 (14)	89.233 (14)
Residuen freigesetzt	110.054 (14)	40.899 (14)	46.134 (12)	14.915 (12)	45.298 (12)	78.195 (12)
Strukturkoeffizient freigesetzt	108.541 (13)	37.945 (13)	45.992 (11)	13.437 (11)	42.463 (11)	72.847 (11)
χ^2 -Differenz (df-Differenz)	1.513 (14-13)	2.954 (14-13)	0.142 (12-11)	1.478 (12-11)	2.835 (12-11)	5.348 (12-11)

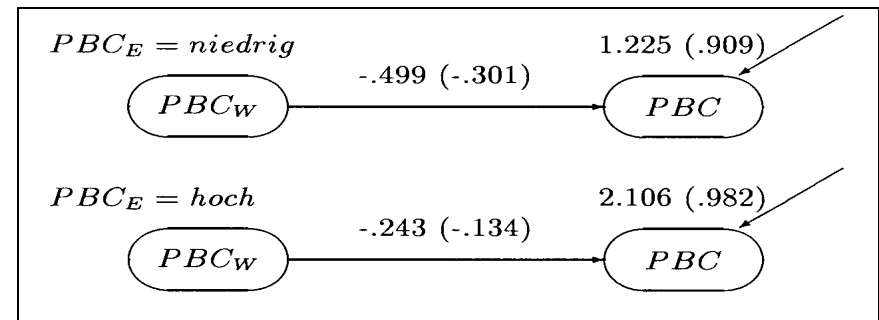
Die Zellen beinhalten χ^2 -Werte des Likelihood-Ratio-Test und in Klammern die Freiheitsgrade. Eine χ^2 -Differenz > 2.0 ist auf dem 5%-Niveau signifikant.

Die entsprechenden Modelle werden für die Konstrukte Einstellung und subjektive Norm berechnet. Wenn jeweils beide multiplen Gruppenvergleiche zu einem signifikant unterschiedlichen Regressionskoeffizient führen, dann ist von einem bedeutsamen Interaktionseffekt zwischen den Erwartungen und den Bewertungen auszugehen. Tabelle 2 zeigt die χ^2 -Tests für die berechneten multiplen Gruppenvergleiche. Für jedes Konstrukt sind, wie erörtert, je zwei Gruppenvergleiche durchgeführt worden. Die hierarchischen Modellanalysen starten mit einem Invarianzmodell, bei dem kein Parameter zwischen den Gruppen variiert. Bei der zweiten Modellvariante werden die Meßfehler der Indikatoren zwischen den Gruppen freigesetzt, bei der dritten die Residuen der abhängigen latenten Variablen. Bis zur dritten Modellvariante sind in allen Analysen signifikante Modellverbesserungen zu verzeichnen. Die vierte und letzte Modellvariante beinhaltet die Variation des Strukturkoeffizienten β , der die Existenz eines Interaktionseffektes nachweist. Drei signifikante Modellverbesserungen zwischen der dritten und vierten Modellvariante werden ermittelt (vgl. die letzte Zeile in Tabelle 2). Eine betrifft das Modell für die Einstellung (χ^2 -Differenz von 2.954) und zwei das Modell für die wahrgenommene Verhaltenskontrolle (χ^2 -Differenzen von 2.835 und 5.348). Diese Resultate zeigen, daß ein bedeutender Interaktionseffekt zwischen den Kontrollerwartungen und den Kontrollbewertungen

gen besteht, während für die Wert×Erwartungsmodelle der Einstellung und der subjektiven Norm ein solcher Effekt nicht vorliegt.

Die Abbildungen 3 und 4 zeigen die Regressionskoeffizienten und Fehlervarianzen der beiden multiplen Gruppenvergleiche für das Konstrukt wahrgenommene Verhaltenskontrolle. Unter der Restriktion, daß die Faktorenladungen zwischen beiden Gruppen invariant sind, werden für den ersten Gruppenvergleich folgende Regressionskoeffizienten ermittelt: $\beta_{niedrig} = -.499$ (t-Value = -4.986) für die Gruppe mit niedrigen Kontrollerwartungen und $\beta_{hoch} = -.243$ (t-Value = -2.317) für die Gruppe mit hohen Kontrollerwartungen (vgl. Abbildung 3). Beide Koeffizienten sind signifikant von Null verschieden. Die Koeffizienten zeigen, daß für Personen mit niedriger Kontrollerwartung eine höhere Beziehung zwischen Kontrollbewertungen (PBC_W) und wahrgenommener Verhaltenskontrolle bezüglich des Gebrauchs von Kondomen (PBC) existiert als für Personen mit hoher Kontrollerwartung. Je mehr die Befragten anzeigen, daß bestimmte Umstände zu Schwierigkeiten beim Gebrauch von Kondomen führen, desto niedriger ist ihre wahrgenommene Verhaltenskontrolle. Diese negative Beziehung variiert mit ihren Kontrollerwartungen (PBC_E).

Abbildung 3: Ergebnis des multiplen Gruppenvergleichs für die Beziehung zwischen PBC_W und PBC



$\chi^2=42.463$ (11); RMSEA=.0769; GFI=.971; RMR=.0604

Standardisierte Koeffizienten in Klammern.

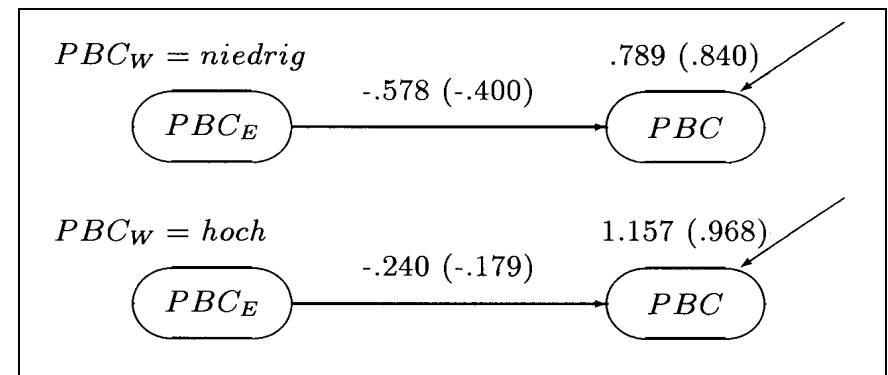
Ergebnisse des Meßmodells werden der Übersicht halber nicht gezeigt.

Für den zweiten Gruppenvergleich (Abbildung 4) werden folgende Regressionskoeffizienten ermittelt: $\beta_{niedrig} = -.578$ (t-Value = -5.162) und $\beta_{hoch} = -.240$ (t-Value = -2.353). Beide Koeffizienten sind auch hier signifikant von Null verschieden. Die Interpretation der Gruppenunterschiede sind ähnlich zum ersten Gruppenvergleich. Für Personen mit niedriger Kontrollbewertung existiert eine höhere Beziehung zwischen Kontrollerwartun-

gen (PBC_E) und wahrgenommener Verhaltenskontrolle bezüglich des Gebrauchs von Kondomen (PBC) als für Personen mit hoher Kontrollbewertung. Je mehr die Befragten bestimmte Schwierigkeiten beim Gebrauch von Kondomen erwarten, desto niedriger ist ihre wahrgenommene Verhaltenskontrolle. Diese negative Beziehung variiert mit ihren Kontrollbewertungen (PBC_W).

Die Goodness-of-Fit Maße der beiden Gruppenvergleiche zeigen eine deutliche Differenz bezüglich der Modellanpassung. Insbesondere das Ergebnis des zweiten Gruppenvergleichs ist wenig zufriedenstellend. Die Freigabe weiterer Parameter würde das Modell zwar verbessern, im Hinblick auf die Bedeutung des Interaktionseffektes würde sich aber an den getroffenen Schlußfolgerungen nichts ändern.

Abbildung 4: Ergebnis des multiplen Gruppenvergleichs für die Beziehung zwischen PBC_E und PBC



$\chi^2=72,847$ (11); RMSEA=.110; GFI=.969; RMR=.0612

Standardisierte Koeffizienten in Klammern.

Ergebnisse des Meßmodells werden der Übersicht halber nicht gezeigt.

4.2 Strukturgleichungsmodelle mit Produkttermen

Die multiplen Gruppenvergleiche haben gezeigt, daß kein Interaktionseffekt zwischen den Wert- und Erwartungsvariablen der Einstellung und der subjektiven Norm vorliegt. Demgegenüber liegt für die Wert- und Erwartungsvariablen der wahrgenommenen Verhaltenskontrolle ein deutlicher Interaktionseffekt vor. Allerdings muß für den multiplen Gruppenvergleich eingeschränkt werden, daß die gewählte Gruppierungsform (ungewichtete Indexbildung und anschließende Medianisierung) eher ad hoc ist und eine andere Vorgehensweise mit den gleichen Variablen zu einem anderen Modellergebnis für den multiplen

Gruppenvergleich führen kann. Dazu kommt noch die Tatsache, daß der unsystematische Meßfehler der Gruppierungsvariable im multiplen Gruppenvergleich nicht berücksichtigt wird. Das wichtigste Argument für die Konstruktion einer latenten Interaktionsvariablen bezieht sich auf die Ermittlung des Strukturkoeffizienten (β_{43} in Gleichung 4), der die Stärke des Einflusses dieser Variablen auf die abhängige latente Variable zeigt. Die Stärke und damit die Bedeutung des Interaktionseffektes kann mit dem multiplen Gruppenvergleich nur indirekt über die Differenz der Parameter zwischen den Gruppen ermittelt werden.

Auch wenn die multiplen Gruppenvergleiche keine Hinweise für Interaktionseffekte zwischen den Wert- und Erwartungsvariablen der Einstellung und der subjektiven Norm geben, sollen diese aus Vergleichsgründen neben den entsprechenden Variablen der wahrgenommenen Verhaltenskontrolle in jeweils getrennten Strukturgleichungsmodellen mit latenten Produkttermen behandelt werden. Zum Vergleich mit den zitierten Simulationsstudien (Baumgartner/Bagozzi 1995; Jöreskog/Yang 1996; Yang Jonsson 1997) werden alle im folgenden diskutierten Strukturgleichungsmodelle mit dem Maximum-Likelihood-Verfahren (ML), dem Generalized-Least-Squares-Verfahren (GLS) und dem Weighted-Least-Squares-Verfahren (WLS) berechnet (Jöreskog/Sörbom 1996a).

Das ML-Schätzverfahren setzt die multivariate Normalverteilung der gemessenen Variablen voraus. Das GLS-Schätzverfahren ist weniger restriktiv, setzt aber voraus, daß die vierten Momente der gemessenen Variablen (multivariate Kurtosis) Null sind. Werden Produkte von gemessenen Variablen gebildet, dann sind die genannten Voraussetzungen nicht erfüllt. Selbst wenn die gemessenen Variablen normalverteilt sind, können deren Produkte nicht normalverteilt sein. Neuere Simulationsstudien (z. B. Hoogland/Boomsma 1998) zeigen, daß die ML-Schätzer relativ robust gegen Abweichungen von der Normalverteilung sind, wenn die Stichprobengröße mehr als 200 Personen umfaßt. Es ist nach wie vor umstritten, ob die Standardfehler bei Verwendung von Produktvariablen für inferenzstatistische Aussagen genutzt werden können (vgl. Yang Jonsson 1997: 13). Als „guidelines“ (Jaccard/Wan 1996: 58) werden die Standardfehler auch bei Verwendung nicht-linearer Komponenten in Strukturgleichungsmodellen allgemein akzeptiert.

Das Problem schiefverteilter Daten wird am besten durch das WLS-Verfahren (vgl. Browne 1984) behandelt. Die Fit-Funktion berücksichtigt eine positiv definierte Gewichtungsmatrix W , die die asymptotischen Varianzen und Kovarianzen der empirischen Varianz/Kovarianzmatrix S beinhaltet. Weniger statistisch gesprochen werden durch die Gewichtungsmatrix die Informationen über die höheren Momente berücksichtigt und damit die Schiefe bei der Schätzung des Strukturgleichungsmodells korrigiert. Allerdings werden zur Schätzung asymptotischer Varianzen und Kovarianzen ausreichende Daten-

mengen benötigt, da ansonsten die Parameterschätzungen negativ beeinflusst werden (vgl. Jöreskog/Sörbom 1996a: 23).⁶

Um das Modell in Abbildung 5 schätzen zu können, wird der Mittelwertvektor \bar{z} , die Varianz/Kovarianzmatrix S und für das WLS-Verfahren die asymptotische Varianz/Kovarianzmatrix W mit PRELIS (vgl. Jöreskog/Sörbom 1996b) berechnet. Die LISREL-Spezifikation für das Strukturgleichungsmodell nach Gleichung (4) lautet folgendermaßen:

$$(8) \quad \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_1\eta_2 \\ \eta_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \psi_{21} \\ \alpha_4 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & 0 \end{pmatrix} * \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_1\eta_2 \\ \eta_4 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{pmatrix}$$

$$\eta = \alpha + B\eta + \zeta$$

Die Varianz/Kovarianzmatrix Ψ wird folgendermaßen spezifiziert:

$$(9) \quad \Psi = \begin{pmatrix} \psi_{11} & & & \\ \psi_{21} & \psi_{22} & & \\ 0 & 0 & \psi_{11}\psi_{22} + \psi_{21}^2 & \\ 0 & 0 & 0 & \psi_{44} \end{pmatrix}$$

⁶ Satorra/Bentler (1994) and Bentler/Dudgeon (1996) widmen sich dem Problem schiefverteilter Daten mit korrigierten χ^2 -Statistiken und korrigierten Standardfehlern. Modellschätzungen können mit der neuen Lisrel-Version 8.30 vorgenommen werden, die aber in einem weiteren Beitrag gesondert diskutiert werden müssen.

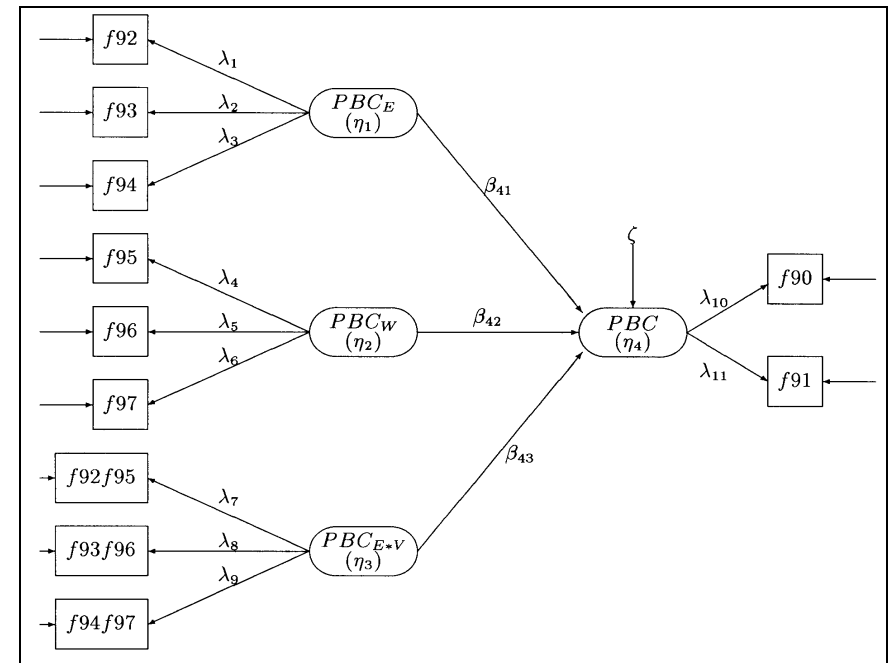
Die Meßfehlervarianzen der drei gemessenen Produktvariablen θ_{77} , θ_{88} und θ_{99} werden jeweils durch folgende Gleichungen restringiert (vgl. Yang Jonsson 1997: 143):

$$(12) \quad \theta_{77} = \tau_4^2 \theta_{11} + \tau_1^2 \theta_{44} + \psi_{11} \theta_{44} + \psi_{22} \theta_{11}^2 \theta_{44}$$

$$(13) \quad \theta_{88} = \tau_5^2 \theta_{22} + \tau_2^2 \theta_{55} + \lambda_{21}^2 \psi_{11} \theta_{55} + \lambda_{52}^2 \psi_{22} \theta_{22}^2 \theta_{55}$$

$$(14) \quad \theta_{99} = \tau_6^2 \theta_{33} + \tau_3^2 \theta_{66} + \lambda_{31}^2 \psi_{11} \theta_{66} + \lambda_{62}^2 \psi_{22} \theta_{33}^2 \theta_{66}$$

Abbildung 5: Empirisches Beispiel für ein latentes Interaktionsmodell angewendet auf ein Konstrukt (PBC) der Theorie des geplanten Verhaltens



Für die Einstellung (ATT) und subjektive Norm (SN) werden nach den Gleichungen 8 und 10 ebenfalls je ein latentes Interaktionsmodell mit den entsprechenden Mittelwertsvektoren und den Varianz/Kovarianzmatrizen geschätzt. Die Modellierung entspricht der Graphik in Abbildung 5. Die Datenbasis beträgt $N = 1120$ für ATT, $N = 711$ für SN und $N = 966$ für PBC. Eine ausreichende Datenbasis ist für alle Schätzverfahren (ML, GLS und WLS) vorhanden. Aus Vergleichsgründen haben alle Modelle die gleiche Anzahl von Parametern und damit die gleiche Anzahl von Freiheitsgraden.⁷

Der Vergleich der Goodness-of-Fit Maße für die drei latenten Interaktionsmodelle (ATT, SN, PBC) berechnet nach den drei erörterten Schätzverfahren (ML, GLS und WLS) führt zu einem relativ eindeutigen Ergebnis (vgl. Tabelle 3).

Tabelle 3: Goodness-of-fit Maße der latenten Interaktionsmodelle

<i>ATT</i>	χ^2	<i>Df</i>	<i>RMSEA</i>	GFI
ML	232.24	27	.0804	.969
GLS	219.60	27	.0902	.968
WLS	63.45	27	.0347	.967
<i>SN</i>	χ^2	<i>Df</i>	<i>RMSEA</i>	GFI
ML	238.76	46	.0760	.951
GLS	227.29	46	.0896	.952
WLS	183.44	46	.0649	.951
<i>PBC</i>	χ^2	<i>Df</i>	<i>RMSEA</i>	GFI
ML	708.62	46	.1220	.897
GLS	523.43	46	.1690	.924
WLS	315.61	46	.0779	.935

Bei jedem Modell wird der schlechteste χ^2 -Wert durch das ML-Verfahren ermittelt, der beste χ^2 -Wert durch das WLS-Verfahren. Der über das GLS-Verfahren berechnete χ^2 -Wert liegt in allen Fällen zwischen den Werten von ML und WLS. Hierbei fällt auf, daß der RMSEA, ermittelt über GLS, bei allen drei Modellen die schlechtesten Werte aufweist. Nach den Empfehlungen von Browne/Cudeck (1993) können die Werte des RMSEA, ermittelt über WLS, für das Einstellungsmodell als gut und für die beiden übrigen Modelle als moderat bezeichnet werden. Weitere Modellmodifikationen würden die Modellanpassungen, insbesondere beim Kontrollmodell (PBC), verbessern. Entscheidende Parameterveränderungen bezüglich der Haupt- und Interaktionseffekte würden daraus aber

⁷ Es ist zu beachten, daß Erwartungs- und Wertmessungen bezogen auf die Einstellung nur auf je zwei gemessenen Variablen basieren. Daher sind hier weniger Parameter zu schätzen.

nicht resultieren. Insgesamt werden mit diesem Ergebnis die Resultate der Simulationsstudien und empirischen Modelle von Baumgartner/Bagozzi (1995) und deren Reanalyse durch Yang Jonsson (1997) bestätigt. Mit einer ausreichenden Stichprobengröße kommt das WLS-Verfahren zu angemessenen Goodness-of-Fit Maßen unter Berücksichtigung der durch die Produktvariablen verletzten Annahme der Normalverteilung. Insbesondere beim ML-Verfahren macht sich die Verletzung dieser Annahme durch die hohen χ^2 -Werte bemerkbar. Tabelle 4 zeigt die unstandardisierten Schätzwerte der Strukturparameter, ihre Standardfehler und t-Werte für alle berechneten Modelle über alle drei Schätzverfahren. Wie aus den Ergebnissen der multiplen Gruppenvergleiche zu erwarten war (vgl. Abschnitt 4.1), ist der Interaktionseffekt (β_{43}) für das Einstellungs- und Normmodell (ATT und SN) nicht signifikant. Der Haupteffekt der Verhaltenserwartungen auf die Einstellung (β_{41} für ATT) ist signifikant für ML, GLS und WLS. Die Ergebnisbewertung erreicht dagegen nur einen signifikanten Effekt bei dem WLS-Verfahren. Ähnlich sind die Ergebnisse für das Normmodell: Der Haupteffekt der normativen Erwartungen auf die subjektive Norm (β_{41} für SN) ist ebenfalls signifikant für ML, GLS und WLS. Der Haupteffekt der normativen Bewertungen auf die subjektive Norm ist weder für ML, noch für GLS und WLS signifikant. Bei Betrachtung der einzelnen Verteilungen der gemessenen Wert- und Erwartungsvariablen des Einstellungs- und Normmodells zeigt sich, daß alle Verteilungen sehr schief sind (vgl. auch die entsprechenden Summenindizes in Tabelle 1). Wenn überwiegend hohe Wert- und Erwartungshaltungen bezüglich der Einstellung zum Kondomgebrauch bzw. der Übereinstimmung mit dem Lebenskontext zu verzeichnen sind, dann werden – statistisch gesehen – kaum signifikante Effekte der jeweiligen Interaktionsvariablen auftreten.

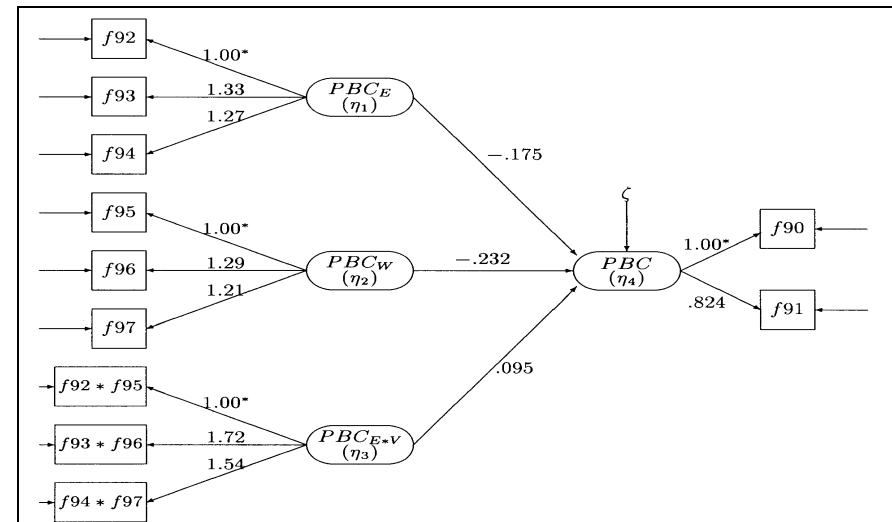
Die Ergebnisse für das Kontrollmodell (PBC) unterscheiden sich deutlich von dem Einstellungs- und Normmodell. Der Effekt der latenten Interaktionsvariable PBC_{E*V} ist signifikant für alle drei Schätzverfahren. Die Schätzwerte variieren zwischen .095 (WLS) und .122 (GLS). Dies bedeutet, daß der Effekt von den Kontrollererwartungen auf die Verhaltenskontrolle (β_{41}) und der Effekt von den Kontrollbewertungen auf die Verhaltenskontrolle (β_{42}) konditional abhängig voneinander sind. Die WLS-Schätzungen für das Kontrollmodell sind zur Verdeutlichung in Abbildung 6 eingetragen.

Der Interaktionseffekt (.095) bedeutet einerseits, daß die Stärke der Beziehung zwischen Kontrollererwartungen und Verhaltenskontrolle (-.175) abhängig von dem Niveau der Kontrollbewertung ist. Andererseits bedeutet der Interaktionseffekt auch, daß die Stärke der Beziehung zwischen Kontrollbewertungen und Verhaltenskontrolle (-.232) abhängig ist von dem Niveau der Kontrollererwartungen. Eine getrennte Interpretation der Haupteffekte (β_{41} und β_{42}) ist demnach nur möglich, wenn jeweils einer dieser Parameter auf Null fixiert wird.

Tabelle 4: Schätzungen, Standardfehler und T-Werte der Strukturparameter für die latenten Interaktionsmodelle

<i>ATT</i>	ML		GLS		WLS	
Parameter	Schätzung	T-Wert	Schätzung	T-Wert	Schätzung	T-Wert
β_{41}	.929	17.16	.937	16.82	.974	14.23
β_{42}	.052	1.22	.078	1.62	.153	2.40
β_{43}	.063	1.56	.053	0.95	-.085	-0.47
<i>SN</i>	ML		GLS		WLS	
Parameter	Schätzung	T-Wert	Schätzung	T-Wert	Schätzung	T-Wert
β_{41}	.485	9.94	.440	9.15	.486	12.14
β_{42}	.032	0.64	.065	1.32	.066	1.59
β_{43}	.013	0.89	.014	1.07	.026	1.38
<i>PBC</i>	ML		GLS		WLS	
Parameter	Schätzung	T-Wert	Schätzung	T-Wert	Schätzung	T-Wert
β_{41}	-.258	-2.33	-.211	-2.73	-.232	-4.15
β_{42}	-.175	-1.29	-.148	-2.08	-1.70	-3.28
β_{43}	.118	4.18	.122	5.27	.095	3.19

Abbildung 6: WLS Schätzungen des latenten Interaktionsmodells für das Konstrukt PBC



Nichtlineare Restriktionen sind aus Gründen der Übersicht in der Abbildung nicht aufgeführt; die mit einem * versehenen Parameter sind zur Identifikation des Modells fixiert.

Inhaltlich kann folgende Interpretation bezüglich des Kontrollmodells getroffen werden: Die Stärke der Beziehung zwischen den Erwartungen, daß bestimmte Umstände (Erhältlichkeit, unkooperativer Partner und Ungeschicklichkeit) Schwierigkeiten mit Kondomen verursachen und der Kontrolle über den Gebrauch von Kondomen zur Kontrazeption, ist abhängig von dem Ausmaß, wie diese Schwierigkeiten bewertet werden. Oder, die Stärke der Beziehungen zwischen den Bewertungen, daß bestimmte Umstände Schwierigkeiten mit Kondomen verursachen und der Kontrolle über den Gebrauch von Kondomen zur Kontrazeption, ist abhängig von dem Ausmaß, wie diese Schwierigkeiten erwartet werden. Bei Betrachtung der einzelnen Verteilungen der gemessenen Wert- und Erwartungsvariablen des Kontrollmodells zeigt sich, daß diese teilweise U-förmig sind (vgl. auch die entsprechenden Summenindizes in Tabelle 1). Die unterschiedlichen Wert- und Erwartungshaltungen bezüglich der wahrgenommenen Verhaltenskontrolle zum Kondomgebrauch lassen – statistisch gesehen – Raum für konditionale Abhängigkeiten zwischen Kontrollerwartungen und Kontrollbewertungen.

5. Zusammenfassung und Schlußfolgerungen

Produkte von Variablen, die auf Wert- und Erwartungsmessungen basieren, sind oft integrative Bestandteile theoretischer Konzepte. Ein Beispiel dafür ist die Theorie des geplanten Verhaltens. Strukturgleichungsmodelle bieten angemessene Modellierungsmöglichkeiten dieser theoretischen Vorgaben. Die Konzeptualisierung einer Meßtheorie für Produktvariablen, die einhergeht mit der Konstruktion latenter Produktvariablen, ermöglicht die Schätzung des Interaktionseffektes unter Berücksichtigung der Konstruktvalidität der manifesten Produktindikatoren.

Es existieren kaum Arbeiten, die Produktvariablen in Strukturgleichungsmodellen bezogen auf die Theorien überlegten Handelns bzw. geplanten Verhaltens behandeln. Van den Putte/Hoogstraaten (1997: 327) resümieren zwar die Problematik der angemessenen Modellierung, verzichten aber auf Produktvariablen und verkürzen das von ihnen untersuchte Wert×Erwartungsmodell der subjektiven Norm auf die Erwartungskomponente. Die von Baumgartner/Bagozzi (1995) und von Yang Jonsson (1997) vorgeschlagenen Meßmodelle zur Berücksichtigung von Produktvariablen in Strukturgleichungsmodellen zeigen demgegenüber eine verbesserte Modellierung der Zusammenhänge zwischen den Konstrukten der Theorie überlegten Handelns bzw. geplanten Verhaltens auf. Allerdings beschränken sich ihre Analysen nur auf ein theoretisches Konstrukt der Theorie überlegten Handelns bzw. geplanten Verhaltens, nämlich auf die Einstellung. Deren Modellprüfung und Reanalysen basieren zudem auch noch auf einer relativ kleinen Stichprobe.

Mit diesem Beitrag sollte eine Formulierung von latenten Interaktionsmodellen erfolgen, die neben der Einstellung auch die Konstrukte subjektive Norm und wahrgenommene Verhaltenskontrolle aus der Theorie des geplanten Verhaltens berücksichtigen. Zudem sollte zur Schätzung eine ausreichende Datenbasis zur Verfügung stehen.

Wert- und Erwartungsmessungen für alle genannten Konstrukte der Theorie des geplanten Verhaltens sind in der repräsentativen Stichprobe dieser Untersuchung enthalten. Für jedes Konstrukt wurde gemäß den Gleichungen 1, 2 und 3 ein nichtlineares Strukturgleichungsmodell formuliert. Diesen Modellen wurden Analysen mit Hilfe von multiplen Gruppenvergleichen vorangestellt. Die Gruppenvergleiche zeigen, daß der Einfluß der Interaktion zwischen Kontrollerwartungen und Kontrollbewertungen (PBC_{E*V}) auf die Verhaltenskontrolle bedeutsam ist und die Signifikanz des Interaktionseffektes durch das nichtlineare Strukturgleichungsmodell bestätigt wird (vgl. Abbildung 6). Die Analysestrategie von Yang Jonsson (1997), multiple Gruppenvergleiche den aufwendigen Modellformulierungen für die nichtlinearen Strukturgleichungsmodelle (vgl. die Restriktionen in den Gleichungen 8 bis 14) voranzustellen, hat sich hier bestätigt: In den beiden Modellen (Einstellung und Norm), wo keine Interaktionseffekte aufgrund des Gruppenvergleichs festzustellen waren, erwies sich die Effektstärke der latenten Interaktionsvariable als nicht signifikant. In dem Kontrollmodell, wo der Gruppenvergleich auf einen bedeutsamen Interaktionseffekt hinwies, erwies sich die Effektstärke der latenten Interaktionsvariable als signifikant (vgl. Tabelle 4).

Der Vergleich zwischen den Schätzungen ermittelt durch drei in LISREL zur Verfügung stehenden Schätzverfahren (ML, GLS und WLS) ergibt keine Differenzen für eine substantielle Interpretation der latenten Interaktionsmodelle. Die Bedeutung der Effekte durch Vergleiche zwischen der Signifikanzstärke (T-Werten) führt bei allen Verfahren zu dem gleichen inhaltlichen Ergebnis. Die Differenzen zwischen den Schätzverfahren sind allerdings auf dem Hintergrund der Modellanpassung (Goodness-of-fit) und damit für statistische Diskussionen bedeutsam: Wo die statistischen Voraussetzungen durch die Verwendung von Produktvariablen am stärksten verletzt sind (ML), ist die Differenz zwischen empirischen und geschätzten Varianzen und Kovarianzen auch am größten. Die Verwendung der asymptotischen statistischen Theorie durch die Verwendung von WLS-Schätzverfahren und asymptotischer Varianz/Kovarianzmatrix führt zur Berechnung von korrekten Signifikanzwerten und Goodness-of-fit Maßen.

Zukünftige Analysen sollten die Integration der hier für jedes theoretische Konzept getrennt berechneten Interaktionsmodelle behandeln und dann diese integrierten Interaktionsmodelle mit dem linearen Teil des theoretischen Modells (d. h. die kausalen Effekte von Einstellung, Norm, und Verhaltenskontrolle auf das Verhalten) verknüpfen. Eine

Erweiterung und Integration der Beziehung zwischen Intention und Verhalten im Längsschnitt, wie sie in Reinecke et al. (1996) diskutiert wird, würde sich hieran anschließen. So könnte schrittweise eine auch statistisch angemessene und nicht nur auf lineare Beziehungen beschränkte Modellierung der Beziehungen zwischen den Konstrukten der Theorie des geplanten Verhaltens erfolgen.

Allerdings muß einschränkend darauf hingewiesen werden, daß die Aufnahme von Interaktionstermen in LISREL-Modellen immer die Berücksichtigung nicht-linearer Restriktionen bei der Modellkonstruktion erfordert, die u. U. zu Schwierigkeiten bei der Modellschätzung führen können. Die angesprochene Integration der einzelnen Interaktionsmodelle würde auch mit einer wachsenden Zahl von nicht-linearen Restriktionen einhergehen. Konvergenzprobleme bei der Modellschätzung sind zu befürchten, die auch im Rahmen der Simulationen von Yang Jonsson (1997: 28) erörtert werden. Auch in dieser Untersuchung sind Konvergenzprobleme aufgetreten, die aber in allen Fällen durch den Wechsel des Schätz-Algorithmus in LISREL (vom Davidon-Fletcher-Powell Algorithmus auf den Fisher Scoring Algorithmus) beseitigt werden konnten.

Abschließend sollen noch kurz Alternativen zu den hier behandelten Methoden erörtert werden. Hier ist einerseits das von Klein et al. (1997) entwickelte Verfahren für latente moderierte Strukturgleichungsmodelle (LMS) zu nennen, daß die für Interaktionsmodelle spezifische Verteilungssituation aus den Rohdaten berücksichtigt und ausnutzt. Der entwickelte Algorithmus liefert über die Berechnung bedingter Erwartungswerte eine ML-Schätzung der Parameter. Auf die in LISREL üblichen nicht-linearen Restriktionen kann verzichtet werden. Verschiedene Simulationsstudien weisen eine höhere Effizienz der durch LMS ermittelten Parameterschätzungen auf, die im Vergleich zu LISREL auch weniger anfällig für die Stichproben- und Verteilungsproblematik von Interaktionsvariablen sind (für weitere Ausführungen, vgl. Mossbrugger et al. 1997; Schermelleh-Engel et al. 1998). Die entwickelte Software ist bis jetzt allerdings nur als Demo-Version verfügbar.

Die andere Alternative zur Schätzung von latenten Interaktionsmodellen bezieht sich auf den zweistufigen Kleinstquadrateschätzer (two-stage least squares, abgekürzt 2SLS), der von Bollen (1995) vorgeschlagen wurde und auch keine aufwendigen nichtlinearen Restriktionen zur Modellierung benötigt. Ähnlich zum WLS-Verfahren wird keine multivariate Normalverteilung der Daten vorausgesetzt. Allerdings zeigen Simulationsstudien, daß die Effizienz der Schätzer erst bei großen Datenmengen erreicht wird. Der zweistufige Kleinstquadrateschätzer ist mittlerweile auch in PRELIS (vgl. Jöreskog/Sörbom 1996b) verfügbar, so daß nicht auf andere Statistikprogrammpakete, wie z. B. SAS ausgewichen werden muß (zu Beispielen und Programminputs durchgeführt mit SAS vgl. Bollen/Paxton

1998). Ein Vergleich mit dem hier vorgestellten Verfahren wäre allerdings über den Umfang dieser Arbeit hinausgegangen.

Korrespondenzadresse

*Jost Reinecke
Institut für Soziologie/Abt. II
Scharnhorststr. 121
48151 Münster
E-mail: reineck@uni-muenster.de*

Literatur

- Ajzen, I., 1988: Attitudes, personality, and behavior. Chicago: The Dorsey Press.
- Ajzen, I., 1991: The theory of planned behavior. *Organizational Behavior & Human Decision Processes*, 50: 179-211.
- Ajzen, I./Fishbein, M., 1980: Understanding attitudes and predicting social behavior. Englewood-Cliffs, NJ: Prentice-Hall.
- Arbuckle, J./Wothke, W., 1999: AMOS 4.0 User' Guide. Chicago: Smallwaters Corporation.
- Bandura, A., 1977: Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84: 191-215.
- Bandura, A., 1982: Self-efficacy mechanism in human agency. *American Psychologist*, 37: 122-147.
- Baumgartner, H./Bagozzi, R. P., 1995: Specification, estimation, and testing of moment structure models based on latent variates involving interactions among the exogenous constructs. *Sociological Methods & Research*, 24: 187-213.
- Bentler, P. M., 1995: EQS 5.0. Structural Equations Program Manual. Encino, CA: Multivariate Software.
- Bentler, P. M./Dudgeon, P., 1996: Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology*, 47: 563-592.
- Bollen, K. A., 1995: Structural equation models that are nonlinear in latent variables: A least-squares estimator. *Sociological Methodology*, 25: 223-251.
- Bollen, K./Paxton, P., 1998: Interactions of latent variables in structural equation models. *Structural Equation Modeling*, 5: 267-293.
- Browne, M. W., 1984: Asymptotically distribution - free methods for the analysis of covariance structures. *British Journal of Mathematical & Statistical Psychology*, 37: 62-83.

- Browne, M. W./Cudeck, R., 1993: Alternative ways of assessing model fit. S. 136-162 in: Bollen, K. A./Long, J. S. (Eds.), *Testing structural equation models*. Newbury Park: Sage.
- Frey, D./Stahlberg, D./Gollwitzer, P. M., 1993: Einstellung und Verhalten: Die Theorie des überlegten Handelns und die Theorie des geplanten Verhaltens. S. 361-398 in: Frey, D./Irlle, M. (Hrsg.), *Theorien der Sozialpsychologie*. Bern: Huber.
- Hayduk, L. A., 1987: *Structural equation modeling with LISREL*. Baltimore: John Hopkins University Press.
- Hoogland, J. J./Boomsma, A., 1998: Robustness studies in covariance structure modeling. *Sociological Methods & Research*, 26: 329-367.
- Hübner, M./Münch, K./Reinecke, J./Schmidt, P., 1998: *Kontrazeption und Sexualität Jugendlicher und junger Erwachsener. Eine repräsentative Wiederholungsbefragung. Reihe: Forschung und Praxis der Sexualaufklärung und Familienplanung. Bundeszentrale für gesundheitliche Aufklärung, Köln.*
- Jaccard, J./Wan, C. K., 1996: *LISREL approaches to interaction effects in multiple regression*. Newbury Park: Sage
- Jöreskog, K. G./Sörbom, D., 1996a: *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Jöreskog, K. G./Sörbom, D., 1996b: *PRELIS 2 User's reference guide*. Chicago: Scientific Software International.
- Jöreskog, K. G./Yang, F., 1996: Non-linear structural equation models: The Kenny - Judd model with interaction effects. S. 57-88 in: Macoulides, G. A./Schumacker, R. E. (Eds.), *Advanced structural equation modeling. Issues and techniques*. Hillsdale: Erlbaum.
- Kenny, D. A./Judd, C. M., 1984: Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96: 201-210.
- Klein, A./Moosbrugger, H./Schermelleh-Engel, K./Frank, D., 1997: A new approach to the estimation of latent interaction effects in structural equation models. S. 479-486 in: Bandilla, W./Faulbaum, F., (Eds.), *SoftStat '97. Advances in Statistical Software 6*. Stuttgart: Lucius & Lucius.
- Kunz, V., 1997: *Theorie rationalen Handelns. Konzepte und Anwendungsprobleme*. Opladen: Leske und Budrich.
- Moosbrugger, H./Schermelleh-Engel, K./Klein, A., 1997: Methodological problems of estimating latent interaction effects. *Methods of Psychological Research Online*, 2: 2.
- Reinecke, J., 1997: *AIDS-Prävention und Sexualverhalten. Die Theorie des geplanten Verhaltens im empirischen Test*. Opladen: Westdeutscher Verlag.
- Reinecke, J./Schmidt, P./Ajzen I., 1996: Application of the theory of planned behavior to adolescents' condom use: A panel study. *Journal of Applied Social Psychology*, 26: 749-772.

Satorra, A./Bentler, P. M., 1994: Corrections to test statistics and standard errors in covariance structure analysis. S. 399-419 in: von Eye, A./Clogg, C. C. (Eds.), *Latent Variable Analysis: Applications for Developmental Research*. Thousand Oaks: Sage.

Schermelleh-Engel, K./Klein, A./Moosbrugger, H., 1998: Estimating nonlinear effects using a latent moderated structural equations approach. S. 203-238 in: Schumacker, R. E./Marcoulides, G. A., (Eds.), *Interaction and nonlinear effects in structural equation modeling*. Mahwah: Erlbaum.

van den Putte, B./Hoogstraten, J., 1997: Applying structural equation modeling in the context of the theory of reasoned action. *Structural Equation Modeling*, 4: 320-337.

Yang Jonsson, F., 1997: Non-linear structural equation models. Simulation studies of the Kenny-Judd model. Uppsala: Studia Statistica Upsaliensia.

Anhang

Die Itembezeichnungen werden in Klammern angegeben und entsprechen den Bezeichnungen aus dem Originaldatensatz. Die Itemformulierungen sind dem für die weiblichen Befragten konstruierten Fragebogen entnommen.

Zwei Items messen Verhaltenserwartungen (b_i) für die Anwendung des Kondoms zur Schwangerschaftsverhütung: *Der Gebrauch des Kondoms würde mich vor einer Schwangerschaft schützen (f081)*, und *Der Gebrauch des Kondoms würde für mich gemeinsame Verantwortung für die Schwangerschaftsverhütung bedeuten (f083)*. Die Endpunkte der Skalen werden mit *sehr unwahrscheinlich* und *sehr wahrscheinlich* bezeichnet.

Zwei Items messen die entsprechenden Ergebnisbewertungen (e_i) für die Anwendung des Kondoms zur Schwangerschaftsverhütung: *Die Vermeidung einer Schwangerschaft ist zur Zeit für mich überhaupt nicht wichtig . . . sehr wichtig (f078)*, und *Die gemeinsame Verantwortung für die Schwangerschaftsverhütung ist für mich überhaupt nicht wichtig . . . sehr wichtig (f080)*.

Drei Items messen die normativen Erwartungen (n_j) bezüglich der Anwendung des Kondoms zur Schwangerschaftsverhütung: *Meine besten Freunde sind der Ansicht, daß mein Partner Kondome zur Verhütung nehmen sollte (f101)*, *Mein Partner ist der Ansicht, daß er Kondome zur Verhütung nehmen sollte (f102)*, und *Meine Eltern sind der Ansicht, daß mein Partner Kondome zur Verhütung nehmen sollte (f103)*. Die Endpunkte der Skalen werden mit *sehr unwahrscheinlich* und *sehr wahrscheinlich* bezeichnet.

Drei Items messen die entsprechenden Motivationen (m_i) zur Konformität mit den Erwartungen der Bezugspersonen: *Wie groß ist im allgemeinen Ihre Bereitschaft, das zu tun, was . . . Ihre besten Freunde bezüglich der Verwendung von Kondomen für richtig halten*

(f107); . . . Ihr Partner bezüglich der Verwendung von Kondomen für richtig hält (f108); und . . . Ihre Eltern bezüglich der Verwendung von Kondomen für richtig halten (f109). Die Endpunkte der Skalen tragen die Bezeichnung *überhaupt keine Bereitschaft* und *sehr große Bereitschaft*.

Drei Items messen die Kontrollerwartungen (c_k) bezüglich der Anwendung des Kondoms zur Schwangerschaftsverhütung: *Wie wahrscheinlich erschweren Ihnen bestimmte Umstände den Gebrauch von Kondomen zur Schwangerschaftsverhütung, wie z.B. die Erhältlichkeit (f092), wie z.B. ein unkooperativer Partner (f093), und wie z.B. Ungeschicklichkeit (f094)*. Die Endpunkte der Skalen werden mit *sehr unwahrscheinlich* und *sehr wahrscheinlich* bezeichnet.

Drei Items messen die entsprechenden wahrgenommenen Fähigkeiten (p_k) bezüglich der Anwendung des Kondoms zur Schwangerschaftsverhütung: *Wie oft können bestimmte Umstände dafür verantwortlich sein, keine Kondome zur Schwangerschaftsverhütung zu gebrauchen, wie z.B. die Erhältlichkeit (f095), wie z.B. ein unkooperativer Partner (f096), und wie z.B. Ungeschicklichkeit (f097)*. Die Befragten konnten auf einer Skala von *nie-mals* bis *oft* antworten.

Drei Items werden zur globalen Messung der Einstellung gegenüber bestimmten Verhalten (ATT) verwendet: *Halten Sie die Anwendung von Kondomen zur Schwangerschaftsverhütung für sehr schlecht . . . sehr gut (f085), sehr nutzlos . . . sehr nützlich (f086), und sehr unwichtig . . . sehr wichtig (f087)*.

Zwei Items werden zur globalen Messung der subjektiven Norm (SN) verwendet: *Wie gut oder wie schlecht finden es die Personen, die Ihnen wichtig sind, wenn Ihr Partner Kondome zur Schwangerschaftsverhütung anwendet (f088), und Die Personen, die mir wichtig sind, denken, daß mein Partner Kondome zur Schwangerschaftsverhütung benutzen sollte (f089)*. Die Endpunkte der Skala für das erste Item werden mit *sehr schlecht* und *sehr gut* bezeichnet, die Endpunkte der Skala für das zweite Item werden mit *sehr unwahrscheinlich* und *sehr wahrscheinlich* bezeichnet.

Zwei Items stehen zur globalen Messung der wahrgenommenen Verhaltenskontrolle (PBC) zur Verfügung: *Für wie schwierig halten Sie die Anwendung von Kondomen zur Schwangerschaftsverhütung (f090), und Für wie wahrscheinlich halten Sie es, daß Ihr Partner gegenwärtig in der Lage ist, Kondome zur Schwangerschaftsverhütung richtig anzuwenden (f091)*. Die Endpunkte der Skala für das erste Item werden mit *sehr schwierig* und *überhaupt nicht schwierig* bezeichnet, die Endpunkte der zweiten Skala werden mit *sehr unwahrscheinlich* und *sehr wahrscheinlich* bezeichnet.

DIE STICHPROBENZIEHUNG BEI ZUWANDERERBEFRAGUNGEN*

KURT SALENTIN

Die Frage wird behandelt, welchen Einfluß vier unterschiedliche Ziehungsverfahren auf die Repräsentativität von Zuwandererstichproben haben und welchen Aufwand sie verursachen. Melderegisterauszüge nach Staatsangehörigkeit schließen eingebürgerte Migranten aus und bedingen einen mit der regionalen Streuung und der Zahl einbezogener Kommunen steigenden finanziellen und Verwaltungsaufwand. Auswahlen auf der Grundlage von Gebietseinheiten wie das Random-Route-Verfahren kommen nur in Frage, wenn ein Screening vorgeschaltet wird. Mit sinkendem Anteil der Zuwandererpopulation an der Wohnbevölkerung werden sie ineffizienter und teurer. Das Schneeballverfahren begünstigt überproportional die Ziehung von Individuen mit großen Kontaktnetzwerken und ergibt keine Wahrscheinlichkeitsstichproben. Namensorientierte Verfahren sind kostengünstig bei geringerer Aktualität des Auswahlbestands und Abstrichen hinsichtlich seiner Vollständigkeit. Für die einzelnen Verfahren werden praktische Erfahrungen aus der Umfrageforschung geschildert.

Four sampling methods are discussed with regard to their ability to produce representative selections of immigrant populations and to the costs and effort involved. Excerpts from German local authority registers of residents based on nationality exclude naturalized persons, i.e. foreigners, who have taken German citizenship. Moreover, the greater the regional spread and the number of target municipalities involved, the higher costs become. Area sampling procedures such as random route procedures are only feasible in conjunction with extensive screening. They are inefficient and expensive for highly dispersed target populations. Snowball sampling tends to favour the inclusion of persons with large social networks and does not produce probability samples. Procedures based on utilising non-German names from directories are cheap but have other drawbacks – some small sections of the target population may be omitted and directories may be out of date. Practical experiences using each of the procedures are reported.

* Der Autor dankt zwei anonymen Gutachtern dieser Zeitschrift für wichtige Hinweise.

1. Einleitung

Das zentrale Gütekriterium einer Stichprobe ist erfüllt, wenn sie „ein verkleinertes Abbild der Grundgesamtheit hinsichtlich der Heterogenität der Elemente und hinsichtlich der Repräsentativität der für die Hypothesenprüfung relevanten Variablen“ darstellt (Friedrichs 1979: 125). In diesem Beitrag wird die Frage behandelt, inwiefern sich mit unterschiedlichen Auswahl- und Stichprobenverfahren für Zuwanderer eine Annäherung an dieses Ideal erreichen läßt. An konkreten Beispielen werden dabei auch allgemeine Probleme der Stichprobenziehung in besonderen Populationen illustriert.

Solange sie von Untersuchungen der Wohnbevölkerung ausgehen kann, muß sich die Literatur zu Auswahl- und Stichprobenverfahren am Problem der Grundgesamtheit nicht lange aufhalten. Die Definition einer Grundgesamtheit der Zuwanderer ist jedoch in hohem Maß von Setzungen abhängig, die unmittelbare Konsequenzen für das Spektrum und die Ausprägung der *relevanten Variablen* besitzen. Schon in terminologischen Ungeheimheiten drückt sich dieses Problem aus: Die Bezeichnung „Ausländer“ greift, wie Boos-Nünning (1990: 16) anmerkt, nur noch, wenn man rechtliche Aspekte in den Vordergrund stellt. Eingebürgerte Zuwanderer, etwa Deutschstämmige aus Mittel- und Osteuropa, sind in manchen Lebensbereichen durchaus mit Asylbewerbern vergleichbar, obwohl sie keine Ausländer sind. Der Begriff des „Einwanderers“ ist politisch hochgradig besetzt und kann nur gegen erhebliche Widerstände verwendet werden. „Zuwanderer“ im engeren Sinn sind lediglich Migranten der ersten Generation, das heißt Personen ausländischer Herkunft mit ständigem Aufenthalt im Inland. Die rechtliche, soziale und ökonomische Situation ihrer Nachkommen, der zweiten und folgenden Generationen, die im Aufnahmeland der Eltern ihre Heimat haben, gleicht sich aber nur langsam den Alteingesessenen an, wie diachrone Untersuchungen über Türken, Italiener und andere Gruppen in der Bundesrepublik zeigen (zum Beispiel Fendrich 1986; Mehrländer et al. 1996; Seifert 1997). Für viele sozialpolitische Fragestellungen ist es mithin geboten, den Aspekt der Abstammung zu beachten - ohne daß damit normativen Abstammungsprinzipien das Wort geredet wird. Im Konzept der „ethnischen Gruppe“ schließlich ist Max Webers (1964: 307) Definition zufolge der „subjektive Glaube an eine Abstammungsgemeinsamkeit“ zentral. Bereits Weber weist indessen darauf hin, daß ethnisches Bewußtsein auch ohne reale Grundlage durch politische Strukturen künstlich geschaffen werden kann und andererseits auch nicht in jeder immigrierten Gruppe vorhanden sein muß. Ob und in welcher Intensität also der Gemeinsamkeitsglaube gegeben ist, läßt sich nur empirisch ermitteln. Für die Definition der Grundgesamtheit ist dieser Aspekt folglich ungeeignet, obwohl er für viele Zielkonstrukte von Bedeutung sein dürfte. Die Bezeichnung „Herkunftsgruppe“ bzw. „-minderheit“ ist am ehesten geeignet, den gemeinten Personenkreis zu umreißen. Für praktische Zwecke wird häufig eine technische Definition ausreichen: Zuwanderer

sollen alle aus einem bestimmten Gebiet (oder generell aus dem Ausland) stammenden und in einem angebbaren Zeitraum (zum Beispiel nach dem Zweiten Weltkrieg) im Inland ansässig gewordenen Personen einschließlich ihrer Nachkommen sein. Die Differenzierung nach Herkunftsgebieten ist zumeist sinnvoll, weil innerhalb der Gesamtheit aller Zuwanderer mit der Herkunft Unterschiede in sozioökonomischen und anderen Variablen einhergehen.

Auf welche Weise diese Personenkreise für die empirische Forschung erreichbar sind, soll im folgenden erörtert werden. Auf Verfahren der bewußten Auswahl, etwa der Mitglieder von Migrantenvereinigungen oder der Klienten von Beratungsstellen, soll nicht eingegangen werden, weil sie wegen der in Vereinen organisierten Interessen etc. allzu leicht inhaltliche Verzerrungen bewirken.

2. Melderegisterstichproben

In den Melderegistern der Kommunen wird die Staatsangehörigkeit aller Einwohner gespeichert. Melderegisterstichproben, auch unter der behördlichen Bezeichnung *Gruppenauskunft* bekannt, können nach Staatsangehörigkeiten gezogen werden, wenn der Nutzer der Stichprobe ein öffentliches Interesse nachweist. Die rechtliche Grundlage hierfür bilden das Melderecht des Bundes (§21 (3) des Melderechtsrahmengesetzes in der Fassung vom 24. Juni 1994, siehe auch Ordemann 1988) und die entsprechenden Gesetze der Länder. Ein öffentliches Interesse der nichtkommerziellen Sozialforschung wird von den Kommunen in der Regel anerkannt, wenn belegt werden kann, daß die Zielgruppe auf andere Weise nicht erreichbar ist, und wenn der Verwendungszweck dargelegt wird. Dieses Verfahren hat den Vorteil, daß jeweils die Gesamtheit der in einer Kommune gemeldeten Personen einer Staatsangehörigkeit (oder eine Teilmenge) erreicht werden kann; es lassen sich Altersvorgaben machen, um beispielsweise keine Kinder in die Stichprobe aufzunehmen oder bestimmte Altersgruppen kontrolliert überzupräsentieren. Melderegisterauszüge sind daher eigentlich das Verfahren der Wahl. Leider sind drei Nachteile zu beachten.

a) Trotz gesetzlicher Meldepflicht ist die faktische Übereinstimmung zwischen Aufenthalt und behördlicher Meldung nicht gewährleistet. Hormuth/Brückner (1985: 53) weisen darauf hin, daß zum Beispiel Um- und Anmeldungen oft unterbleiben, weil Wohnortwechsler noch am alten Wohnort wählen oder günstigere Versicherungstarife in Anspruch nehmen wollen. Es ist auch bekannt, daß Ausländer es wegen erwarteter Schwierigkeiten mit den Behörden in manchen Zielkommunen vorziehen, Umzüge nicht zu melden. Dies betrifft vor allem Personen mit unsicherem ausländerrechtlichen Status. Die größte Diskrepanz dürfte dadurch zustande kommen, daß sich Ausländer nicht abmelden, wenn sie

das Bundesgebiet verlassen. Einen groben Eindruck von der Größenordnung des Fehlers vermittelt ein Vergleich der Ausländerzahlen im Bundesgebiet zwischen dem Jahr 1986, für das die Angaben auf Zensusfortschreibungen beruhen, und dem Folgejahr 1987, für das Anpassungen an die Volkszählungsergebnisse vorliegen (Statistisches Bundesamt 1995: 16ff.). Die ausgewiesenen Zahlen für alle Ausländer *sanken* um sechs Prozent, obwohl in den Jahren davor und danach jeweils ein *Anstieg* um drei bis sechs Prozent verzeichnet worden war. Die Übererfassung schwankte zwischen den Nationalitäten und reichte für Länder wie Ghana an 30 Prozent heran.

b) Herkunftsgruppen wie Flüchtlinge, die im Asylverfahren nach festgelegten Quoten über Bundesländer verteilt und innerhalb dieser dann Kommunen zugewiesen werden, aber zunehmend auch die Gruppen aus dem Mittelmeerraum (aus der Türkei, Griechenland, den Nachfolgestaaten Jugoslawiens, Italien und Spanien), leben nicht konzentriert in großen Städten, sondern in beträchtlichem Umfang auch in Kleinstädten und auf dem Land. Der ländliche Raum muß einbezogen werden, wenn echte Zufallsstichproben benötigt werden. (Nur Zufallsstichproben gewährleisten in der Praxis die Erfüllung der Repräsentativitätsanforderung.) In einzelnen Landgemeinden hält sich aber jeweils nur eine kleine Zahl von ihnen auf. Wenn Zuwanderer in ländlichen Gegenden befragt werden sollen, sind Kontakte zu zahlreichen Kommunen (in manchen Bundesländern zu Kreisen) nötig, um zu hinreichenden Stichprobenumfängen zu gelangen, da die Stichprobe alleine vom Einwohnermeldeamt einer jeden selbständigen Kommune (oder den Ausländerämtern der Kreise) gezogen werden kann. Der zeitliche Vorlauf muß in jedem Fall großzügig bemessen sein. Der organisatorische und finanzielle Aufwand steigt drastisch, wenn dispergierte (in geringer räumlicher Konzentration lebende) Gruppen mit Meldeamtstichproben erreicht werden sollen. Zum einen werden von jedem einzelnen Amt pauschale oder von der EDV-Laufzeit abhängige Gebühren in Rechnung gestellt.¹ Die für die regionale Streuung maßgebliche Zahl der einbezogenen Gemeinden wird daher zur zentralen Variablen für die Höhe der Kosten. Zum anderen müssen stets die gleichen Verhandlungen aufgenommen werden, um zziehungstechnische und datenschutzrechtliche Fragen mit verschiedenen Beteiligten zu klären, was erfahrungsgemäß besonders dann erheblichen Aufwand verursacht, wenn die Kommunen nie zuvor mit der empirischen Forschung kooperiert haben. Auch in größeren Städten sind viele Herkunftsgruppen nur in so kleiner Zahl vertreten, daß Stichproben mit genügender Fallzahl aus mehreren lokalen Melderegisterauszügen zusammengestückelt werden müssen, wobei die gleichen Probleme entstehen. In der Praxis ist auch die deutsche Wohnbevölkerung mit großflächig

¹ Sie belaufen sich beispielsweise je Kommune auf 2.000 DM bis 30.000 DM für eine Stichprobe dreier Nationalitäten mit einem N von je 2.000 Personen (eigene Recherchen).

gestreuten Meldeamtsstichproben nur unter beträchtlichem Aufwand zu erreichen. Kirschner (1984: 119) führt aus, daß bei der Planung der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS), die bekanntlich ursprünglich noch Haushalte mit ausländischem Vorstand ausschloß, dieses Verfahren aus Kosten- und Kapazitätsgründen verworfen wurde. Seit 1994 wird für den ALLBUS jedoch der erhöhte Aufwand wegen einiger Nachteile des bis 1992 genutzten alternativen Verfahrens (ADM-Design, siehe Abschnitt 3) in Kauf genommen (Koch 1997).

c) Auswahlmerkmal in behördlichen Datenbeständen ist die Staatsangehörigkeit, die unter sozialwissenschaftlichen Gesichtspunkten ein nur eingeschränkt taugliches Erkennungskriterium für Zuwanderer darstellt. So ist die in manchen Kontexten wichtige Differenzierung nach Volksgruppen (etwa Kurden versus Türken) nicht im vorhinein realisierbar. Zum anderen sind eingebürgerte Migranten und ihre Nachkommen im Melderegister nicht mehr erkennbar.² Problematisch an diesem Umstand ist für die meisten Gruppen gegenwärtig noch nicht primär die quantitative Unterausschöpfung, die sich bei Kenntnis der Einbürgerungszahlen recht gut beziffern läßt.³ (Sie wird allerdings schon in naher Zukunft stärker ins Gewicht fallen, da die Zahl der Eingebürgerten aus den meisten Gruppen mit Ausnahme der mittel- und osteuropäischen Herkunftsländer sich in den neunziger Jahren alle zwei bis fünf Jahre verdoppelt.) Bedenklich ist vielmehr, daß es bei Ausschluß der Eingebürgerten auf dem Wege der Stichprobenziehung unmöglich wird, Veränderun-

2 In manchen Kommunen registrieren die Einwohnermeldeämter neben der aktuellen auch die frühere Staatsangehörigkeit, allerdings nur lückenhaft, da es sich nicht um eine für die behördliche Statistik notwendige Eintragung handelt. Anderenorts werden Änderungen der Staatsangehörigkeit, wie alle Veränderungen individueller Daten, in jährlichen Historik-Dateien gespeichert. Diese Informationen sind indessen nicht für alle Jahre vorhanden und werden überdies in unterschiedlichen Behörden gehalten. Für einen Zugriff auf diese Informationen fehlt schließlich die Rechtsgrundlage. Nach §21 (2), (3) des Melderechtsrahmengesetzes des Bundes dürfen sich sog. erweiterte Melderegisterauskünfte und Gruppenauskünfte zwar auf die (aktuelle) Staatsangehörigkeit erstrecken. Da Auskünfte über frühere Staatsangehörigkeiten durch das Gesetz jedoch nicht abgedeckt sind, ist eine Weitergabe nach Auskunft beteiligter Ämter ausgeschlossen.

3 Wenn E die Zahl Eingebürgerter einer bestimmten früheren Staatsangehörigkeit von 1974 bis 1998 und S die Anzahl der Personen mit dieser Staatsangehörigkeit im Bundesgebiet Ende 1998 angibt, ergeben sich für die folgenden beispielhaft ausgewählten Gruppen folgende Anteile $R = E/(S+E)$: Türken 8 Prozent, Ex-Jugoslawen 5 Prozent (1981-1995), Sri Lanker (Tamilen) 11 Prozent, Vietnamesen 22 Prozent. Deutschstämmige (Spät-)Aussiedler aus Mittel- und Osteuropa lassen sich zu circa 60 Prozent nach der Übersiedlung umgehend naturalisieren, während der restliche Anteil den bei der Einreise ausgestellten Personalausweis in Verbindung mit dem sicheren Aufenthaltsrecht als ausreichend ansieht oder das Verfahren später einleitet (Lederer 1997: 58, Fn. 4). Als EU-Angehörige mit sicherem Aufenthaltsrecht lassen sich nur wenige Italiener und Griechen einbürgern. Quellen: Beauftragte der Bundesregierung für Ausländerfragen 1998, Lederer 1997: 65 und diverse Mitteilungen des Ausländerzentralregisters.

gen in Einstellungen und Sozialbeziehungen zu verfolgen, die sich im Anschluß an die Einbürgerung vollziehen. „Seit ich die deutsche Staatsbürgerschaft besitze, fühle ich mich nicht mehr als Fremder“, berichtet ein frisch Eingebürgerter anlässlich eines Empfangs durch den Bundespräsidenten.⁴ Da dem Wechsel der Staatsbürgerschaft wahrscheinlich eine beträchtliche symbolische Bedeutung für den Übergang in eine andere Gesellschaft zukommt, ist die Befragung Eingebürgerter für viele Erkenntnisinteressen unverzichtbar. Überdies handelt es sich bei den Eingebürgerten um eine arrivierte Schicht der Zuwanderer, da die Einbürgerung den gesetzlichen Regelungen⁵ entsprechend ein Mindestmaß deutscher Sprachkenntnisse, gesicherten Lebensunterhalt ohne Bezug bestimmter Sozialleistungen und langjährigen Aufenthalt mit legalem Rechtstitel voraussetzt.

Die Ausländerstichprobe des Sozio-oekonomischen Panels (SOEP, siehe Hanefeld 1987), die für Haushalte mit einem Vorstand türkischer, jugoslawischer (inzwischen nach Nachfolgestaaten differenziert), italienischer, griechischer und spanischer Nationalität steht, beruht auf Auszügen der Ausländerkarteien der Kreise (in einigen Bundesländern der Gemeinden). Der Personenauswahl vorgeschaltet ist eine Auswahl dieser Gebietseinheiten auf Grundlage der im Ausländerzentralregister gespeicherten Verteilungsdaten für die Nationalitäten. Hanefelds (1987: 178ff.) Ausführungen lassen den enormen organisatorischen Aufwand bei dieser Methode erahnen. Sie erwähnt auch, daß eine einzelne Behörde für 40 Adressen 2.333,75 DM in Rechnung stellte. Im übrigen kommt dem SOEP aufgrund seines Paneldesigns zugute, daß etwaige Veränderungen im Umfeld von Einbürgerungen erfaßt werden. Da Einbürgerungen zur Zeit seiner Planung in den fünf Herkunftsgruppen so gut wie keine Rolle spielten, kam es in den ersten Wellen nicht zu nennenswerten Lücken; für eine heute neu anzulegende Studie käme indessen eine Stichprobe alleine aus den Ausländerkarteien nicht mehr in Frage.

3. Auswahl auf der Grundlage von Gebietseinheiten

Wenn listenbasierte Verfahren, für die Melderegisterstichproben ein Beispiel sind, nicht zum Einsatz kommen können, weil Listen nicht existieren oder nicht zugänglich sind, werden in der empirischen Forschung häufig wohnstandortbasierte Verfahren eingesetzt. Ihnen liegt die Überlegung zugrunde, daß man für Stichproben der Wohnbevölkerung keine Personenlisten braucht, wenn man einen zufälligen Zugriff auf Wohnungen realisieren kann. Sie sind in der Regel Teil mehrstufiger Auswahlverfahren, in denen zuerst Gebietseinheiten (Bundesländer, Gemeinden etc.) zufällig oder nach bestimmten Kriterien

⁴ Neue Westfälische, 29.11.98.

⁵ §§8 und 9 des Reichs- und Staatsangehörigkeitsgesetzes sowie die Einbürgerungsrichtlinien.

festgelegt und innerhalb derselben dann Wohnungen nach festen Regeln (Begehungsanweisungen) ausgewählt werden. Ein bekanntes Beispiel ist das Random-Route-Verfahren (Verfahren des Zufallsweges), das vorschreibt, wie von zufällig vorgegebenen Startadressen in einer bestimmten Richtung gegangen und auf dem Weg jeder n-te Haushalt notiert wird, wobei das Verhalten des Adreßsammlers an Kreuzungen, Abbiegungen etc. exakt geregelt ist. Die wichtigsten Umfragen in der Bundesrepublik basieren auf dem sogenannten ADM-Design (des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e. V., siehe Schaefer 1979). Es sieht eine dreistufige Auswahl vor: erstens von Stimmbezirken, geschichtet unter anderem nach Gemeindegrößenklasse, zweitens von Haushalten innerhalb eines Stimmbezirks nach der Methode des Zufallsweges und drittens einer wiederum zufällig ausgewählten Person innerhalb des Haushalts.

Zuwanderer werden auf diese Weise in dem Anteil einbezogen, den sie an der Gesamtbevölkerung haben, wie etwa Schwarze und andere Gruppen (außer spanischsprechenden Minderheiten) im US-amerikanischen *Panel Study of Income Dynamics* (PSID, siehe Hill 1992). Ähnlich verfährt das *Panel Studie van Belgische Huishoudens* (PSBH, siehe Jacobs et al. 1991). PSID, PSBH und SOEP benutzen damit echte Wahrscheinlichkeitsstichproben. Für Analysen kleiner Minderheiten eignen sich solche Samples wegen des für sie zu geringen Umfangs nicht. So enthält das SOEP circa zwei Prozent „sonstige“ Ausländerhaushalte (ohne Vorstand aus einem der fünf Mittelmeerländer, siehe Hanefeld 1987). In der Welle 1997 entfallen laut SOEPINFO auf diese Restkategorie insgesamt 190 Personen 43 verschiedener Nationalitäten. Selbst nach großzügigen Zusammenfassungen sind auf dieser Grundlage kaum Aussagen über West- oder Osteuropäer, Außereuropäer oder etwa Flüchtlinge (abgesehen von Balkanflüchtlingen) möglich.

Prinzipiell kann die Methode für die Ziehung von Zuwandererstichproben abgewandelt werden, indem etwa in die Begehungsanweisungen die Vorschrift aufgenommen wird, nur die Adressen von Zuwandererhaushalten zu sammeln, indem also ein *Screening* vorgechaltet wird. Das britische *National Survey of Ethnic Minorities* (Modood et al. 1997) gewann sein Sample auf diese Weise. Wahlbezirke in England und Wales wurden je nach dem Anteil ethnischer Minderheiten in drei Schichten unterteilt, aus denen eine disproportional geschichtete Stichprobe (der Bezirke) gezogen wurde. In jedem Bezirk wurden dann Adressen aufgesucht, die zufällig aus einem Verzeichnis (dem *Postcode Address File*) entnommen worden waren, um zu ermitteln, ob dort Personen aus ethnischen Minderheiten wohnen. In Gebieten mittlerer Minderheitenanteile mußten die Interviewer zusätzlich jeweils die fünf Adressen zu beiden Seiten der vorgegebenen Anschrift ansteuern. In Gebieten niedriger Minderheitendichte wurden die Bewohner zufällig ausgewählter Adressen gefragt, ob in der Nähe Angehörige der Minderheiten leben, und aus den Nennungen ein oder zwei in Frage kommende Haushalte zufällig ausgewählt. Im PSID

werden sogenannte „Latino“-Haushalte (Mexikaner, Kubaner und Puertorikaner) in einer Zusatzstichprobe berücksichtigt, die in Gebieten mit überdurchschnittlichem Bevölkerungsanteil dieser Gruppen durch Random-Route gezogen wurde. Latino-Haushalte in Gebieten mit geringer Konzentration dieser Gruppen sind aus der Befragung ausgeschlossen, wofür explizit Kostengründe genannt werden (Duncan et al. 1995: 2). Eine ähnliche Einschränkung müssen Mehrländer et al. (1996: 474) bei einer Befragung vietnamesischer Vertragsarbeiter im Gebiet der ehemaligen DDR vornehmen.

Diese Methode ist mit einigen allgemeinen Fehlern behaftet. So tritt ein sogenanntes *Undercoverage*-Problem auf, weil die Anstaltsbevölkerung, also Heimbewohner und eine Reihe anderer Gruppen außerhalb von Privathaushalten, darunter auch Flüchtlinge in Unterkünften, ausgeschlossen werden (siehe zum Beispiel Schnell 1991; Diekmann 1998: 357ff.). Bedingt durch den Modus der praktischen Durchführung haben Untermieter sowie Haushalte in Hinterhofwohnungen und an anderen bei der Begehung schwer erkennbaren Adressen eine geringere Auswahlwahrscheinlichkeit. Wie Koch (1997) zeigt, besteht stets die Gefahr, daß die Interviewer, denen ja keinen konkreten Zielhaushalte vorgegeben werden, leicht erreichbare und kooperationswillige Teilnehmer bevorzugen. Alte, Ledige und Alleinlebende, Personen mit geringer Schulbildung und niedriger Berufsposition sowie Ausländer werden deutlicher unterrepräsentiert als es bei melderegisterbasierten Verfahren der Fall ist (die von solchen Effekten ebenfalls nicht frei sind). Daneben treten bei Zuwandererstichproben weitere Schwierigkeiten auf:

a) Das Screening wird um so ineffizienter und um so teurer, je kleiner der Bevölkerungsanteil einer Zuwanderergruppe ist. Zwar kann der Ausländeranteil in bestimmten Quartieren industriell geprägter Städte 50 Prozent übersteigen. Man muß sich aber vor Augen führen, daß Spitzenwerte des Anteils der größten ausländischen Nationalität in der Bundesrepublik, der Türken (einschließlich der Kurden aus dem Staatsgebiet der Türkei) sich auf städtischer Ebene nur in einer Größenordnung von fünf bis acht Prozent bewegen (Statistisches Bundesamt 1995: 22f.). Alle anderen Gruppen machen in der Regel weniger als ein Prozent der Wohnbevölkerung aus. Man kann leicht überschlagen, daß hier der Aufwand für das Screening die Kosten der Befragung weit übersteigt. Und da für die meisten großen Zuwanderergruppen Indizes der räumlichen Segregation seit den achtziger Jahren sinken (Friedrichs 1998), ist mit einer weiteren Verschlechterung des Verhältnisses von Aufwand und Ertrag zu rechnen. Durch Aufnahme einer Screener-Frage in mehrere Bus-Befragungen eines Instituts läßt sich unter günstigen Umständen noch eine vertretbare Lösung finden. Die Grenzen sind jedoch rasch erreicht, wie ein einfaches Beispiel zeigt. In Berlin, einer Hochburg der Vietnamesen in der Bundesrepublik, waren am

31.12.98 7.855 Personen vietnamesischer Staatsangehörigkeit registriert.⁶ Eine Dreisatzrechnung ergibt, daß bei 3.358.235 Einwohnern in der Stadt für eine Stichprobe von nur 300 Personen der Zielgruppe - von Teilnahmeverweigerungen einmal ganz abgesehen - 128.258 Screenings stattfinden müßten.

b) Einige technische Probleme wie die Suche in umfangreichen Namenslisten zur Identifikation der Gruppenzugehörigkeit nach Türschildern lassen sich inzwischen mit tragbaren PCs lösen (siehe dazu Abschnitt 5). Wenn dagegen persönliche Vorsprachen notwendig sind, weil für die Zielgruppe solche Verfahren nicht funktionieren, wird man bei Recherchen in der Nachbarschaft nicht immer alle Bewohner persönlich ansprechen können. Die Befragung ortsansässiger Informanten (etwa der Nachbarn, Ausländerbeauftragten oder Ausländervereine, wie bei Mehrländer et al. 1996: 475) birgt ähnliche Probleme wie das noch zu behandelnde Schneeballverfahren, weil sie vornehmlich die Namen von Personen mit umfangreichen Kontakten, entweder untereinander oder zu den Angehörigen der Mehrheitsgesellschaft, generiert. Da Nachbarn in vielen Fällen wohl nur Personen auffälliger Hautfarbe nennen werden, ihre Herkunft aber nicht kennen, ist das Verfahren ungeeignet, spezifische, vor allem im äußeren Erscheinungsbild unauffällige Gruppen zu identifizieren.

c) Werden wegen des dann vertretbaren Aufwands nur Wohngebiete mit hoher Ausländerdichte in die Auswahl einbezogen, kommt es zu Verzerrungen der Sozialstruktur der Zuwanderer. Sie leben vorwiegend nicht deshalb räumlich konzentriert, weil sie in der Nähe von Landsleuten wohnen wollen, sondern weil sie in bestimmten Stadtteilen erschwingliche Mieten vorfinden (Schuleri-Hartje 1984; Mehrländer et al. 1996). Auf Zuwanderer in „besseren Wohnlagen“ mit niedrigem Ausländeranteil zu verzichten hieße, wirtschaftlich bessergestellte Schichten unterproportional zu erreichen.

4. Das Schneeballverfahren

Die Träger seltener, ungewöhnlicher oder sozial unerwünschter Merkmale bilden Populationen, die die Stichprobenziehung vor besondere Probleme stellen. Listen stehen für sie oft nicht zur Verfügung, doch zugleich sind auch Zufallsverfahren unpraktikabel, weil ein Screening aufgrund der hohen Kosten und der möglicherweise geringen Akzeptanz nicht in Frage kommt. Mit dem Schneeballverfahren sind manche dieser Populationen dennoch erreichbar. Wenn die Angehörigen der Population untereinander Beziehungen unterhalten, kann man sich diese für die sukzessive Zusammenstellung einer Stichprobe zunutze ma-

⁶ Angaben des Statistischen Landesamtes Berlin in TOP-Berlin International 4/1999. Die Zahl der Eingebürgerten vietnamesischer Herkunft in Berlin ist nicht bekannt.

chen. Ausgehend von einer Initialstichprobe werden die bekannten Merkmalsträger gebeten, weitere Individuen zu benennen, die für eine Befragung angesprochen werden können (siehe die Beispiele bei Gabler 1992). Auf diese Weise kann entweder in mehreren Durchgängen fortgefahren werden, bis eine hinreichend große Stichprobe erzielt wurde, oder es kann eine Auswahlgesamtheit erzeugt werden, aus der dann zufällig eine Erhebungsstichprobe zu ziehen ist. Die Initialstichprobe kann auf verschiedene Weise gewonnen werden, zum Beispiel durch Kontakte zu Organisationen der Merkmalsträger oder Anfragen im Umfeld von Erhebungen bei Allgemeinbefragungen (Sudman/Kalton 1986: 411ff.).

Echte Zufallsstichproben kommen dabei aber nur unter optimalen Bedingungen zustande. Wenn das Verfahren nicht so weit fortgeführt werden kann, daß die Auswahlgesamtheit die Population annähernd vollständig umfaßt und damit quasi eine Totalerhebung vorgenommen wird, sind selektive Ausfälle wahrscheinlich. Die Stichprobe ist eng an die Struktur der Netzwerke gebunden, die die Merkmalsträger miteinander bilden. Individuen unterscheiden sich hinsichtlich des Umfangs ihrer persönlichen Netze und der Dichte ihrer Interaktionen. Wer viele Kontakte pflegt, hat eine größere Wahrscheinlichkeit, von Personen genannt zu werden, die bereits in die Stichprobe gelangt sind, und wer sozial isoliert ist, wird vom Schneeballverfahren übergangen (Sudman/Kalton 1986: 411ff.). Die Partizipation an Beziehungsnetzen ist aber ein Merkmal, daß bei Zuwanderern wie bei Einheimischen mit vielen anderen Variablen korreliert. In einer Untersuchung Gilanshahs (1990) geht es beispielsweise um diverse Einstellungen zur Eigengruppe und zur Entstehung einer ethnischen Gemeinde bei Iranern in Minneapolis-St. Paul. Da die Autorin keine Möglichkeit besitzt, eine Zufallsauswahl der Zielgruppe zu erreichen, befragt sie Mitglieder von Freundschaftsnetzen. Zum einen ist deshalb in ihrer Stichprobe die Bildungselite überrepräsentiert, und zum anderen ist ungesichert, inwiefern die berichteten Einstellungsangaben auf diejenigen Migranten übertragbar sind, die sich nicht an iranischen Netzwerken beteiligen. Weil hier die Zielvariablen in die Stichprobenziehung eingehen, muß mit tautologischen Befunden gerechnet werden. Bei Zuwanderern gilt ferner: Je weniger sie Beziehungen zu Landsleuten unterhalten und je mehr sie ihre Kontakte zu Einheimischen hin verlagern, desto geringer ist ihre Auswahlwahrscheinlichkeit, wenn Zuwanderer die Initialstichprobe bilden. Die soziale Integration in die Aufnahmegesellschaft ist nun aber erst recht ein Kriterium, das nicht in die Stichprobenziehung eingehen darf.

Bereits bei der Festlegung der Ausgangsstichprobe werden u. U. die Weichen für weitere Verzerrungen gestellt, weil Schneeballverfahren Gefahr laufen, in den Teilen der Sozialstruktur zu zirkulieren, in die sie anfänglich einsteigen. Zuwandererpopulationen können intern a) horizontal und b) vertikal ebenso geschichtet sein wie ihre Umgebung, obwohl dies für Einheimische wenig augenfällig ist und manchmal auf Kriterien basiert, die in der

Aufnahmegesellschaft keine Entsprechung finden (etwa auf Clan- oder Kasten-Strukturen). a) Migration findet meist kettenförmig statt. Abgesehen von Pioniermigranten erhalten Migrationsaspiranten Informationen über das Zielland meist von Personen aus der selben Herkunftsregion. Sie lassen sich in räumlicher Nähe zu ihnen nieder, und Teile von Dörfern werden so buchstäblich verpflanzt. Auch in Folgegenerationen gehen diese Strukturen nicht rasch verloren. Es bilden sich Segmente ähnlicher Herkunftsorientierung, die ein Schneeballverfahren nur über wenige Brücken überwinden kann (zu diesem Problem allgemein: Gabler 1992: 52f.). b) Orthogonal zu diesen Differenzierungen liegen vertikale Schichtungen aufgrund von Bildung, Einkommen oder auch Beruf, über deren Grenzen hinweg weniger Beziehungen bestehen als innerhalb einer Schicht.

Mit ähnlichen Schwierigkeiten muß übrigens bei Quotenvorgaben gerechnet werden (sowie bei Random-Quota-Mix-Verfahren, siehe den Zugang zu ausländischen Arbeitnehmern aus dem Mittelmeerraum in der BRD bei Mehrländer et al. 1996: 12ff.). Wenn Interviewern aufgetragen wird, Personen mit bestimmten Merkmalskombinationen zu befragen (die in einem definierten Verhältnis zur Verteilung der Merkmale in der Grundgesamtheit stehen), werden sie sich in der Regel der Kontakte bedienen müssen, die ihnen bereits befragte Individuen nennen. Wegen der naheliegenden Orientierung an Netzwerkkontakten kommt dabei indirekt eine Schneeballstichprobe zustande.

Die Erfahrungen im Sozio-oekonomischen Panel (SOEP) bestätigen die Skepsis gegenüber dem Schneeballverfahren. Das SOEP befragt seit 1992 Zuwanderer, die in den Jahren 1984 bis 1994 nach Westdeutschland gekommen sind (ohne Differenzierung nach Herkunftsländern und einschließlich Übersiedlern aus der DDR). Die Ziehung der Zuwanderer-Stichprobe D ist komplex angelegt (Schupp/Wagner 1995). Ein erster Teil der Adressen wurde per Screening im Zusammenhang mit Bus-Befragungen (Mehrthemenbefragungen) des Instituts gewonnen, das die SOEP-Erhebung im Feld implementiert (siehe Infratest Sozialforschung 1994). Ein zweiter Teil stammte aus einer Schneeballergänzung durch Adreßnennungen der Haushalte des ersten Teils. Diese Aufstockung war notwendig, weil per Screening zu wenige Zuwanderer identifiziert worden waren. In Auswertungen stellte sich einerseits heraus, daß die Stichprobe nichtdeutsche Zuwanderer generell untererfaßt (was unter anderem mit dem Einsatz eines deutschsprachigen Erhebungsinstruments zusammenhängt). Die Schneeballergänzung hatte diesen Effekt andererseits jedoch noch verstärkt, da Aussiedler, die im Gegensatz zu anderen Zuwanderern im Familienverband migrieren, überdurchschnittlich viele Schneeballadressen genannt hatten (Schupp/Wagner 1995: 22). So wie hier unterschiedliche Reichweiten der Netzwerke die Relation zwischen Herkunftsgruppen verfälschten, verzerren sie in anderen Zielgruppen das Verhältnis unterschiedlich sozial aktiver Persönlichkeitstypen. Für die weitere Ausdehnung der SOEP-Stichprobe D wurde der Schluß gezogen, auf Schneeballergänzungen

zu verzichten und nur noch Screening-Adressen einzusetzen - selbst um den Preis eines deutlich höheren Aufwands. Doch leider dürften den wenigsten Forschern die nötigen Ressourcen zur Verfügung stehen, wenn sie eine solche Entscheidung zu treffen haben.

Manche Zielgruppen lassen sich mit vertretbaren Mitteln wohl nur per Schneeball erreichen. Blahusch (1992) untersuchte die Integration von Flüchtlingen in den Arbeitsmarkt und definierte seine Grundgesamtheit als alle Flüchtlinge, die aufgrund des ausländer- und arbeitsrechtlichen Status prinzipiell eine Erwerbstätigkeit aufnehmen können. Der Umfang dieses Personenkreises ist unbekannt, und amtliche Register oder dergleichen existieren nicht. In Stadt und Landkreis Fulda wurde deshalb eine Totalerhebung im Schneeballverfahren versucht, ergänzt durch Kontrollgruppen in Frankfurt und Kassel im Umfeld von Beratungsstellen. Blahusch (1992: 48) berichtet von Angst und Skepsis gegenüber der Befragung, denen er durch den Einsatz von Projektmitarbeitern begegnete, die auch in der Flüchtlingsberatung tätig waren. Überraschend nennt er dann aber nur eine einzige Verweigerung bei 270 Interviews, was auf Selektionseffekte im Vorfeld der eigentlichen Befragung, nämlich im Schneeballverfahren, schließen läßt.

Ob das Verfahren für andere Befragungsformen als für face-to-face-Interviews genutzt werden kann, ist zweifelhaft. Da deutsche Sprachkenntnisse bei vielen Zuwanderern nicht vorausgesetzt werden können, ist der Einsatz muttersprachlicher Feldmitarbeiter anzuraten. Diese werden aber oft nur schwer glaubhaft machen können, daß sie weder im Auftrag deutscher Behörden noch bestimmter Exilorganisationen unterwegs sind, denen Migranten mißtrauen. Die Erfahrungen aus einer Studie des Autors des vorliegenden Beitrags mit Flüchtlingen, in der die Befragten auf als Interviewer tätige unbekannte Personen derselben Herkunftsgruppe sehr mißtrauisch reagierten, gibt aber Anlaß zu der Befürchtung, daß nur die wenigsten Zuwanderer bereit sind, Fremden Namenslisten ihrer Landsleute anzuvertrauen. Für die reine Sammlung Teilnahmewilliger analog zum Adreßrandom scheint die Methode also ungeeignet.

5. Namensbasierte Verfahren

Als letztes Stichprobenverfahren soll eine Technik vorgestellt werden, die in der Umfrageforschung selten eingesetzt wird, obwohl sie zum Standardrepertoire der gewerblichen Wirtschaft gehört, wenn es darum geht, private Haushalte zu erreichen. Das Verfahren macht sich die Tatsache zunutze, daß Telefonbücher mit gewissen Abstrichen Verzeichnisse der Haushalte der Wohnbevölkerung bilden und, daß viele Zuwandererhaushalte an ihren Namen zu erkennen sind. Hier kann nur das Grundprinzip erläutert werden, während Details und eine Schilderung der Erfahrungen in empirischen Untersuchungen bei Salentin (1994, 1997), Andreß/Lipsmeier/Salentin (1995) und Marhenke (1997) zu finden sind.

Die namensorientierte Identifizierung von Zuwanderern ist nicht neu (siehe zum Beispiel Borhek 1970; Driedger/Church 1974); neu sind aber die Möglichkeiten des großflächigen, EDV-gestützten Samplings. Die Verzeichnisse der Telefonanschlusshaber liegen nicht nur in gedruckter, sondern auch in maschinenlesbarer Form vor. Teils werden die Daten von einer Tochter der Deutschen Telekom selbst vermarktet, teils verschaffen sie sich Adreßverlage durch Scannen der Telefonbücher.⁷ Für den maschinellen Zugriff stehen bundesweite Verzeichnisse der Inhaber von Telefonanschlüssen und damit der Haushaltsvorstände grundsätzlich in zwei Formen zur Verfügung: 1. Adreßverlage vermieten Namen, Anschriften und Telefonnummern von Privatpersonen und Firmen vornehmlich für Direktmarketingzwecke (adressierte Werbung). Sie reichern ihre Bestände mit vielerlei Zusatzinformationen an, die gewisse Rückschlüsse auf die Lebensverhältnisse der an einem Wohnstandort ansässigen Haushalte zulassen. Dies ermöglicht es beispielsweise, in Stichproben bestimmte Kaufkraft- oder Altersgruppen überproportional zu berücksichtigen, wenn dies sinnvoll ist. 2. Die Inhalte einiger frei verkäuflicher Telefonauskunfts-CD-ROMs lassen sich vollständig, einige neue Ausgaben leider nur noch auszugsweise, in externe Dateien übertragen. Auf diese kann dann mit Datenbank-Programmiersystemen beliebig zugegriffen werden. Trotz vieler gegenteiliger Einschätzungen ist damit ein Adreßpool gegeben, der sich auch für sozialwissenschaftliche Stichprobenziehungen eignet. Die meisten Einwände lassen sich entkräften, wie gleich zu zeigen ist.

Für viele Herkunftsgruppen lassen sich nun namensbasierte heuristische Erkennungsalgorithmen konzipieren, die die Selektion von Zuwandererhaushalten in beliebig definierbaren Gebieten erlauben. Diese Algorithmen machen sich die Tatsache zunutze, daß Migranten für ihre Herkunftsländer typische Familiennamen wie auch ihre Nachkommen der zweiten und oft auch noch späterer Generationen charakteristische Vornamen tragen. Manche außereuropäische Ethnien benutzen zwar nicht das System aus Vor- plus Nachnamen (siehe zum Beispiel Vorsatz 1995), sind aber nichtsdestoweniger an Teilnamen oder markanten Silben zu erkennen. Bei der maschinellen Auswahl werden stets einige „falsche“ Namen ausgefiltert, so daß die Grobstichprobe von Muttersprachlern manuell nachselektiert werden muß, doch läßt sich auf diese Weise eine Auswahl durchaus akzeptabler Güte erzielen. Eine vollständige Identifikation der jeweiligen Populationen wird wegen abweichender Schreibweisen und ungewöhnlicher oder seltener Namen nicht erreicht, was aber unproblematisch ist, weil keine disproportionalen Ausfälle bestimmter Merkmalsträger zu erwarten sind. Dagegen kommt dem Verfahren zugute, daß in allen

⁷ Teils gelangen sie an die Hersteller von CD-ROMs auch auf Wegen, die Gegenstand gerichtlicher Auseinandersetzungen sind.

Gesellschaften bestimmte Namen häufig auftreten, weshalb auch mit begrenzten Namenslisten hohe Erkennungsraten zu erzielen sind. Mit den 100 häufigsten deutschen Vornamen können beispielsweise bereits 57 Prozent aller deutschen Telefonanschlüßinhaber identifiziert werden, mit 500 Vornamen 87 Prozent (eigene Auswertungen). Die Entwicklung der Algorithmen ist Aufgabe des Forschers. Direktmarketingverlage übernehmen zwar gegen Kostenübernahme die Programmierung der Stichprobenziehung, doch verfügt keines der vom Autor angeschriebenen überregional tätigen Unternehmen über fertige Programme. Wer die Möglichkeit hat, komplexe Datenbankabfragen selbst vorzunehmen, wird in der Regel durch eigene Programmierung über CD-ROM-Exporte Kosten sparen, da die Unternehmen alleine für diese Arbeiten Beträge von bis zu 10.000 DM je Gruppe in Rechnung stellen.

Nun zu einige Einwänden gegen diese Methode.

a) Denkbare Verzerrungen durch die Ausblendung von Haushalten ohne Telefonanschluß halten sich wegen der hohen Telefondichte in der Bundesrepublik in engen Grenzen. Je nach Haushaltstyp haben im Westen 97,1 bis 99,7 Prozent der Haushalte der Wohnbevölkerung einen Anschluß, im Osten liegen die Anteile nur um wenige Prozentpunkte darunter (Statistisches Bundesamt 1998: 547f.). Die offizielle Statistik weist leider Nationalitäten nicht separat aus; bei Ausländern wird man aber von einer geringeren Ausstattung mit Telefonanschlüssen ausgehen müssen.

b) Die Nutzbarkeit der Telefonbücher als Adreßquelle wird in der Hauptsache aufgrund der Existenz anonymer Anschlüsse in Abrede gestellt. Über den Anteil der Anschlüsse ohne Eintrag ins Telefonbuch bzw. mit einem Eintrag ohne vollständige Anschrift sind unterschiedliche Angaben im Umlauf.⁸ Offenbar steigt er nach Liberalisierung der Bestimmungen seit 1992 an. Häder (1996) untersucht, ob sich nichteingetragene oder anonyme Anschlußinhaber (im englischen Sprachgebrauch *nonpublished phone subscribers* oder *unlisted phone numbers*) von eingetragenen unterscheiden. Für die ostdeutschen Bundesländer, in denen nach der Wende das Telefonnetz beinahe neu entstand und wo die meisten Teilnehmer über die Eintragung ins Telefonbuch frei entscheiden konnten, kann sie einen direkten Vergleich zwischen anonymen und eingetragenen Anschlüssen anstellen. Eine Logit-Analyse ergibt, daß am stärksten die Gemeindegrößenklasse die Neigung zum Nichteintrag vorhersagt (in Großstädten gibt es die höchste Quote anonymer Anschlüsse, weil dort offenbar die Furcht vor Belästigungen größer ist), gefolgt vom Vor-

⁸ Nach unseren Recherchen (Salentin 1997) beläuft sich der Anteil vollkommen anonymer Anschlüsse auf circa vier Prozent; Häder (1996) nennt 3,7 Prozent. Eine durch von der Heyde (1997: 200) publizierte Infratest-Studie weist 23 Prozent anonyme oder mit fehlender/unvollständiger Anschrift versehene Anschlüsse aus.

handensein von Kindern im Haushalt: Wenn genau ein Kind zum Haushalt gehört, ist ein anonymer Anschluß am wahrscheinlichsten, was die Autorin mit dem Wunsch der Eltern nach störungsfreiem Schlaf erklärt. Es zeichnet sich nur eine schwache Polarisierung nach dem Einkommen ab: Sowohl die unterste wie auch die oberste Einkommensklasse neigt etwas eher zum Nichteintrag. (Auch bei Blankenship (1977) bleiben die Ergebnisse hinsichtlich sozialer Statusmerkmale inkonsistent.) Eingetragene und nichteingetragene Befragtengruppen unterscheiden sich statistisch nicht hinsichtlich Einstellungsdimensionen, der politischen Partizipation und der Beteiligung an lokalen Netzwerken (Häder 1996: 63). von der Heyde (1997) ermittelt eine in Ost- und Westdeutschland konsistent geringere Eintragsbereitschaft der sozial Schwächeren, wenngleich mit geringem Gewichtungsaufwand dennoch hinsichtlich Erwerbsstatus, Haushaltsnettoeinkommen etc. abbildungstreue Telefonstichproben zu erzielen seien. Da im Osten jedoch Schichtungs- und Haushaltstypeneffekte bei manchen Merkmalskombinationen zu dramatischen, der Korrektur durch Gewichtung nicht mehr zugänglichen Ausfällen kumulieren, rät er von der Nutzung der Telefonbücher als Adreßbasis ab.

c) Ein potentielles Problem des Samplings auf Telefonbuchbasis ist entstanden, seit der Gebrauch von Mobiltelefonen (Handys) in den neunziger Jahren stark zunimmt. Die Neigung, die Rufnummer der Handys in Telefonbüchern veröffentlichen zu lassen, ist gering. Zwar wollen (oder können) die Netzbetreiber keine genauen Angaben über den Anteil der Anschlüsse machen, deren Rufnummern unveröffentlicht bleiben.⁹ Schätzungen dieser Unternehmen und eines Adreßverlages zufolge, der mit Daten aus Telefonverzeichnissen handelt,¹⁰ beträgt der Anteil der in die Telefonbücher eingetragenen Handy-Rufnummern nur 10 bis 30 Prozent des gesamten Handy-Bestandes. Man kann aus dieser geringen Quote allerdings wiederum nicht auf eine bedeutende Unterrepräsentation der Handy-Nutzer in den Telefonbüchern schließen, da der größere Teil von ihnen ein Mobiltelefon zusätzlich zum Festnetzanschluß nutzt. Es kommt offenbar nur selten vor, daß neben dem Mobiltelefon kein Festnetztelefon benutzt wird, was aufgrund der beträchtlich höheren Kosten beim Mobiltelefonieren auch plausibel ist. Wenn aber Festnetzanschlüsse der Mobiltelefonbesitzer vorhanden sind, gibt es keine Anzeichen dafür, daß der Anteil der nicht eingetragenen Anschlüsse unter ihnen überdurchschnittlich hoch ist.

d) Die Aktivitäten der neuen Telefongesellschaften neben der Deutschen Telekom beeinflussen nicht die Veröffentlichung der Anschlußdaten. Die Dienste dieser Anbieter können zur Zeit mit wenigen Ausnahmen nur für Ferngespräche in Anspruch genommen

⁹ Es wurden dazu am 18.08.98 telefonische Anfragen bei T-Mobil in Bonn, dem Betreiber von D1, und bei Mannesmann in Düsseldorf, dem Betreiber von D2, gestellt.

¹⁰ Schober Direktmarketing, Ditzingen, Anfrage vom 18.08.98.

werden; fast alle Festnetzanschlußinhaber sind nach wie vor Kunden der Deutschen Telekom und werden folglich auch in den Telefonbüchern geführt. Auch mit der Freigabe des Marktes für Ortsgespräche wird sich in dieser Beziehung nichts ändern, weil weiterhin die Herausgabe eines einheitlichen Gesamttelefonbuchs vorgesehen ist.

e) Ein Nachteil der telefonbuchbasierten Verfahren liegt in der geringeren Aktualität der Daten. Nach Angaben der Deutschen Telekom (siehe auch Fuchs 1994: 154) ändern sich im Gesamtbestand der Telefonanschlüsse jährlich 30 Prozent der Eintragungen. Obwohl keine differenzierten Aufschlüsselungen vorliegen, dürfte indessen der größere Teil der Änderungen auf die gewerblichen Einträge entfallen und unter den Privatanschlüssen weniger Fluktuation herrschen. Ferner sind nicht alle Änderungen der Einträge identisch mit Anschlußinhaberwechseln, da im oben genannten Prozentsatz ja auch Personenstandsänderungen enthalten sind. Schließlich sind Änderungen vor allem in den besonders mobilen Bevölkerungsgruppen, zum Beispiel bei jüngeren, in Ausbildung befindlichen Personen, anzunehmen (Schröder 1989: 107ff.). Es ist aber nicht zu leugnen, daß die Daten gegenüber Melderegisterstichproben im Lauf der Zeit an Aktualität verlieren, weil Telefonbücher, aus denen die Daten gewonnen werden, nur im jährlichen Rhythmus aufgelegt werden. Allerdings relativiert sich dieser Nachteil, wenn man sich vergegenwärtigt, daß, wie Hormuth/Brückner (1985) argumentieren, auch Melderegister nicht alle Personen enthalten, die in einer Kommune faktisch wohnen. Diese Autoren stellen sogar die These auf, daß die Anmeldung eines Telefons in der Regel der polizeilichen Meldung vorausgeht, weil sie zu den ersten Dingen gehört, die nach einem Umzug wesentlich erscheinen.

f) Nicht in allen Herkunftsgruppen ist eine hinreichend zuverlässige Namenserkennung möglich. Während sie bei Türken, Griechen, Italienern, Personen aus den Nachfolgestaaten Jugoslawiens und vielen anderen Gruppen funktioniert, stößt sie beispielweise bei iberischen und muslimischen Namen an ihre Grenzen, da diese von Zuwanderern aus mehreren Kontinenten geführt werden.

g) Da die Inhaber von Telefonanschlüssen meist Haushaltsvorstände sind, werden in Stichproben Männer über 25 Jahre überrepräsentiert. (Andererseits sind die Anschlüsse vieler Witwen noch auf den Namen der Ehemänner gemeldet.) Die namentlich eingetragenen Frauen sind nicht mit der weiblichen Wohnbevölkerung identisch, da unter ihnen vor allem Hausfrauen fehlen. Es ist daher unerlässlich, in einer nachgeschalteten Auswahlstufe für eine zufallsgesteuerte Bestimmung der befragten Person innerhalb des Haushalt zu sorgen. Auch lassen sich nur selten beide Ehepartner in Mischehen zwischen Zuwanderern und Deutschen als Anschlußinhaber eintragen. Häufige Fehlzuordnungen sind bei Partnerschaften unter Beteiligung solcher Zuwanderergruppen zu erwarten, die

als sog. Heiratsmigranten (Ruenkaew 1997) in die Bundesrepublik einreisen, etwa Thailänderinnen, Filipinas, Ghanaer und Ghanaerinnen, die trotz numerisch geringer Präsenz eine große Zahl von Mischehen eingehen (siehe Statistisches Bundesamt 1996: 92f.)

h) Generell lassen abgekürzte oder fehlende Vornamen keine Zuordnung zu. Gewerbliche und private Anschlüsse sind mitunter nicht eindeutig zu trennen.

Viele praktische Vorteile der Methode liegen dagegen auf der Hand. Zunächst entfällt die Einschränkung auf den rechtlichen Ausländerstatus, da auch Eingebürgerte anhand ihrer Namen bestimmten Gruppen zugeordnet werden können. In vielen Fällen ist eine Differenzierung nach Volksgruppenzugehörigkeit realisierbar. Zu erwähnen ist daneben die leichte Verfügbarkeit für das gesamte Bundesgebiet zu niedrigen Kosten. Es lassen sich beliebige Gebietsauswahlen treffen, sei es nach Bundesländern, Gemeindegrößen oder sonstigen Merkmalen. Auch der ländliche Raum wird problemlos erfaßt. Außer dem Anschaffungspreis einer CD-ROM fallen lediglich Aufwendungen für Programmierarbeiten an.

Um die Tauglichkeit dieses Verfahrens zu prüfen, hat eine Bielefelder Arbeitsgruppe bei einer Studie zur Situation tamilischer Flüchtlinge versuchsweise mit geringem Aufwand ortsansässige tamilische Telefoninhaber identifiziert. Dazu wurden alle 160.000 Telefonanschlüsseinträge Bielefelds aus der CD-ROM „D-Info 97“ in eine Datei exportiert. Mit Hilfe einer Liste typischer Namensbestandteile ließen sich darin 137 tamilische Namen auffindig machen.¹¹ In der Stadt lebten Ende 1996 (zum Zeitpunkt der Herausgabe der CD) 650 Personen mit der Staatsangehörigkeit Sri Lankas. In dieser Zahl fehlen zwar die Eingebürgerten, sie enthält aber statt dessen in etwa dem gleichen Umfang Angehörige nicht-tamilischer Volksgruppen dieser Nationalität, so daß sich Unter- und Überschätzung der Zielgruppe die Waage halten. Der Einfachheit halber läßt sich ein Telefonanschluß mit einem Haushalt gleichsetzen. Der durchschnittliche Haushaltsumfang in der Gruppe ist nicht exakt bekannt. Da es sich überwiegend um Paare im Alter zwischen 25 und 45 Jahren mit Kindern handelt und Einpersonenhaushalte die Ausnahme bilden, ist eine Schätzung von drei Personen je Haushalt realistisch. Somit wurden hochgerechnet 411 Personen erfaßt, was einer Ausschöpfung der Population von über 63 Prozent entspricht. Diese Rate dürfte sich mit entsprechendem Namensrecherche- und Programmieraufwand weiter steigern lassen.

¹¹ Uns liegen umfangreiche Listen türkischer, italienischer, jugoslawischer, griechischer, vietnamesischer und tamilischer Namen vor.

6. Resümee

Melderegisterstichproben haben den Vorteil, alle in einer Kommune gemeldeten Personen einer Staatsangehörigkeit zu erfassen. Leider schließen sie eingebürgerte Migranten aus, und immer, wenn Merkmale untersucht werden, in denen mit der Einbürgerung Veränderungen anzunehmen sind, müssen Verzerrungen befürchtet werden. Behördliche Meldung und faktischer Aufenthalt stimmen bei Ausländern mehr noch als bei Deutschen nicht unbedingt überein, so daß auch unbrauchbare Adressen gezogen werden. Der ländliche Raum ist mit diesem Verfahren nur schwer zu erschließen, weil eine große Zahl von Kommunen einbezogen werden muß. Die Kosten steigen mit der gewünschten regionalen Streuung. Melderegisterstichproben eignen sich daher vorwiegend, wenn die Beschränkung auf wenige städtische Populationen vertretbar ist. Bei einer Wiederholungsbefragung einer identischen Stichprobe lassen sich schließlich Umzüge in andere Kommunen nur mit Recherchen außerhalb der ursprünglichen Quelle der Adressen nachverfolgen.

Auf der Grundlage der Begehung von Gebietseinheiten sind wegen des Screening-Aufwands Stichproben nur für wenige Herkunftsgruppen rationell zu ziehen. Man handelt sich mit Versuchen, den Aufwand zu reduzieren, leicht Verzerrungen der Sozialstruktur ein; Recherchen in der Nachbarschaft führen zur Unterrepräsentation kontaktparmer Personen. Das Risiko derartiger Schiefen scheint beim Schneeballverfahren noch größer zu sein, weil es vollständig auf Netzwerkkontakten innerhalb einer Population basiert. Schneeballziehungen ergeben im allgemeinen keine Zufallsstichproben. Sie sind nichtsdestoweniger der einzige Weg, mit geringem Aufwand Gruppen wie Spätaussiedler zu erreichen, die sich nur sehr lückenhaft auf dem Weg über die Staatsangehörigkeit oder über Namen identifizieren lassen. Namensorientierte Verfahren auf der Grundlage von Telefonteilnehmerverzeichnissen erfassen die Haushalte unabhängig von der Staatsangehörigkeit und Beziehungen der Zuwanderer untereinander. Sie sind sehr kostengünstig, erlauben flexible regionale Differenzierungen und ermöglichen es mit gewissen Einschränkungen, Umzüge im Bundesgebiet für eine Zweitbefragung zu verfolgen. Die vorliegenden Erkenntnisse legen allerdings den Schluß nahe, daß bei sozial Schwächeren und jungen Familien Abbildungsdefizite auftreten. Ein weiterer Nachteil liegt in der geringeren Aktualität der Grundgesamtheit gegenüber allen anderen Verfahren.

Die Erörterung hat verdeutlicht, daß kein einzelnes Verfahren gleichzeitig auf eine vollständige Grundgesamtheit zurückgreifen und verzerrungsfreie Stichproben garantieren kann. Vor- und Nachteile der Techniken müssen daher in Anhängigkeit von Zielpopulationen, untersuchten Merkmalen und verfügbaren Ressourcen im Einzelfall abgewogen und gegebenenfalls Optimierungen durch Methodenkombinationen angestrebt werden.

Ein optimales Verfahren gibt es nicht: Wer Zuwanderer befragen will, wird sich für das kleinere Übel entscheiden müssen.

Korrespondenzadresse

*Dr. Kurt Salentin
Universität Bielefeld
Fakultät für Soziologie
33501 Bielefeld
E-Mail salentin@sozjur.uni-bielefeld.de*

Literatur

- Andreß, H.-J./Lipsmeier, G./Salentin, K., 1995: Erfahrungen mit Direktmarketingadressen: Bevölkerungsumfragen im unteren Einkommensbereich. ZUMA-Nachrichten 37: 163-187.
- Beauftragte der Bundesregierung für Ausländerfragen 1998: Daten und Fakten zur Ausländersituation. Bonn.
- Blahusch, F., 1992: Flüchtlinge auf dem deutschen Arbeitsmarkt: Eine empirische Untersuchung am Beispiel Hessens. Frankfurt: Haag und Herchen.
- Blankenship, A. B., 1977: Listed Versus Unlisted Numbers in Telephone-Survey Samples. Journal of Advertising Research 17: 39-42.
- Boos-Nünning, U., 1990: Einwanderung ohne Einwanderungsentscheidung: Ausländische Familien in der Bundesrepublik Deutschland. Aus Politik und Zeitgeschichte B 23-24: 16-25.
- Borhek, J. T., 1970: Ethnic-Group Cohesion. American Journal of Sociology 76: 33-46.
- Diekmann, A., 1998: Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt (4., durchgesehene Auflage).
- Driedger, L./Church, G., 1974: Residential Segregation and Institutional Completeness: A Comparison of Ethnic Minorities. Canadian Review of Sociology and Anthropology 11: 30-52.
- Duncan, G. J./Hill, M. S./Morgan, J. N./Hofferth, S. J., 1995: A Panel Study of Income Dynamics: Procedures and Tape Codes. (Documentation) 1990 Interviewing Year. Volume I: Procedures and Tape Codes, Wave XXIII. A Supplement. Ann Arbor, Michigan: University of Michigan, Institute for Social Research.
- Fendrich, P., 1986: Integration kommt voran. Bundesarbeitsblatt 27: 9-12.
- Friedrichs, J., 1979: Methoden der empirischen Sozialforschung. Reinbek: Rowohlt.
- Friedrichs, J., 1998: Ethnic Segregation in Cologne, Germany, 1984-94. Urban Studies 35: 1745-1763.

- Fuchs, M., 1994: Umfrageforschung mit Telefon und Computer: Einführung in die computergestützte telefonische Befragung. Weinheim: Beltz.
- Gabler, S., 1992: Schneeballverfahren und verwandte Stichprobendesigns. ZUMA-Nachrichten 31: 47-69.
- Gilanshah, F., 1990: The Formation of Iranian Community in the Twin Cities from 1983-89. *Wisconsin Sociologist* 27: 11-17.
- Häder, S., 1996: Wer sind die „Nonpubs“? Zum Problem anonymer Anschlüsse bei Telefonumfragen. ZUMA-Nachrichten 39: 45-68.
- Hanefeld, U., 1987: Das Sozio-ökonomische Panel: Grundlagen und Konzeption. Frankfurt: Campus
- Hill, M. S., 1992: The Panel Study of Income Dynamics: A User's Guide. Newbury Park: Sage.
- Hormuth, S. E./Brückner, E., 1985: Telefoninterviews in Sozialforschung und Sozialpsychologie: Ausgewählte Probleme der Stichprobengewinnung, Kontaktierung und Versuchsplanung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 37: 526-545.
- Infratest Sozialforschung 1994: SOEP '94 Zuwanderer-Befragung (Teilstichprobe D 1): Methodenbericht. München: Infratest Sozialforschung.
- Jacobs, T./Loots, I./Marynissen, R./Scheipers, T., 1991: Panel Studie van Belgische Huishoudens: Onderzoeksplan. o.O. (mimeo, Internet-Version: <http://hgins.uia.ac.be/psbh/>)
- Kirschner, H.-P., 1984: ALLBUS 1980: Stichprobenplan und Gewichtung. S. 114-182 in: K.-U. Mayer/P. Schmidt (Hrsg.), *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften: Beiträge zu methodischen Problemen des ALLBUS 1980*. Frankfurt: Campus (Monographien Sozialwissenschaftliche Methoden, Bd. 5)
- Koch, A., 1997: ADM-Design und Einwohnermelderegister-Stichprobe: Stichprobenverfahren bei mündlichen Bevölkerungsumfragen. S. 99-116 in: S. Gabler/J. H. P. Hoffmeyer-Zlotnik (Hrsg.), *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Lederer, H. W., 1997: Migration und Integration in Zahlen: Ein Handbuch. Bonn: Beauftragte der Bundesregierung für Ausländerfragen.
- Marhenke, W., 1997: Telefonanschlußdaten als Auswahlgrundlage. S. 207-220 in: S. Gabler/J. H. P. Hoffmeyer-Zlotnik (Hrsg.): *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Mehrländer, U./Ascheberg, C./Ueltzhöffer, J., 1996: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik (Repräsentativuntersuchung '95). Berlin, Bonn, Mannheim: Bundesministerium für Arbeit und Sozialordnung (Forschungsbericht Nr. 263).
- Modood, T./Berthoud, R./Lakey, J./Nazroo, J./Smith, P./Virdee, S./Beishon, S., 1997: *Ethnic Minorities in Britain: Diversity and Disadvantage*. London: Policy Studies Institute.

- Ordemann, H.-J., 1988: Paßrecht, Ausweisrecht, Melderecht des Bundes. München: Beck.
- Ruenkaew, P., 1997: Thailändische Frauen heiraten nach Deutschland: Motive und Hintergründe von Heiratsmigration. *Forschung an der Universität Bielefeld* 16: 32-38.
- Salentin, K., 1994: Direktmarketingadressen für Umfragezwecke? *ZUMA-Nachrichten* 35: 54-65.
- Salentin, K., 1997: Adressen und Zusatzinformationen: Erfahrungen mit Stichproben aus Adreßverlagen des Direktmarketing. S. 49-62 in: W. Sodeur (Hrsg.): *Regionale Analyse mit kleinen Gebietseinheiten*. Opladen: Leske & Budrich.
- Schaefer, F., 1979: *Muster-Stichproben-Pläne für Bevölkerungsstichproben in der Bundesrepublik Deutschland und West-Berlin*. München: Verlag Moderne Industrie
- Schnell, R., 1991: Wer ist das Volk? Zur faktischen Grundgesamtheit bei „allgemeinen Bevölkerungsumfragen“: Undercoverage, Schwererreichbare und Nichterreichbare. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 43: 106-137.
- Schröder, H., 1989: Sample-Planung und Sample-Realisierung bei Telefonumfragen: Theorie und Praxis einer medizinsoziologischen Studie. S. 105-114 in: G. Lüschen (Hrsg.): *Gesundheit und Krankheit in der BRD und den USA*. München: Oldenbourg.
- Schuleri-Hartje, U.-K., 1984: Wohnen im fremden Land. *Zeitschrift für Ausländerrecht und Ausländerpolitik* 4: 90-94.
- Schupp, J./Wagner, G., 1995: Die Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP). *Vierteljahrshefte zur Wirtschaftsforschung* 64: 16-25.
- Seifert, W., 1997: Integration von Ausländern. S. 579-589 in: Statistisches Bundesamt (Hrsg.): *Datenreport 1997: Zahlen und Fakten über die Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung.
- Statistisches Bundesamt 1995: *Bevölkerung und Erwerbstätigkeit*. Fachserie 1, Reihe 2, Ausländer, 1994. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt 1996: *Bevölkerung und Erwerbstätigkeit*. Fachserie 1, Reihe 1, Gebiet und Bevölkerung, 1995. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt 1998: *Statistisches Jahrbuch für die Bundesrepublik Deutschland 1998*. Stuttgart: Metzler-Poeschel.
- Sudman, S./Kalton, G., 1986: New Developments in the Sampling of Special Populations. *Annual Review of Sociology* 12: 401-429.
- von der Heyde, Christian 1997: Random-Route und Telefon: Struktur von Telefonhaushalten. S. 196-206 in: S. Gabler/J. Hoffmeyer-Zlotnik (Hrsg.): *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Vorsatz, R., 1995: Die vietnamesische Minderheit. S. 532-545 in: C. Schmalz-Jacobsen/G. Hansen (Hrsg.): *Ethnische Minderheiten in der Bundesrepublik Deutschland: Ein Lexikon*. München: Beck.

Weber, M., 1964: *Wirtschaft und Gesellschaft: Grundriß der verstehenden Soziologie*. Köln, Berlin: Kiepenheuer & Witsch.

KOSTENGÜNSTIGER ZUGANG ZU AMTLICHEN MIKROZENSUSDATEN DER JAHRE 1989, 1991, 1993 UND 1997

Beim Statistischen Bundesamt können faktisch anonymisierte Mikrozensusdaten der Jahre 1989, 1991, 1993 und 1997 zu Sonderkonditionen (Bereitstellungsgebühr von jeweils 130 Mark) für Forschungszwecke bestellt werden. Ermöglicht wurde dies durch eine Vereinbarung zwischen dem Statistischen Bundesamt und der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS e.V.), welche die Grundkosten für die Bereitstellung der Mikrozensusdaten übernimmt.

Die Daten der amtlichen Statistik sind für die empirischen Sozial- und Wirtschaftswissenschaften eine der wichtigsten Informationsquellen im Rahmen der Analyse wirtschafts- und gesellschaftspolitischer Zusammenhänge. Dies gilt insbesondere für den Mikrozensus, der als größte Mehrthemenumfrage der Bundesrepublik Deutschland vielfältige Informationen über die wirtschaftliche und soziale Lage der Bevölkerung enthält.

Das Bundesstatistikgesetz von 1987 erlaubt den statistischen Ämtern die Weitergabe von Individualdaten an die Wissenschaft, wenn sie so anonymisiert sind, daß sie Auskunftgebenden nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können. Man spricht auch von faktisch anonymisierten Datenfiles oder Scientific-Use-Files. Während in der Vergangenheit hohe Kosten für die Weitergabe der Daten eine breite Nutzung durch die Forschung erschwerten, können seit 1997 durch eine gemeinsame Initiative des Statistischen Bundesamtes, des Bundesministeriums für Bildung und Forschung und der GESIS amtliche Daten, wie z.B. der Mikrozensus von 1995 und 1996, in Form von Scientific-Use-Files für Forschungszwecke kostengünstig bezogen werden. Diese Vereinbarung hat dazu geführt, daß bisher rund 100mal von dem Angebot der Mikrozensusdaten Gebrauch gemacht wurde.

Um diese Verbesserung bei der Nutzung der Daten weiter zu fördern und den Zeitraum bis zu einer eventuellen längerfristigen Konsolidierung zu überbrücken, haben die GESIS und das Statistische Bundesamt eine Ausweitung der für die Forschung zugänglichen Datenbasis des Mikrozensus um die Erhebungen 1989, 1991, 1993 und 1997 vereinbart. Die GESIS übernimmt dabei die Grundkosten für die Bereitstellung dieser Mikrozensusdaten.

damit sie das Statistische Bundesamt der Wissenschaft zu einem Preis von jeweils 130 Mark zur Verfügung stellen kann. Die Daten von 1989, 1991 und 1993 können ab sofort bezogen werden; die Daten des Mikrozensus 1997 werden voraussichtlich ab Anfang nächsten Jahres zugänglich sein.

Um eine effiziente wissenschaftliche Nutzung der faktisch anonymisierten Mikrozensusdaten zu ermöglichen, stehen interessierten Wissenschaftlern Ansprechpartner im Statistischen Bundesamt und im Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim zur Verfügung. Zu den Scientific-Use-Files sind im World Wide Web (<http://www.zuma-mannheim.de/data/microdata>) u.a. folgende Informationen abrufbar: Datensatzbeschreibung, Fragebogen, rechtliche Voraussetzungen der Datenlieferung und Bestellmöglichkeiten.

Nähere Auskünfte erteilen:

Norbert Schwarz
Statistisches Bundesamt, Zweigstelle Bonn
Gruppe IX B – Mikrozensus
Graurheindorfer Str. 198
53117 Bonn
Tel.: 01888/643-8692, Fax –8962
E-mail: mikrozensus@statistik-bund.de

Bernhard Schimpl-Neimanns
ZUMA, Abteilung Mikrodaten
Postfach 12 21 55
68072 Mannheim
Tel.: 0621/1246-263, Fax: -100
E-mail: schimpl-neimanns@zuma-mannheim.de

THESEN ZUR VERBESSERUNG DER INFORMATIONELLEN INFRASTRUKTUR IN DER BUNDESREPUBLIK DEUTSCHLAND

Durch das Pilotprojekt zur Bereitstellung faktisch anonymisierter Mikrodaten ist die Nutzung dieser nunmehr kostengünstig zugänglichen Daten erheblich gestiegen. Darüber hinaus werden weitergehende Möglichkeiten zur Verbesserung des Verhältnisses zwischen amtlicher Statistik und Forschung verstärkt in Fachkreisen und selbst in den Medien diskutiert. Daher haben - auf Anregung und mit finanzieller Förderung des Bundesministeriums für Bildung und Forschung (BMBF) - das Deutsche Institut für Wirtschaftsforschung (DIW), die GESIS und das Statistische Bundesamt ein Symposium zum Thema „Kooperation zwischen Wissenschaft und amtlicher Statistik - Praxis und Perspektiven“ organisiert, das am 31. Mai und 1. Juni 1999 in Wiesbaden stattfand. Ziel der Veranstaltung war es, den Status Quo der Zusammenarbeit zu erörtern sowie über Perspektiven für deren Weiterentwicklung zu diskutieren. Um hierbei Modelle der Zusammenarbeit in anderen Ländern näher kennenzulernen, wurden insbesondere auch mehrere ausländische Referenten eingeladen, über die Erfahrungen in ihren Ländern zu berichten. In einer abschließenden Paneldiskussion wurden die wesentlichen Ergebnisse der Tagung noch einmal zugespitzt.¹

Erfreulicherweise sind schon bald nach der Veranstaltung weitere Initiativen zur intensiveren Kooperation zwischen amtlicher Statistik und Forschung ergriffen worden. Direkt im Anschluß an das Symposium haben die Teilnehmer der Paneldiskussion gemeinsam Thesen zur Verbesserung der informationellen Infrastruktur in der Bundesrepublik Deutschland erarbeitet. Vor dem Hintergrund des Erfahrungsaustauschs auf dem Symposium und auf Basis dieser Thesen hat die Bundesministerin für Bildung und Forschung, Frau Edelgard Bulmahn, eine Kommission einberufen, die den Auftrag hat, Lösungsvorschläge zu den angesprochenen Fragen und Problembereichen zu erarbeiten. Diese Kommission hat im Oktober 1999 ihre Arbeit aufgenommen. Im folgenden dokumentieren wir diese Entwicklung mit dem Abdruck der Thesen der Teilnehmer der Paneldiskussion sowie einer Mitteilung des BMBF zur Einsetzung der Kommission.

¹ Die Beiträge sind in folgendem Band erschienen: Kooperation zwischen Wissenschaft und amtlicher Statistik - Praxis und Perspektiven. Beiträge zum Symposium am 31. Mai/1. Juni 1999 in Wiesbaden. Band 34 der Schriftenreihe Forum der Bundesstatistik, herausgegeben vom Statistischen Bundesamt. Metzler-Poeschl, Stuttgart 1999 (Bestell-Nr.: 1030434-99900; ISBN: 3-8246-0608-9; 32,80 DM).

Thesen zur Verbesserung der informationellen Infrastruktur in der Bundesrepublik Deutschland

1. Die Bedeutung der informationellen Infrastruktur für Staat, Wirtschaft und Gesellschaft ist sowohl quantitativ wie qualitativ gestiegen. Fehlende Informationen verstellen politische Lösungswege oder behindern Entwicklungen. Wegen der grundlegenden Bedeutung aktueller, aufbereiteter Daten für die verantwortliche Gestaltung der öffentlichen Angelegenheiten in der Demokratie finden Fragen der Statistik in jüngster Zeit wieder mehr politische Aufmerksamkeit (z. B. das Green Paper der Britischen Regierung von 1998 „Statistics – A matter of trust“).
2. Wissenschaft und amtliche Statistik sind zentrale Teile dieser informationellen Infrastruktur. Ihre stärkere Zusammenarbeit ist Erfolgsbedingung für die Funktionsfähigkeit der informationellen Infrastruktur in unserer Demokratie.
3. Zur Erfüllung ihrer gemeinsamen Aufgabe bedürfen beide der Unabhängigkeit mit entsprechender institutioneller Ausgestaltung. Die Sozial- und Wirtschaftswissenschaften sollten berücksichtigen, daß sie Verantwortung für die informationelle Infrastruktur mittragen.
4. Der Gesetzgeber hat im Rahmen der Novellierung des Bundesstatistikgesetzes im Jahr 1987 mit der Aufnahme des § 16 Abs. 6 anerkannt, daß eine besondere Beziehung zwischen Statistik und Wissenschaft besteht, damit die Datengrundlagen für die Gesellschaft besser genutzt werden und dadurch die statistischen und wissenschaftlichen Instrumente und Programme in wechselseitiger Hilfe fortentwickelt werden können.
5. Um zu nachhaltigen Formen der Zusammenarbeit zu kommen, ist es ratsam, die vielfältigen Erfahrungen aus den bereits durchgeführten Pilotexperimenten mit der amtlichen Statistik und aus der internationalen Zusammenarbeit, auch bei der Finanzierung, zu sichten. Insbesondere die institutionellen Ausgestaltungen der letzten Jahre in den Niederlanden, Norwegen und Kanada sollten herangezogen werden, um entsprechende Lösungen zu finden, die auch den Grundsätzen des Datenschutzes gerecht werden.
6. Jede tragfähige Lösung muß einbeziehen, daß die politischen Gestaltungsaufgaben heute nicht mehr nur national wahrgenommen werden können. Der europäische und internationale Vergleich nicht nur der Daten, sondern auch des institutionellen Kontext bildet die notwendige Voraussetzung verantwortlicher Gestaltung. Deshalb dürfen Praktiken des Datenzugangs die notwendige international vergleichende Forschung nicht behindern.
7. Statistik und Wissenschaft brauchen einander, gerade in der internationalen Einordnung als Partner in einer solchen informationellen, unabhängigen Infrastruktur. Wis-

senschaft muß deshalb im Eigeninteresse für die Stärkung der wissenschaftlichen Analyse- und Methodenkapazität der amtlichen Statistik eintreten.

8. Die empirische Wirtschafts- und Sozialforschung und die amtliche Statistik können einander viel geben. Gegenseitige Studien- und Forschungsaufenthalte bieten dazu hervorragende Möglichkeiten. Die institutionellen und finanziellen Voraussetzungen sind zu schaffen. Modelle für eine systematische Weiterbildung der Datenproduzenten und Datennutzer sollten in Kooperation zwischen Universitäten, Forschungsinstituten und statistischen Ämtern erprobt werden.

Die zu diesen Thesen führenden Diskussionen des Symposiums „Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven“ am 31. Mai/1. Juni 1999 in Wiesbaden sollten von einer Kommission weiterverfolgt werden, die Lösungsvorschläge zu den inhaltlichen, institutionellen und ressourcenbezogenen Fragen erarbeitet.

Teilnehmer der Podiumsdiskussion „Perspektiven der Kooperation von Wissenschaft und amtlicher Statistik in Deutschland“ auf dem Symposium „Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven – am 31. Mai/1. Juni 1999 in Wiesbaden: *Johann Hahlen* (Statistisches Bundesamt, Wiesbaden), *Prof. Dr. Richard Hauser* (Universität Frankfurt/Main), *Prof. Dr. Karl Ulrich Mayer* (Moderation; Max-Planck-Institut für Bildungsforschung, Berlin), *Prof. Dr. Wolfgang Zapf* (Wissenschaftszentrum Berlin für Sozialforschung), *Prof. Dr. Klaus F. Zimmermann* (Forschungsinstitut zur Zukunft der Arbeit, Bonn), *Hansvolker Ziegler* (Bundesministerium für Bildung und Forschung, Bonn).

* * * * *

Mitteilung des Bundesministeriums für Bildung und Forschung zur Einsetzung der Kommission

Amtliche Statistik und Wissenschaft müssen als zentrale Teile der informationellen Infrastruktur stärker zusammenwirken - Kommission zur Verbesserung der informationellen Infrastruktur nimmt ihre Arbeit auf

Zur Verbesserung der Zusammenarbeit zwischen Wissenschaft und Statistik hat Frau Bundesministerin Edelgard Bulmahn eine Kommission berufen, in der Sozial- und Wirtschaftswissenschaftler mit Vertretern der amtlichen Statistik konkrete Empfehlungen ausarbeiten sollen.

Nur auf der Grundlage verlässlicher Daten können die Sozial- und Wirtschaftswissenschaften in Deutschland treffsichere Analysen erstellen und damit die Handlungsgrund-

lagen für die Politik verbessern. Über viele der neueren gesellschaftlichen und wirtschaftlichen Entwicklungen, wie die Flexibilisierung der Wirtschafts- und Arbeitsverhältnisse oder den Umfang unsteter Beschäftigung, wissen wir viel zu wenig, erst recht nicht über ihre Auswirkungen auf andere Lebensbereiche.

Dabei verfügen wir in Deutschland über eine im internationalen Vergleich sehr leistungsfähige amtliche Statistik. Auch die Sozial- und Wirtschaftswissenschaften weisen eine beachtliche, allerdings disziplinar aufgesplittete Kapazität in der empirischen Forschung auf. Was uns fehlt, ist ein systematischeres Zusammenwirken von Wissenschaft und Statistik bei der Erfüllung ihrer jeweiligen Aufgaben.

Wissenschaft und Statistik brauchen dabei einander. Die gegebenen rechtlichen Möglichkeiten der Zusammenarbeit müssen besser genutzt und in beständigere Organisationsformen übergeführt werden. Es kann nicht angehen, daß in Deutschland für viel Steuergeld Daten gesammelt, aber sie anschließend vor der Nutzung durch die ebenfalls mit Steuergeld finanzierten Wissenschaft weggeschlossen werden, die damit Problemlösungswissen für die Politik erarbeiten könnte. Zu klären sind deshalb Fragen des Datenzugangs für die Wissenschaft, einschließlich geeigneter wissenschaftlicher und organisatorischer Vorkehrungen dafür, dass dies mit den strengen deutschen und europäischen Datenschutzbestimmungen im Einklang steht.

Da jedoch das einzig Beständige in Wirtschaft und Gesellschaft der Wandel ist, muß auch die statistische Erfassung mit diesem Wandel Schritt halten. Deshalb braucht die amtliche Statistik im Gegenzug Rat und Unterstützung der Wissenschaft, damit ihre Datenerfassung stets dem Wandel folgt und ihn sogar rechtzeitig feststellen kann.

Die jetzt eingesetzte Kommission zur Verbesserung der informationellen Infrastruktur soll im Laufe des Jahres 2000 konkrete Ergebnisse zu den inhaltlichen, organisatorischen und ressourcenbezogenen Fragen einer verbesserten Zusammenarbeit zwischen Wissenschaft und amtlicher Statistik vorlegen. Die gut funktionierenden institutionellen Lösungen etwa in den Niederlanden, Norwegen oder auch Kanada zeigen den Weg und unterstreichen den Nachhol- und Handlungsbedarf in Deutschland. Dabei kann bereits auf der Erfahrung mit einigen Pilotprojekten der Zusammenarbeit aufgebaut werden, die insbesondere mit Unterstützung des BMBF in den letzten Jahren durchgeführt worden sind.

Mit der Berufung dieser Kommission setzt Bundesministerin Bulmahn die Ergebnisse eines Symposiums um, das zu dem Thema „Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven“ am 31. Mai und 1. Juni 1999 in Wiesbaden stattfand, und bei dem sich Vertreter der empirischen Sozial- und Wirtschaftsforschung,

der amtlichen Statistik, der Registerdaten und der Verwaltung über die Notwendigkeit eines engeren Zusammenwirkens einig waren.

Die Kommission wird von *Prof. Dr. Hans-Jürgen Krupp*, dem Präsidenten der Landeszentralbank in der Freien und Hansestadt Hamburg, in Mecklenburg-Vorpommern und in Schleswig-Holstein in Kooperation mit *Johann Hahlen*, dem Präsidenten des Statistischen Bundesamtes geleitet. Die weiteren Mitglieder sind:

- aus der amtlichen Statistik: Herr Präsident *Eckart Hohmann* vom Hessischen Statistisches Landesamt,
- aus dem Bereich der Registerdaten: Herr *Dr. Axel Reimann*, vom Verband Deutscher Rentenversicherungsträger (VDR),
- aus der Wirtschaftswissenschaft: Herr *Prof. Dr. Richard Hauser* von der Johann Wolfgang Goethe-Universität in Frankfurt am Main und Herr *Prof. Dr. Gert Wagner* vom Deutschen Institut für Wirtschaftsforschung (DIW) in Berlin und der Europa-Universität Viadrina in Frankfurt (Oder),
- aus der Statistikwissenschaft: Herr *Prof. Dr. Gerhard Armingier* von der Gesamthochschule Wuppertal,
- aus der Soziologie: Herr *Prof. Dr. Walter Müller* von der Universität Mannheim und Herr *Prof. Dr. Karl Ulrich Mayer* vom Max-Planck-Institut für Bildungsforschung Berlin und
- aus dem Ausland: Herr *Willem F.M. de Vries*, Stellvertretender Generaldirektor von Statistics Netherlands und künftiger stellvertretender Direktor des Statistikdepartements der Vereinten Nationen.

FÖRDERPREIS DES STATISTISCHEN BUNDESAMTES FÜR ZUMA-DISSERTATION

Heike Wirth, Soziologin am Zentrum für Umfragen, Methoden und Analysen, erhielt für ihre Doktorarbeit zum Thema „Bildung, Klassenlage und Partnerwahl. Eine empirische Analyse der bildungs- und klassenspezifischen Heiratsbeziehungen“ einen Förderpreis des Statistischen Bundesamtes. Der Förderpreis, der mit 2000 Euro dotiert ist, wurde im Rahmen des 8. Wissenschaftlichen Kolloquiums des Statistischen Bundesamtes und der Deutschen Statistischen Gesellschaft am 18. November in Wiesbaden verliehen.

1999 hat das Statistische Bundesamt erstmals den Gerhard-Fürst-Preis ausgeschrieben. Mit dem Preis, der nach dem ersten Präsidenten des Statistischen Bundesamtes benannt wurde, sollen herausragende Dissertationen und herausragende Diplom- bzw. Magisterarbeiten ausgezeichnet werden, die theoretische Themen mit einem engen Bezug zum Aufgabenspektrum der amtlichen Statistik behandeln oder die empirische Fragestellungen unter Nutzung von Daten der amtlichen Statistik untersuchen. Es kann sich dabei ebenso um Arbeiten aus der theoretischen Statistik oder aus der Wirtschafts- und Sozialstatistik handeln wie um wirtschaftswissenschaftliche oder um sozialwissenschaftliche Abhandlungen. Mit der Auslobung des Preises möchte das Statistische Bundesamt „die Beschäftigung mit Fragen der amtlichen Statistik im Rahmen der universitären Ausbildung fördern und die Zusammenarbeit zwischen Wissenschaft und amtlicher Statistik intensivieren“.

Auf Beschluß des Gutachtergremiums wurden 1999 zwei Förderpreise verliehen. Die Dissertation von Heike Wirth „Bildung, Klassenlage und Partnerwahl. Eine empirische Analyse zum Wandel der bildungs- und klassenspezifischen Heiratsbeziehungen“ wurde als überdurchschnittlich gute Arbeit bewertet, die die wissenschaftliche Diskussion im Bereich der amtlichen Statistik befruchtet hat. Der zweite Förderpreis ging an Yuanhua Feng, wissenschaftlicher Mitarbeiter an der Universität Konstanz, der für seine Dissertation „Kernel- and Locally Weighted Regression with Application to Time Series Decomposition“ ausgezeichnet wurde. Der Förderpreis wurde im Rahmen des 8. Wissenschaftlichen Kolloquiums des Statistischen Bundesamtes und der Deutschen Statistischen Gesellschaft am 18. November in Wiesbaden durch den Präsidenten des Statistischen Bundesamtes, Johann Hahlen verliehen.

Im Gutachtergremium sind folgende Wissenschaftler vertreten: Prof. Dr. *Hans Wolfgang Brachinger* (Seminar für Statistik, Universität Freiburg/Schweiz), Prof. Dr. *Ullrich Hei-*

lemann (Rheinisch-Westfälisches Institut für Wirtschaftsforschung (Essen), Prof. Dr. *Richard Hauser* (Institut für Volkswirtschaftslehre, Universität Frankfurt am Main), Prof. Dr. *Walter Müller* (Lehrstuhl für Methoden der empirischen Sozialforschung und angewandte Soziologie, Universität Mannheim), Prof. Dr. *Werner Neubauer* (Institut für Statistik und Mathematik, Universität Frankfurt am Main).

* * * * *

Bildung, Klassenlage und Partnerwahl. Eine empirische Analyse zum Wandel der bildungs- und klassenspezifischen Heiratbeziehungen.

Die Dissertation von Heike Wirth beschäftigt sich mit der Analyse von bildungs- und klassenspezifischen Heiratsbeziehungen. Die gesellschaftliche Entwicklung in Deutschland seit Ende des Zweiten Weltkrieges ist durch eine Vielzahl von Veränderungsprozessen gekennzeichnet. Die Auswirkungen dieser Modernisierungsprozesse auf die sozialen Strukturen in der Bundesrepublik werden kontrovers diskutiert. Grob skizziert sind hierbei zwei Hauptargumentationslinien zu unterscheiden. Die erste Argumentationsrichtung ist im wesentlichen durch die Leitthesen der 'Individualisierung' und 'Entstrukturierung' geprägt. Die Erweiterung von individuellen Handlungsoptionen infolge der Anhebung des materiellen Wohlstands wie auch allgemeiner Modernisierungsprozesse habe dazu geführt, so die These, daß traditionelle Ungleichheitskriterien immer weniger Einfluß auf individuelles Verhalten und die Strukturierung von sozialen Beziehungen ausüben. Es komme zu einem Verfall von traditionellen Strukturen und kollektiven Orientierungsmustern, während individualisierte Lebensbeziehungen und expressive Verhaltenselemente an Bedeutung gewinnen. Folglich sei es auch kaum mehr möglich, gesellschaftliche Phänomene empirisch mit den typischen Merkmalen sozialer Ungleichheit wie Bildung, Schicht- oder Klassenzugehörigkeit adäquat zu beschreiben oder gar theoretisch zu verorten. Die zweite Argumentationsrichtung hält dieser These entgegen, daß die Auswirkungen von objektiv weiterbestehenden Ungleichheiten auf die Gesellschaftsstruktur unterschätzt würden bzw. unberücksichtigt blieben. Diese Kritik beruht auf empirischen Untersuchungen, die belegen, daß Unterschiede in den Lebenschancen auch in der Gegenwartsgesellschaft in einem engen Zusammenhang mit den traditionellen Ungleichheitskriterien wie Bildungsniveau und Schicht- oder Klassenzugehörigkeit stehen.

Die Arbeit bezieht sich in ihrer theoretischen Verortung insofern auf diese Debatte, als der Frage nachgegangen wird, ob sich die vermutete Entstrukturierung auch in einem Bereich zeigt, dem für die Reproduktion sozialer Ungleichheit große Bedeutung zugeschrieben wird: der bildungs- und klassenspezifischen Partnerwahl. In der neueren Forschung wird die Durchlässigkeit gruppenspezifischer Heiratsbeziehungen - gemessen an dem Ausmaß,

in welchem zwischen sozialen Gruppen (*Heterogamie*) bzw. innerhalb sozialer Gruppen (*Homogamie*) geheiratet wird - neben beruflicher Mobilität als ein wesentlicher Indikator für die Offenheit einer Gesellschaft gesehen.

Durchgeführt wurde die Studie mit Daten der Volkszählung 1970 und dem Mikrozensus 1993, der es erlaubt, auch die bisher noch wenig erforschte Situation in der ehemaligen DDR in die Analyse einzubeziehen.

Im Kontext der Ausgangsfrage nach der Durchlässigkeit von sozialen Verkehrskreisen bildete die Frage nach den zentralen Entwicklungstendenzen der *bildungsspezifischen* Heiratsbeziehungen einen ersten Schwerpunkt der Arbeit. Der Wandel von bildungsspezifischen Heiratsbeziehungen ist eine Dimension für die Überprüfung der von der Individualisierungsthese angenommenen Öffnungstendenzen unserer Gesellschaft. Der Kern der Entstrukturierungsdiskussion konzentriert sich allerdings auf die infolge von fortgesetzter Modernisierung und Individualisierung vermutete nachlassende Prägekraft der Klassenzugehörigkeit auf individuelle Handlungsorientierungen. Folgt man Vertretern der Individualisierungsthese, dann kommt dem Klassenbegriff in der Gegenwartsgesellschaft kaum noch Wirklichkeitsgehalt zu, die soziale Wahrnehmbarkeit von Klassen bspw. im Sinne ihrer Abgrenzung durch „*Kontakt-, Hilfs- und Heiratskreise*“ ist nicht mehr gegeben.

Als zentraler Befund der verschiedenen bildungsspezifischen Analysen ist festzuhalten, daß die vorliegenden Ergebnisse insbesondere für Westdeutschland in eindeutiger Weise der These einer sozialen Annäherung (Entstrukturierungsthese) zwischen den Angehörigen verschiedener Bildungsgruppen widersprechen. Die Analyse der relativen Homogamieraten als maßgeblicher Indikator für Öffnungs- oder Schließungsprozesse zeigt, daß die Neigung, Partner mit dem gleichen formalen Bildungsabschluß zu heiraten, das die Heiratsbeziehungen dominierende Muster darstellt und sich hieran über die Kohorten hinweg kaum etwas verändert. Die stärkste Schließung findet sich bei der privilegiertesten Bildungsgruppe, den Akademikern, und der am wenigsten privilegierten Bildungsgruppe, den Hauptschulabsolventen ohne berufliche Ausbildung. Am oberen und unteren Ende der Bildungsskala ist demnach eine Kumulierung der jeweils hohen bzw. niedrigen sozio-kulturellen und ökonomischen Ressourcen von Männern und Frauen zu beobachten.

Insgesamt deuten die Ergebnisse in der Tendenz auf eine in der Gegenwart stärkere Polarisierung der bildungsspezifischen Heiratsbeziehungen in Westdeutschland hin. Die Bildungsexpansion ist nicht von einer Öffnung der bildungsspezifischen Heiratskreise begleitet, sondern die soziale Distanzierung zwischen den oberen und unteren Bildungsgruppen ist eher größer geworden. Im Zuge der allgemeinen Höherqualifizierung ist es zu keiner 'Verschmelzung' der mittleren und oberen Bildungsgruppen gekommen, d.h. es ist keine Öffnung der sozialen Heiratskreise zu beobachten. Gemessen an den bildungs-

spezifischen Heiratskreisen ist die westdeutsche Gesellschaft damit auch in der Gegenwart als eine 'Geschlossene Gesellschaft' zu sehen. Für Ostdeutschland ist festzuhalten, daß die allgemeine Strukturierung der Heiratsbeziehungen in ihren Grundzügen dem Muster in Westdeutschland ähnelt. Auch bei den ostdeutschen Kohorten ist die Neigung zur bildungshomogamen Partnerwahl das dominierende Charakteristikum, gefolgt von der Tendenz, daß beide Partner den gleichen allgemeinbildenden Ausbildungsabschluß aufweisen.

Der zweite Schwerpunkt der Arbeit ist die Analyse von klassenspezifischen Heiratsbeziehungen. Hier ist festzustellen, daß die klassenspezifischen Heiratsbeziehungen im wesentlichen ein ähnliches Muster der sozialen Durchlässigkeit aufweisen wie intergenerationale Mobilitätsprozesse. Charakteristisch hierfür ist zum einen die starke Konzentration der Heiratsbeziehungen auf Partner der eigenen Klassenzugehörigkeit. Diese tritt - ähnlich wie in Studien zur intergenerationalen Mobilität belegt - am deutlichsten bei den Dienstklassenangehörigen und den traditionellen Arbeiterklassen zutage. Gleichfalls weist die Strukturierung der Heiratsbeziehungen zwischen Angehörigen unterschiedlicher Klassen ein den Mobilitätsprozessen ähnliches Muster auf. Dieses ist zum einen dadurch gekennzeichnet, daß den Chancen von Angehörigen der Arbeiterklassen, in die Dienstklassen einzuheiraten, offensichtlich erhebliche Barrieren gesetzt sind. Zum anderen sind die Heiratsbeziehungen durch eine massive Blockbildung gekennzeichnet: Angehörige des 'white-collar' Blocks heiraten ebenso vorwiegend untereinander, wie dies Angehörige des 'blue-collar' Blocks tun. Entgegen der weitverbreiteten Vermutung, daß ein Durchbrechen dieser traditionellen 'Kragengrenze', wenn nicht in großem Maßstab, so doch hinsichtlich der Heiratsbeziehungen zwischen Facharbeitern und Frauen in Angestelltenpositionen zu erwarten ist, zeigt sich unter Kontrolle der Klassenverteilungen von Männern und Frauen keine erhöhte positive Affinität zwischen diesen beiden Gruppen. Auch wenn die Bedingtheit zwischen Bildungsniveau und Klassenposition kontrolliert wird, ist eine klare Klassendimension in den Heiratsbeziehungen zu beobachten. Diese äußert sich insbesondere in der weiterhin starken Separierung der Heiratskreise in einen manuellen und nicht-manuellen Block.

Die Modernisierungsprozesse zwischen den 70er und den 90er Jahren haben nicht zu einer grundsätzlich anderen Zusammensetzung der klassenspezifischen Heiratsbeziehungen geführt. Ein Verschmelzen der Heiratskreise von der Arbeiterschaft und den Angestellten bzw. das Entstehen einer diffusen Mitte läßt sich nicht erkennen. Selbst wenn es in Hinblick auf den materiellen Wohlstand sicherlich zu einer Überlappung zwischen Arbeitern und Angestellten gekommen ist und die individuellen Handlungsoptionen zugenommen haben, scheint sich dies nach den vorliegenden Ergebnissen kaum auf die sozialen Verkehrskreise dieser

Gruppen ausgewirkt zu haben. Die gesellschaftliche Entwicklung in den letzten Jahrzehnten hatte somit nahezu keinen Einfluß auf die klassenspezifische Partnerwahl.

Wie sind diese Befunde von - auch in der Gegenwart - stark durch die Bildungs- und Klassenzugehörigkeit geprägten Heiratsbeziehungen zu erklären, angesichts der Ergebnisse anderer Untersuchungen, die am Beispiel von Heiratsannoncen aufzeigen, daß der Bezug auf Merkmale der sozialen Position für die Partnersuche deutlich an Bedeutung verloren hat? In der modernen - auf gemeinsamen Interessen, gegenseitigem Verständnis und emotionaler Geborgenheit aufbauenden - Partnerschaft scheint der Bezug auf das eigene 'social standing' *„increasingly become illegitimate in the process of searching for a true love partner“*. Stattdessen scheint die Partnersuche heutzutage vor allem an Kriterien des Lebensstils, der äußerlichen Erscheinung und Freizeitaktivitäten orientiert. Die einfachste Erklärung für diesen scheinbar widersprüchlichen Befund ist, daß auch die Herausbildung von unterschiedlicher Lebensführung und unterschiedlichen Lebensstilen in einem engen Zusammenhang mit ökonomischen Ressourcen steht. So zeigen andere Untersuchungen eine klare Beziehung zwischen dem praktizierten Lebensstil und der sozio-ökonomischen Position. In diesem Kontext erübrigen sich somit explizite Hinweise auf die eigene soziale Position, da diese bereits implizit bei der Beschreibung der eigenen Interessen enthalten sind. Weiterhin ist zu berücksichtigen, daß bereits über die Wahl der Zeitung, in welcher inseriert wird, in der Regel eine erste soziale Selektion der potentiellen Partner vorgenommen wird. Schließlich ist zu berücksichtigen, daß der Weg der Partnersuche über Heiratsannoncen letztendlich bedeutet, daß innerhalb der eigenen sozialen Verkehrskreise kein akzeptabler Partner gefunden wurde und daher unter Umständen die Bereitschaft, Partner mit einer anderen sozialen Position zu akzeptieren größer ist, sofern sonstige Randbedingungen (z.B. gleiche Freizeitinteressen) gegeben sind.

NSDSTAT

EIN PROGRAMM ZUR SCHNELLEN DATENANALYSE

Ab sofort steht die Windows Version von NSDstat zur Verfügung. NSDstat ist ein Statistikanalyseprogramm, das für die Exploration und Standardanalyse von Datensätzen eingesetzt werden kann. Selbst große Datensätze, etwa ein ALLBUS Datensatz mit 3000 Fällen, können in Sekundenschnelle ausgezählt, tabuliert und in ihren Ergebnissen graphisch dargestellt werden. Vor allem an Geschwindigkeit übertrifft es die Standardstatistikpakete um ein Vielfaches. Wegen seiner einfachen Handhabung ist es auch für Kurse zu Statistik und Methoden der Empirischen Sozialforschung geeignet.

NSDstat ist ein Produkt des Norwegian Social Science Data Service in Bergen (NSD) und ist in den deutschsprachigen Ländern über ZUMA zu beziehen. Das Programm kann unter Windows 95/98 oder NT eingesetzt werden.

Eine besondere Eigenschaft von NSDstat ist die enge Verknüpfung aller numerischen Ausgaben (Tabellen) mit graphischen Darstellungen. Die Graphiken können gedruckt oder in andere Programme, z.B. in WinWord übernommen werden. Besonders gut ist die Möglichkeit zwischen numerischer Ausgabe und Graphik hin- und herzuschalten. Auch sind kartographische Ergebnisdarstellungen möglich.

Augenfällig ist die Konzentration auf das Wesentliche und der didaktische Impetus: Das Programm ist relativ klein, enthält wesentliche Instrumente für eine ausführliche Exploration neuer Datensätze und grundlegende statistische Prozeduren. Zu jeder statistischen Prozedur gibt es Informationen in Hilfsfenstern zu den statistischen Eigenschaften bzw. zu den verwendeten Formeln. Für den Unterricht weiterhin sehr hilfreich sind die Möglichkeiten der statistischen Deskription, wie z.B. das Einblenden einer Normalverteilung in die Graphik einer tatsächlichen Häufigkeitsverteilung und das parallele Unterlegen dieser Graphik mit Box-Whisker-Diagrammen.

Die Einfachheit der Handhabung zusammen mit den graphischen und statistischen Möglichkeiten machen deshalb NSDstat zu einem guten Instrument der Grundausbildung in Statistik und Datenanalyse.

Das Programm kann beim Zentrum für Umfragen, Methoden und Anlaysen (ZUMA), z. Hd. Cornelia Züll, Postfach 122155, 68072 Mannheim, Tel. 0621-1246-111, E-mail: nsdstat@zuma-mannheim.de bezogen werden. Eine Demo-Version steht zur Verfügung. Informationen zu NSDstat einschließlich der Demo-Version und eines Bestellformulars

sind auch auf den ZUMA www-Seiten zu finden: <http://www.zuma-mannheim.de/software/nsdstat/>.

Preise*

Akademische Nutzung	
Einzelkopie	240.-
5 Kopien	600.-
10 Kopien	900.-
25 Kopien	1500.-
50 Kopien	2100.-
Studentenversion	90.-**

Kommerzielle Nutzung	
Einzelkopie	300.-
5 Kopien	750.-
10 Kopien	1125.-
25 Kopien	1875.-
50 Kopien	2625.-

* Alle Preise sind in DM angegeben. Die Preise sind Netto-Preise, d.h. sie werden zzgl. Mehrwertsteuer in Rechnung gestellt.

** Zum Erhalt einer Studentenversion ist eine Immatrikulationsbescheinigung erforderlich.

TESTDATEN ZUM IAB-BETRIEBSPANEL ERHÄLTlich

Seit einem halben Jahr ist am Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) in Nürnberg eine Schalterstelle beim IAB-Betriebspanel eingerichtet, die es der wissenschaftlichen Forschung ermöglichen soll, die darin enthaltenen Daten zu nutzen. Ziel der Position ist es einerseits Informationen über den Datensatz bereitzustellen bzw. deren Nutzungsmöglichkeiten zu beschreiben, andererseits Auswertungswünsche von externen Forschern zu bearbeiten. Bislang konnten die Datennutzer ihre Auswertungswünsche nur anhand eines Codebooks formulieren. Nun stehen den Forschern außerdem „Testdaten“ zur Verfügung, die die Struktur der originalen Angaben exakt wiedergeben und somit die Programmierung von Auswertungswünschen ermöglichen.

Das IAB-Betriebspanel ist eine seit 1993 jährlich wiederholte Panelbefragung von Betriebseinheiten. Seit 1996 enthält der Datensatz auch Betriebe aus den neuen Bundesländern und Ost-Berlin. Grundgesamtheit der repräsentativen Stichprobe sind alle Betriebe mit mindestens einem sozialversicherungspflichtigen Beschäftigten. Die Ziehung der Zufallsstichprobe erfolgte auf Grundlage einer Betriebsdatei, die aus der Beschäftigtenstatistik der Bundesanstalt für Arbeit generiert wurde. Die Samples enthalten pro Befragungswelle jeweils über 4000 Fälle für Ost- und Westdeutschland. Die sechste Welle West bzw. die dritte Welle Ost aus 1998 enthält insgesamt 9194 Betriebe, wobei sich 4289 Fälle auf die neuen Bundesländer und Ost-Berlin und 4905 Fälle auf das frühere Bundesgebiet beziehen. Die Rücklaufquote der mündlichen Befragung beträgt außer für die erste Welle Ost immer über 70 Prozent und lag 1998 insgesamt bei 73 Prozent. Da in den neuen Bundesländern und Berlin eine höhere Stichprobe gezogen wurde, ist es dort nicht nur möglich, Auswertungen für die gesamte Bundesrepublik Deutschland oder die beiden Landesteile, sondern auch separat für die einzelnen Länder durchzuführen. Der Fragenkatalog umfaßt alle Bereiche des betrieblichen Handelns. Erfäßt werden außerdem alle Branchen und Betriebsgrößenklassen.²

² Weitere Informationen zum IAB-Betriebspanel finden sich u. a. bei: Bellmann, L. (1997): Das Betriebspanel des IAB. S. 169 - 182 in: Hujer, R./Rendtel, U./Wagner, G. (Hrsg.): Wirtschafts- und Sozialwissenschaftliche Panel-Studien. Sonderheft zum Allgemeinen Statistischen Archiv 30, Göttingen.

Aus datenschutzrechtlichen Gründen ist es leider nicht möglich, die bisher erhobenen Daten, Wissenschaftlern außerhalb des IAB zur Verfügung zu stellen. Um aber trotzdem die Arbeit mit dem IAB-Betriebspanel zu ermöglichen, ist eine sogenannte Schalterstelle beim IAB eingerichtet worden, die Auswertungswünsche entgegen nimmt und die Ergebnisse an die externen Interessenten weiterleitet. Durch die Testdatensätze ist es nun möglich, eigene Programme zur Auswertung zu erstellen. Diese werden dann innerhalb des IAB mit den Originaldaten durchgeführt. Nach der obligatorischen datenschutzrechtlichen Überprüfung erfolgt die Übermittlung der Resultate.

Die Testdaten sind bei der Schalterstelle kostenlos erhältlich. Sie bestehen aus 7 SPSS 9.0 Dateien (pro Welle ein Datensatz plus ein Organisationsfile) und einer Textdatei, die den Einstieg in die Arbeit mit den Daten ermöglicht und unbedingt vor Beginn gelesen werden sollte. Die Datensätze sind so konstruiert, dass sie die Struktur der Originaldaten abbilden. Ebenso sind die gesamten Metainformationen des IAB-Betriebspanels in den Files enthalten. Auswertungen sind z. Z. für SPSS 9.0, STATA 6.0, TSP 4.3 und GAUSS 3.2.38 möglich.

Weitere Informationen über die Arbeit mit dem IAB-Betriebspanel und den Zugang zur Schalterstelle können sie aus dem Codebook zum IAB-Betriebspanel entnehmen oder direkt bei der Schalterstelle bekommen. Die Schalterstelle beim IAB-Betriebspanel ist unter der folgenden Adresse zu erreichen:

*Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit
z. Hd. Herrn Dr. Arnd Kölling
Regensburger Str. 104; 90327 Nürnberg
Tel.: 0911 / 179 – 3174; Fax: 0911 / 179 – 3297
E-mail: arnd.koelling@iab.de*

BUCHBESPRECHUNGEN

Sowohl die akademische als auch die angewandte Forschung der vergangenen Jahre zum Thema „Lebensstile“ vermittelt ein gespaltenes Bild: Zum einen ist dem Konzept des Lebensstils mit großer Hoffnung und Faszination begegnet worden, zum anderen ist die Bilanz der Leistungsfähigkeit dieses Konzeptes für sozialwissenschaftliche Analysen eher ernüchternd. Vor diesem Hintergrund diagnostiziert Peter H. Hartmann einen „gewissen Rückgang der Lebensstil-Euphorie“ und bemerkt treffend, daß „sich ein Forschungsfeld, dessen prominenteste Vertreter mit weitgehenden Behauptungen, Wahrheitsansprüchen und Gesellschaftsdiagnosen angetreten waren, vor der theoretischen Systematisierung und empirischen Prüfung seiner Aussagen gewissermaßen klammheimlich zu verabschieden“ droht (S. 12). Solch eine Systematisierung und Prüfung zentraler Aussagen der Lebensstilsoziologie nimmt Hartmann in seinem Buch in Angriff. Ihm geht es dabei um eine Darstellung, Kritik und Weiterentwicklung der Lebensstilforschung.

PETER H. HARTMANN

Lebensstilforschung. Darstellung, Kritik und Weiterentwicklung

Leske + Budrich, Opladen, 1999

308 Seiten, 54 Mark, ISBN 3-8100-2514-3

Im Anschluß an das Einleitungskapitel steht in den Kapiteln 2 bis 4 die Darstellung im Vordergrund. Hartmann hebt sich dabei in erfreulicher Weise von den überblickshaften Darstellungen des Forschungsstandes ab, wie sie in den letzten Jahren verschiedentlich vorgelegt worden sind. Dies wird bereits in Kapitel 2 deutlich, in dem er sich um eine Systematisierung definitorischer Aspekte des Lebensstils bemüht. Über die Heranziehung kulturtheoretischer Literatur gelangt er zu drei (Lebens-) Stildefinitionen. Nach der Ausdrucksdefinition sei es konstitutiv für einen Lebensstil, dass er im sozialen Handeln Sinn oder Bedeutungsinhalte ausdrücke. Nach der Formdefinition werde ein Lebensstil durch ähnliche formale Merkmale bestimmt. Nach der Identifizierbarkeitsdefinition schließlich stehe die äußere, symbolisch vermittelte Erkennbarkeit von Menschen als Repräsentanten eines Lebensstils im Vordergrund. Die Unterscheidung dieser Definitionen ist Hartmann zufolge deshalb bedeutsam, weil in operationaler Hinsicht die Verwendung einer Ausdrucksdefinition für die Erhebung subjektiver Merkmale, die einer Form- oder vor allem Identifizierbarkeitsdefinition für die Erhebung objektiver Merkmale des Lebensstils spreche (S. 30). Auch die Frage, inwieweit Stile als bewußt gewählt oder als unbewußt wirkend angesehen werden könnten, stehe mit diesen grundlegenden Definitionen in Verbin-

dung. Erstere Position sei eher mit einer Ausdrucksdefinition, letztere eher mit einer Formdefinition des Lebensstils vereinbar.

In Kapitel 3 rezipiert Hartmann Lebensstilansätze in der Marktforschung weit intensiver, als dies bislang in der soziologischen Lebensstiltliteratur geschehen ist. Er stellt solche Lebensstiltypologien vor, die von den jeweiligen Instituten vergleichsweise gut dokumentiert wurden und sich vergleichsweise breiter Anwendung erfreuen. Dies sind der AIO-Ansatz (Conrad & Burnett), der VALS-Ansatz (Mitchell), die Socio- bzw. Euro-Styles (Cathelat bzw. GfK) und der Sinus-Ansatz sowie die jeweiligen Weiterentwicklungen dieser Ansätze. Soweit dies möglich ist, versucht der Autor das methodische Vorgehen der Institute zu eruieren. Allerdings stößt er dabei auf Grenzen, da die kommerziellen Institute in der Regel keine Auskunft über die statistischen Algorithmen der Zuweisung von Personen zu Lebensstiltypen geben und oftmals keine Gütekriterien für ihre Typologien diskutieren. Insofern geht Hartmanns Botschaft dahin, kommerziell entwickelten Typologien – wie den in der Soziologie sehr beliebten Sinus-Milieus – mit größerer Skepsis gegenüberzustehen.

Im vierten Kapitel wendet sich der Verfasser Lebensstilansätzen in der akademischen Sozialforschung zu. Hartmann kommt hier das Verdienst zu, über den Tellerrand der deutschen Lebensstilforschung hinauszublicken und auch Befunde vor allem aus der amerikanischen (z.B. DiMaggio, Peterson) und niederländischen (z.B. deGraaf, Ganzeboom) Forschung einzubeziehen, in der in stärkerem Maß einzelne Lebensstilelemente als Variablen untersucht worden sind, während man sich in Deutschland eher auf die Entwicklung von Typologien konzentriert hat. Der Autor verfolgt hierbei weniger eine ausgreifende Darstellung der verschiedenen Lebensstilansätze, sondern stellt einzelne Ansätze – vor allem die von Bourdieu, Lüdtke, Schulze und Spellerberg – mit der Intention vor, empirisch testbare Hypothesen zu explizieren, die zum Teil im empirischen Teil des Buches überprüft werden. Ein „Neuling“ der Lebensstilsoziologie wird an dieser Stelle möglicherweise zusätzlich die Originalliteratur oder andere Überblicksdarstellungen konsultieren müssen. Andererseits bietet Hartmanns Vorgehen dem Leser den Vorteil, zentrale Hypothesen zur Ausgestaltung und Wirkung von Lebensstilen gebündelt vorzufinden. Beispielsweise diskutiert der Verfasser die Bedeutung kulturellen Kapitals für die Reproduktion sozialer Ungleichheit und bewertet sie im Licht empirischer Befunde (Kap. 4.1), referiert Hypothesen zur Trägheit von Lebensstilen (Kap. 4.3.1) und greift die Diskussion auf, inwieweit soziale Statusgruppen jeweils exklusive Formen des Kulturkonsums betreiben oder aber höhere Statusgruppen als „Allesfresser“ auch Kulturformen niedriger Schichten für sich vereinnahmen (Kap. 4.4.2).

In Kapitel 5 wendet sich Hartmann einem zentralen Problem der Lebensstilforschung zu: dem Vergleich verschiedener empirischer Typologien. Er identifiziert dabei vier Möglichkeiten der Meta-Analyse von Lebensstilstudien: Vergleiche seien erstens möglich nach Etikett und Beschreibung der Typen, zweitens nach den Facetten und Lebensbereichen, die in die Typenbildung einfließen, drittens nach allgemeinen Dimensionen des Lebensstils und viertens nach dem Verwertungszusammenhang der Typologien. In der Literatur wird – wenn überhaupt – der erste Zugang gewählt. Hartmann zeigt, daß solche Typenvergleiche nach Etikett und Beschreibung mit erheblichen Inkonsistenzen behaftet sind, wenn verschiedene Forscher dieselben Typologien vergleichen. Wenn also in der soziologischen Literatur behauptet werde, verschiedene Studien kämen zu „ähnlichen“ Lebensstiltypen, so mangle es diesen Aussagen an empirischem Gehalt (S. 144). Hartmann selbst führt eine Meta-Analyse nach Facetten und Lebensbereichen durch (Kap. 5.1.5). Mit Hilfe von Korrespondenzanalysen gelingt es ihm, einen Raum der Forscher und der Lebensbereiche, aus denen sie vorzugsweise ihre Lebensstilitems generieren, abzubilden. Der Stellenwert dieser Befunde für die Lebensstilforschung bleibt allerdings etwas unklar. Um von diesen Ergebnissen auf eine „Forschermilieuzugehörigkeit“ (S. 157) zu schließen, müssten wohl im Sinne von Bourdieus Studien zum wissenschaftlichen Feld in Frankreich weitere Variablen über das Forschermilieu herangezogen werden.

Sehr stichhaltig ist dagegen die sich in Kapitel 5.2 anschließende Kritik der empirischen Lebensstilforschung. Anhand einer Reihe von Gütekriterien macht Hartmann auf grundlegende Schwächen der empirischen Praxis der Lebensstilsoziologie aufmerksam. Die Kritikpunkte betreffen die unbegründete (aber seiner Meinung nach auch kaum begründbare) Auswahl von Lebensstilvariablen, die teils unpräzise Beschreibung der Dimensionen einzelner Lebensstiltypologien, das Problem der grafischen Abbildung der Typen, die Gefahr von Reifikationen aufgrund der idealisierten Typbezeichnungen sowie eine unzureichende Untersuchung der Kriteriumsvalidität und der Dynamik von Lebensstilen.

Den letzten beiden Punkten widmet der Verfasser in Kapitel 6 eine eigene Pilotstudie, die 1995 mit 129 Befragten in Köln durchgeführt wurde. Er untersucht die statistische Erklärungskraft von Lebensstilen sowie die Dynamik der Alltagsästhetik im Lebenslauf. Dazu versucht Hartmann zunächst alltagsästhetische Schemata und Lebensstiltypen in Anlehnung an die Modelle von Gerhard Schulze zu konstruieren. Die Reproduktion der drei Schemata der Alltagsästhetik (Hochkultur-, Trivial- und Spannungsschema) gelingt ihm recht gut. Schwieriger gestaltet sich dagegen die Reproduktion der fünf sozialen Milieus mittels Clusteranalysen: Ein dem Schulze'schen Integrationsmilieu verwandter Typus der Alltagsästhetik läßt sich nicht identifizieren, während weitgehende Entsprechungen der übrigen vier Milieus aufzufinden sind. Ob dies durch eine Variation von Raum und Zeit

der Studien oder durch grundlegende Probleme clusteranalytisch orientierter Typenbildung zu erklären ist, bleibt offen. Seine weiteren Analysen gründet Hartmann deshalb vorwiegend auf die drei alltagsästhetischen Schemata. Er untersucht deren Erklärungskraft und diejenige verschiedener soziodemographischer Variablen im Hinblick auf 148 abhängige Variablen, die vorwiegend den Bereichen Freizeit, Geschmack und Mediennutzung entstammen. Die alltagsästhetischen Schemata weisen zwar eine Erklärungskraft auf, die diejenige sozioökonomischer Variablen übersteigt, die aber in vergleichbarem Ausmaß auch durch die Variablen „Alter“ und „Schulbildung“ erzielt wird. Dies sei „ein für die Anhänger der These von der Segmentierung der Gesellschaft nach Lebensstilen enttäuschendes Ergebnis.“ (S. 238) Auch wenn der erzielte Befund angesichts der hohen Korreliertheit der Alltagsästhetik mit Alter und Bildung nicht ausgesprochen überraschend erscheint, unterwirft er die Lebensstilforschung – oder zumindest die Schulze’schen Segmentationsüberlegungen – doch einem gewissen Legitimationszwang. Hartmanns Ergebnisse weisen jedenfalls in bislang noch nicht vorgelegter Breite darauf hin, daß die statistische Erklärungsleistung allein möglicherweise kein Argument für die Verwendung von Lebensstilansätzen in der Soziologie sein kann.

Durch die Berücksichtigung von Retrospektivfragen in seiner Umfrage gelingt es Hartmann weiterhin, erstmals in Deutschland eine Kohortenanalyse für die individuelle Stabilität alltagsästhetischen Geschmacks im Lebenslauf durchzuführen und dabei Kohorten- und Alterseffekte zu trennen. Er kann zeigen, daß die Neigung zum Hochkulturschema mit zunehmendem Lebensalter deutlich zunimmt, die zum Spannungsschema geringfügig abnimmt. Daneben ist die Präferenz einer spannungsgeladenen Alltagsästhetik kennzeichnend für jüngere Kohorten. Die Neigung zum Trivialschema folgt dagegen allein einem Kohorteneffekt. Es darf demzufolge darüber spekuliert werden, ob volkstümliche Kulturformen in Zukunft nur noch ein Schattendasein fristen werden. Die Hochkultur hingegen kann offenbar auf die „innere Reifung“ der Popkulturgenerationen hoffen.

Auch wenn Hartmann die empirische Pilotstudie als seinen eigentlichen Beitrag zur Weiterentwicklung der Lebensstilforschung ansieht, enthält seine Arbeit darüber hinaus eine Reihe weiterer sehr hilfreicher Ansatzpunkte für die zukünftige Lebensstilforschung. Die von ihm herausgearbeiteten Definitionsansätze des Lebensstils sollten von Lebensstilforschern genauso aufgegriffen werden wie seine Diskussion von Vergleichsdimensionen und Evaluationskriterien von Lebensstiltypologien. Dazu zählt auch die Forderung nach einer Untersuchung lebensstilspezifischen Gruppenbewußtseins und typenspezifischer Interaktionsmuster (Kap. 5.2.2.5) sowie einer Entwicklung themenspezifischer Lebensstiltypologien (S. 239f). Auch in Hartmanns Arbeit bleiben jedoch zentrale Fragen der Lebensstilforschung unbeantwortet, etwa die einer Identifikation heutzutage dominanter

Lebensstildimensionen, die aus einem dimensional Vergleich empirischer Lebensstilypologien resultieren könnten und anhand derer stärker theoretisch fundierte Lebensstilypologien konstruiert werden könnten. Wünschenswert wäre auch eine eingehendere Auseinandersetzung mit methodischen Problemen der Typenkonstruktion (wie der Festlegung von Typengrenzen) gewesen, die etwa die Verwendung cluster- und korrespondenzanalytischer Verfahren aufwirft. Gleichwohl kann das besprochene Buch als eine der erfreulichsten Publikationen auf dem Gebiet der Lebensstilforschung betrachtet werden.

GUNNAR OTTE (gotte@sowi.uni-mannheim.de)

* * * * *

Das Buch ist als Einführung in die Untersuchung von sozialen und wirtschaftlichen Prozessen mit Hilfe von Computer-Simulationen konzipiert. Es soll die Leser in die Lage versetzen, die einschlägige Literatur zu verstehen und eigene Simulationen durchzuführen.

GILBERT, NIGEL/TROITZSCH, KLAUS G.

Simulation for the Social Scientist

Buckingham/Philadelphia: Open University Press 1999

273 Seiten, Paperb. \$15.99, ISBN 0-335-19744-2

Das erste Kapitel verortet die Computer-Simulation im sozialwissenschaftlichen Forschungsprozeß. Dabei wird klar herausgearbeitet, daß Simulationen theoriegeleitet sein müssen, um Sinn zu machen, und daß - zumindest beim derzeitigen Stand der Theorieentwicklung in den Sozialwissenschaften - der Nutzen von Simulationen eher in der Weiterentwicklung von Theorien als in der Prognose von Ereignissen liegt. Das zweite Kapitel stellt die Grundzüge der Simulations-Methode da und listet die Schritte auf, die üblicherweise jeweils durchlaufen werden müssen.

Die folgenden Kapitel sind jeweils einem Typ von Simulationsmodell gewidmet: system dynamics and world models, microanalytic simulation models, queuing models, multilevel simulation models, cellular automata, multi-agent models, learning and evolutionary models. Dabei werden jeweils die Grundzüge der Methode und ihr Anwendungsbereich (insbesondere auch im Vergleich zu den anderen Simulations-Methoden), geeignete Software, die Modellspezifikation, die Durchführung der Simulation selbst und die Interpretation der Ergebnisse dargestellt. Beispiele aus der Literatur für die Verwendung der jeweiligen Methoden und eine Vielzahl von gut gegliederten Verweisen auf die Literatur sowie insbesondere nützliche Internet-Adressen (teilweise im Anhang) runden die Kapitel ab.

Zwar hat die Verwendung von Computer-Simulationen in den Sozialwissenschaften in den letzten Jahrzehnten stark zugenommen. Dennoch ist diese Methode zur Zeit noch alles andere als ein Standardverfahren, was sich etwa auch an den äußerst seltenen Lehrveranstaltungen zu diesem Gebiet zeigt. Das Buch ist so interessant und gut geschrieben, daß ich ihm einen möglichst breiten Leserkreis wünsche. Nicht nur diejenigen, die die Durchführung eigener Simulationen beabsichtigen, werden das Buch mit Gewinn lesen können, sondern alle, die an der Theoriebildung in den Sozialwissenschaften interessiert sind.

MICHAEL BRAUN (braun@zuma-mannheim.de)