

GESIS

ZUMA

49

November 2001

NACHRICHTEN

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) (www.wgl.de).

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 luettinger@zuma-mannheim.de
Internet www.gesis.org/zuma ZUMA
www.gesis.org/ GESIS

Redaktion: Dr. Paul Lüttinger

ISSN 0941-1670 25. 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: http://www.gesis.org/publikationen/zeitschriften/zuma_nachrichten/.

Druck: Verlag Pfälzische Post GmbH, Neustadt/Weinstraße.
Gedruckt auf chlorfrei gebleichtem Papier.

© ZUMA

In eigener Sache5

FORSCHUNGSBERICHTE

Effekte des Erhebungsverfahrens? Ein Vergleich zwischen einer Web-basierten und einer schriftlichen Befragung zum ISSP-Modul Umwelt.
Wolfgang Bandilla, Michael Bosnjak & Patrick Altdorfer..... 7

Reliabilität von retrospektiven Berufsverlaufdaten. Ein Vergleich zwischen der Privatwirtschaft und dem öffentlichen Dienst anhand von Paneldaten.
Rolf Becker.....29

Der Einfluss der Expertise auf das Antwortverhalten in Delphi-Studien: Ein Hypothesentest.
Knut Blind & Kerstin Cubls.....57

Konsequenzen für die Methodenausbildung aus dem Gutachten der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI).
Walter Müller..... 81

MITTEILUNGEN

The Swiss Household Panel 1999-2003: Data for Research on micro-social change.
M. Budowski, R. Tillmann, E. Zimmermann, B. Wernli, A. Scherpenzeel & A. Gabadinho... 100

Die GESIS im Urteil der Profession: Ergebnisse einer Befragung von Soziologie-professorinnen und -professoren in Deutschland, Österreich und der Schweiz.
Gisbert Binder, Markus Klein, Rolf Porst & Matthias Stahl..... 126

Preis der Fritz Thyssen Stiftung an Bernhard Schimpl-Neimanns (ZUMA)133

Buchbesprechungen

S. KÜHNEL/D. KREBS: Statistik für die Sozialwissenschaften. Grundlagen, Methoden, Anwendungen.....	135
G. ROHWER/U. PÖTTER: Grundzüge der sozialwissenschaftlichen Statistik.....	137
R. OSTERMANN/K. WOLF-OSTERMANN: Statistik für Studierende der Sozialarbeit und Sozialpädagogik.....	141

PUBLIKATIONEN

ZUMA-Arbeitsberichte.....	145
ZUMA-Methodenberichte.....	146
HOW-TO-Reihe.....	147
Buchhinweise.....	148

ZUMA-WORKSHOPS, ERSTES HALBJAHR 2002

Stichproben in Theorie und Praxis, 26. Februar.....	149
Questionnaire Translation, 11. - 12. März.....	150
Eine Einführung in die Analyse linearer Strukturgleichungsmodelle mit LISREL 8.50, 6. - 7. Juni.....	151
Umfragen in den Massenmedien: Ein Grundkurs für Journalisten, 14. Juni.....	152
Das Internet als Setting für experimentelle Forschung in den Sozialwissenschaften, 20. - 21. Juni.....	152
Eine Einführung in Datenmanagement und Datenrecodierung, 26. - 27. Juni.....	153
Fragen verstehen und beantworten: Kognitive und kommunikative Grundlagen von Befragungen, 16. - 17. Juli.....	154
Simulation for Social Science, 23. - 27. September.....	154

DURCHWAHL-RUFNUMMERN

IN EIGENER SACHE

Das Internet gewinnt als Instrument der Datenerhebung in der empirischen Sozialforschung zunehmend an Bedeutung. *WOLFGANG BANDILLA*, *MICHAEL BOSNJAK* und *PATRICK ALTDORFER* berichten im ersten Beitrag der vorliegenden Ausgabe der ZUMA-Nachrichten über die Ergebnisse einer Befragung zum Thema Umwelt, die sowohl traditionell (in diesem Fall schriftlich), als auch WEB-basiert durchgeführt wurde. Im direkten Vergleich der beiden Verfahren zeigen sich zunächst deutliche Unterschiede in den Ergebnissen, und dies sowohl bei den inhaltlichen als auch bei den soziodemografischen Variablen. Nachdem jedoch die Stichprobe der Internetnutzer an die bevölkerungsrepräsentative Stichprobe auf der Basis soziodemografischer Variablen für eine ausgewählte Gruppe von Personen mit hoher Bildung angepaßt wurde, stimmten die Ergebnisse der WEB Befragung mit denen der schriftlichen Variante weitgehend überein.

Ein häufig eingesetztes Verfahren in der Umfrageforschung ist die Erhebung retrospektiver Angaben. Der Beitrag von *ROLF BECKER* behandelt die Zuverlässigkeit retrospektiver Angaben von Befragten zu beruflichen Tätigkeiten in ihrem Berufsverlauf, wobei er insbesondere untersucht, ob Beschäftigte des öffentlichen Dienstes aufgrund der institutionellen Besonderheiten ihrer Beschäftigung weniger zuverlässige Angaben machen als andere Beschäftigte.

Mit einem spezifischen Problem von Delphi-Studien beschäftigen sich *KNUT BLIND* & *KERSTIN CUHLS*. Basierend auf den Ergebnissen einer Delphi-Studie zur globalen Entwicklung von Wissenschaft und Technik untersuchen sie, in welchem Maß das Expertenwissen der Beteiligten das Antwortverhalten beeinflusst.

Die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik, die von der Bundesministerin für Bildung und Forschung im Herbst 1999 berufen wurde, um Lösungsvorschläge zu den inhaltlichen, organisatorischen und ressourcenbezogenen Fragen der Zusammenarbeit zwischen Wissenschaft und Statistik zu erarbeiten, hat im Frühjahr 2001 ein umfangreiches Gutachten vorgelegt. *WALTER MÜLLER*, Mitglied der Kommission, beschreibt die Arbeiten der Kommission und deren Empfehlungen und skizziert einige Folgerungen für die Methodenausbildung an den Universitäten.

1999 wurden ca. 8.000 Personen in über fünftausend Schweizer Haushalten über ihre Lebensbedingungen befragt. Während der folgenden zehn bis 15 Jahre soll diese Befragung jährlich durchgeführt werden. Das Schweizer Haushalts-Panel (SHP) bietet

somit eine hervorragende Datenbasis über die Schweizer Bevölkerung und deckt ein breites Spektrum von Themen und sozialwissenschaftlichen Forschungsansätzen ab. MONIKA BUDOWSKI und ihre Kollegen beschreiben den Datensatz und die Durchführung der bisherigen Wellen.

Im folgenden geht es um einen Beitrag in eigener Sache, und zwar um die Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS), die sich ja aus den drei Instituten Informationszentrum Sozialwissenschaften, Zentralarchiv für empirische Sozialforschung und ZUMA zusammensetzt. Um mehr über die Nutzung sowie die Bewertung der Produkte und Dienstleistungen der GESIS zu erfassen, hatt diese eine Befragung unter deutschsprachigen Soziologieprofessorinnen und -professoren in Deutschland, Österreich und der Schweiz durchgeführt, deren Ergebnisse *GISBERT BINDER, MARKUS KLEIN, ROLF PORST & MATTHIAS STAHL* zusammenfassen. Allen Kolleginnen und Kollegen, die an der Umfrage teilgenommen haben, möchte ich an dieser Stelle ausdrücklich danken.

Alljährlich wird der Preis der Fritz Thyssen Stiftung für die besten sozialwissenschaftlichen Aufsätze in deutscher Sprache vergeben. Dies ist der einzige Zeitschriftenpreis in den Sozialwissenschaften außerhalb des englischsprachigen Bereichs. Um so mehr freue ich mich, daß in diesem Jahr einem ZUMA-Mitarbeiter der zweite Preis verliehen wurde. Bernhard Schimpl-Neimanns wurde für seinen Aufsatz „Soziale Herkunft und Bildungsbeteiligung. Empirische Analysen zu herkunftsspezifischen Bildungsungleichheiten zwischen 1950 und 1989“, der in der Kölner Zeitschrift für Soziologie und Sozialpsychologie erschienen ist, ausgezeichnet.

Personalia gibt es folgende zu berichten: PD DR. MICHAEL HÄDER hat zum 1. Dezember einen Ruf auf eine Professur am Lehrstuhl für Methoden der empirischen Sozialforschung am Institut für Soziologie der Technischen Universität Dresden angenommen. Wir gratulieren Michael Häder herzlich zu dieser Berufung. SABINE KLEIN ist neue Mitarbeiterin beim International Social Survey Programme (ISSP). ZUMA verlassen haben ANKE SCHÖB und ANINA MISCHAU; FERDA KOYMEN wird ihre Tätigkeit zum 31. Dezember bei ZUMA beenden.

Ich wünsche allen Lesern ein friedvolles Weihnachtsfest und viel Glück und Erfolg im kommenden Jahr.

Max Kaase
Vorsitzender des ZUMA e.V.

EFFEKTE DES ERHEBUNGSVERFAHRENS? EIN VERGLEICH ZWISCHEN EINER WEB- BASIERTEN UND EINER SCHRIFTLICHEN BEFRAGUNG ZUM ISSP-MODUL UMWELT.

WOLFGANG BANDILLA, MICHAEL BOSNJAK & PATRICK ALTDORFER¹

Eine inhaltlich identische Befragung (Umwelteinstellungsmodul des International Social Survey Programme, ISSP) wurde sowohl im Rahmen einer bevölkerungsrepräsentativen schriftlichen Befragung als auch in Form einer Web-basierten Version mit Hilfe eines für die Internetnutzerschaft repräsentativen Online-Access-Panels durchgeführt. Im direkten Gruppenvergleich über beide Erhebungsmodi zeigten sich zunächst deutliche Unterschiede über inhaltliche wie auch soziodemografische Variablen. Eine Anpassungsgewichtung der Stichprobe der Internetnutzer an die bevölkerungsrepräsentative Stichprobe auf der Basis soziodemografischer Variablen ergab für den Vergleich über inhaltliche Items ein inkonsistentes Bild. Beschränkt sich dagegen der Vergleich unter beiden Erhebungsmodi auf solche Personen, die über einen ähnlichen Bildungshintergrund verfügen, so sind weder über die Itemmittelwerte noch über ausgewählte Item-Interkorrelationen praktisch bedeutsame Unterschiede ermittelbar. D.h. die Online erhobenen Daten stimmen für eine ausgewählte High-Coverage Gruppe (hier: Personen mit hoher Bildung) weitgehend überein mit denen der schriftlichen Variante.

The module 'environmental attitudes and values' of the International Social Survey Programme (ISSP) was administered both as a paper-and-pencil questionnaire to a representative sample of the German population and as a Web-based survey to an online access panel representative for German Internet users. These

¹ Wolfgang Bandilla und Michael Bosnjak sind Mitglieder der Arbeitsgruppe 'Computer-gestützte Datenerhebung' am ZUMA, Patrick Altdorfer ist Wissenschaftlicher Mitarbeiter beim forsa-Institut (Berlin). Wir danken der Rogator AG (Nürnberg) und dem forsa-Institut für die Unterstützung bei der Realisierung des Web Surveys.

two samples differ significantly with regard to sociodemographic and substantive variables. An attempt to weight the data of the Web-based sample on the basis of distribution characteristics of several socio-demographic variables resulted in rather inconsistent findings for the comparison of substantive items. However, if only respondents with similar education levels are compared, neither relevant differences in item means nor differences in selected inter-item correlations are observable. This means that for high coverage groups (e.g. subjects with a high education level), the data gathered via the Web are basically identical to those obtained in a traditional self-administered mode.

1. Problemstellung und Zielsetzung der Studie

Zunehmend gewinnen Techniken an Bedeutung, die das Internet zum Zweck der Datenerhebung nutzen. Belege hierfür finden sich in einer Reihe aktueller Publikationen, in denen die Möglichkeiten, aber auch die methodischen Probleme dieser neuen Erhebungstechniken thematisiert und diskutiert werden (vgl. u.a. Dillmann 2000; Birnbaum 2001; Reips/Bosnjak 2001; Theobald/Dreyer/Starsetzki 2001).

Übereinstimmend werden bestimmte Vorteile von Internetbefragungen gesehen, wie z.B. kurze Feldzeiten und die weitgehende Automatisierung der Durchführung und Auswertung (Batinic/Bosnjak 2000: 311f.). Ähnlich wie bei schriftlichen postalischen Umfragen fallen durch den Selbstausfüller-Modus keine Kosten für Interviewer an. Eines der methodischen Hauptprobleme besteht jedoch nach wie vor in der vergleichsweise geringen Internetnutzung. So liegt nach den Ergebnissen bevölkerungsrepräsentativer telefonischer Interviews des Instituts forsa in Deutschland der Anteil derjenigen, die das Internet bzw. World Wide Web zumindest hin und wieder nutzen, im Juli 2001 bei 42 Prozent der Gesamtbevölkerung ab 14 Jahren². D.h. weniger als die Hälfte potentieller Zielpersonen einer Bevölkerungsumfrage kann gegenwärtig für eine Internetbefragung erreicht werden (Abdeckungs- bzw. Coverage-Problem). Macht es angesichts einer derartigen Ausgangssituation überhaupt Sinn, das Internet schon jetzt für Umfragen zu nutzen? Zum Vergleich: Telefonische Bevölkerungsumfragen wurden hierzulande erst gegen Ende der 80er Jahre durchgeführt, als in der damaligen Bundesrepublik Deutschland eine Vollversor-

² Beim Institut forsa werden die Daten zur Internetnutzung auf der Basis von ca. 10.000 Interviews monatlich ermittelt. Die Internetnutzung steigt kontinuierlich: Zum Jahresanfang 2001 (Januar) lag sie bereits bei 35 Prozent, Mitte des vorigen Jahres im Juli 2000 bei nur 27 Prozent. Andere Studien kommen z.T. zu anderen Werten. Der Grund hierfür liegt in der Definition der Internetnutzung (vgl. hierzu u.a. Bandilla 1999: 11).

gung der Haushalte mit einem Telefon weitgehend sichergestellt war (Frey/Kunz/Lüschen 1990: 35).

Als methodisch weit weniger problematisch können in der gegenwärtigen Situation solche Erhebungen im Internet angesehen werden, in denen unter Zugrundelegung experimenteller Designs eng eingrenzbara funktionale Zusammenhänge im Mittelpunkt des Erkenntnisinteresses stehen (z.B. kognitions- und sozialpsychologische Experimente; vgl. u.a. Reips 2001). Wie sind jedoch Bevölkerungsumfragen über das Internet zu bewerten, wenn, wie schon erwähnt, ein Großteil der Bevölkerung von der Teilnahme an solchen Umfragen ausgeschlossen ist?

Nach Couper (2000) sind in diesem Zusammenhang vor allem zwei Ansätze diskussionswürdig: (1) „Pre-recruited panels of full population“ und (2) „Pre-recruited panels of Internet users“. Den Auswahlrahmen beim erstgenannten Ansatz bildet die Gesamtbevölkerung. Zur Realisierung dieses Ansatzes werden zufällig ausgewählte Zielpersonen mit der zur Teilnahme an den Befragungen notwendigen Technik ausgestattet. Die Teilnehmer sind Mitglieder eines Access-Panels, d.h. es handelt sich um Personen, die sich bereit erklärt haben, regelmäßig an Befragungen zu unterschiedlichen Themen teilzunehmen. Derart aufgebaute Access-Panels (vgl. hierzu u.a. Hoppe 2000) sind nicht zu verwechseln mit den in den Sozialwissenschaften bekannten Panels, bei denen an denselben Personen zu verschiedenen Zeitpunkten dieselben Variablen erhoben werden (vgl. u.a. Schnell/Hill/Esser 1995: 229). Access-Panels werden vorwiegend in der Marktforschung eingesetzt und zeichnen sich dadurch aus, dass über die Mitglieder dieser Panels eine Vielzahl gespeicherter Informationen vorliegen und die Panelmitglieder kurzfristig für unterschiedliche Befragungen und Tests zur Verfügung stehen (Hoppe a.a.O.: 147).

In Verbindung mit dem Internet wird dieser Ansatz aktuell in den USA von der Firma Knowledge Networks verfolgt, die 1998 vom SPSS-Mitbegründer Norman Nie sowie Douglas Rivers (beide Stanford University) gegründet wurde. Auch hier wird ein Access-Panel benutzt, das auf einer Zufallsstichprobe der US-amerikanischen Bevölkerung basiert, wobei die Panelmitglieder über ein telefonisches Screening sowie postalische Vorkontakte rekrutiert werden. Allen Mitgliedern dieses Panels wird die zur Teilnahme notwendige Hardware kostenfrei zur Verfügung gestellt. Als Gegenleistung wird erwartet, dass sich die Panelteilnehmer maximal einmal pro Woche an einer Kurzbefragung beteiligen. Zum Jahresende 2001 soll das Panel ins-

gesamt 250.000 Mitglieder umfassen.³ Mit dem „omni-Net“-Panel baut das forsa-Institut das erste vergleichbare Panel in Deutschland auf. Im Oktober 2001 umfasste das Panel 2.500 Haushalte, im kommenden Jahr wird es auf 5.000 Haushalte aufgestockt. 2004 soll das Panel rund 50.000 Haushalte umfassen (forsa: 2001).

Der hier beschriebene Ansatz löst zwar das derzeitige Coverage-Problem, dürfte jedoch angesichts der sehr hohen Kosten, und zwar insbesondere im Hinblick auf die erforderlichen Anfangsinvestitionen, auch zukünftig die Ausnahme darstellen. Das zweite von Couper als „Pre-recruited panels of Internet users“ bezeichnete Vorgehen beschränkt sich demgegenüber auf diejenigen Personen, die über einen Zugang zum Internet verfügen. Dadurch, dass die Bereitstellungskosten für Hardware entfallen, hat dieser Ansatz inzwischen vor allem in der Marktforschung eine starke Verbreitung gefunden⁴. Den Auswahlrahmen bilden dabei Internetnutzer. Sie werden über ein telefonisches Screening auf der Basis einer Zufallsstichprobe rekrutiert⁵. Auch hier handelt es sich um Mitglieder eines Access Panels, die sich für regelmäßige Befragungen zur Verfügung stellen.

Vergleicht man die beiden von Couper genannten Ansätze, so stellt der erstgenannte unter methodischen Gesichtspunkten sicherlich das Optimum dar. Allerdings dürfte, wie schon gesagt, eine derartige Vorgehensweise angesichts der hohen Kosten als Alternative zu den bisher bekannten und praktizierten Erhebungstechniken (Face-to-Face und CATI, schriftlich postalisch) nur von eingeschränkter Bedeutung sein. Mittelfristig sehen wir ein starkes Ansteigen der Bedeutung von Umfragen, die entsprechend dem zweiten Ansatz konzipiert sind, d.h. über ein CATI-Screening

3 Ausführliche Informationen zur Methodik und Vorgehensweise finden sich auf den Webseiten des Unternehmens (<http://www.knowledgenetworks.com>).

4 Eine Übersicht aktueller Online-Panels findet sich auf der Website der Arbeitsgruppe CAI bei ZUMA. Zu finden unter: http://www.gesis.org/Forschung/Online_Research/Informationsquellen/links/onlinepanels.htm.

5 Neben dem hier beschriebenen Typus sind *selbstrekrutierte* Online-Access-Panels weit verbreitet. Diese 'Befragtenpools' basieren auf Rekrutierungsansätzen, bei denen durch Hinweise auf eine Registrierungsseite im Internet zur Teilnahme eingeladen wird. Bei dieser Art von Panels sind Selbstselektionseffekte zu erwarten, da es allein den Teilnehmern überlassen bleibt, ob sie sich registrieren oder nicht registrieren. Die im Frühjahr 2001 gemeinsam vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM), der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute (ASI), dem Berufsverband Deutscher Markt- und Sozialforscher (BVM) sowie der Deutschen Gesellschaft für Online-Forschung (DGOF) herausgegebenen "Standards zur Qualitätssicherung für Online-Befragungen" weisen in aller Deutlichkeit auf die Qualitätsmängel dieser Art von Online-Panels hin (vgl. ADM et al. 2001). Die Standards sind abrufbar unter der Adresse <http://www.adm-ev.de>.

rekrutierte Teilnehmer eines Online-Access-Panels. Zwar beschränkt sich die Grundgesamtheit bei diesem Ansatz immer auf die jeweils aktuelle Gruppe von Personen mit Zugang zum Internet. Zu berücksichtigen ist dabei jedoch, dass auch in Deutschland die Internetnutzung stetig zunimmt⁶, was bedeutet, dass sich das Problem des Undercoverage (vgl. Schnell et al. a.a.O.: 258) auch bei Bevölkerungsumfragen zukünftig zunehmend verringern dürfte.

In der hier vorliegenden Studie stellt dieser zweite Ansatz („Pre-recruited panels of Internet users“) die Grundlage der Internetbefragung dar. Die über das Internet erhobenen Daten werden mit Daten verglichen, die über einen schriftlichen Fragebogen in einer allgemeinen Bevölkerungsumfrage erhoben wurden. D.h. es wird mit zwei unterschiedlichen Grundgesamtheiten operiert: Zum einen mit der Grundgesamtheit der Nutzerinnen und Nutzer des Internet und zum anderen mit der Grundgesamtheit der Allgemeinbevölkerung. Aus beiden Grundgesamtheiten wurden Zufallsstichproben gezogen, denen ein inhaltlich identisches Fragenprogramm zur Beantwortung vorgelegt wurde.

Alle bisher bekannten und publizierten Studien zur Internetnutzung⁷ zeigen, dass sich Internetnutzer hinsichtlich Alter, Bildungsabschluss und Geschlecht deutlich von der Allgemeinbevölkerung unterscheiden. Folglich dürfte es nicht überraschen, wenn sich bei einer Vielzahl von Fragen die Antworten der Internetnutzer von denen der Allgemeinbevölkerung unterscheiden. Wie sieht es jedoch aus, wenn für die Gruppe der Internetnutzer mit Anpassungsgewichten operiert wird (vgl. Gabler 1994: 78), also die Struktur der Internetstichprobe hinsichtlich Alter, Bildung und Geschlecht an die Struktur der Bevölkerungsstichprobe angeglichen wird? Kann eine solche Vorgehensweise bei der gegenwärtigen Internetpenetration sinnvoll sein? Treten mögliche Unterschiede in den Ergebnissen auch dann auf, wenn demografische Merkmale der Befragten, wie Bildung, in beiden Stichproben konstant gehalten werden? Im Folgenden wird versucht, auf diese Fragen erste Antworten zu geben.

⁶ Belege hierfür finden sich u.a. in der ARD/ZDF-Online-Studie (van Eimeren/Gerhard/Frees 2001) sowie speziell zur Internetnutzung von Kindern in der Studie KIM 2000 (Feierabend/Klingler 2001).

⁷ Neben den in Fußnote 6 aufgeführten Quellen sind in diesem Zusammenhang auch der schon mehrmals durchgeführten GfK Online-Monitor sowie die ebenfalls schon wiederholt durchgeführte ARD/ZDF Online Studie zu erwähnen; vgl. hierzu auch Bandilla/Bosnjak (2000: 11f).

2. Konzeption und Durchführung der Untersuchung

Im Mittelpunkt der Studie steht die Frage, ob sich die Ergebnisse einer Online-Befragung mit einem telefonisch vorrekrutierten Access-Panel von einer auf traditionell-schriftlichem Weg durchgeführten Befragung einer Zufallsstichprobe der Allgemeinbevölkerung unterscheiden. Beide Befragungen waren im Hinblick auf die Befragungsinhalte sowie den Befragungsablauf identisch.

2.1 Fragenprogramm

Als Fragenprogramm wurde das Modul „Umwelt“ aus dem *International Social Survey Programme* (ISSP) verwendet. Das ISSP⁸ ist ein weltweiter Forschungsverbund, der regelmäßig sozialwissenschaftliche Umfragen mit wechselnden Themenschwerpunkten durchführt. Die Fragebögen sind in der Regel zum Selbst-Ausfüllen konzipiert und werden nationalen Zufallsstichproben der Allgemeinbevölkerung vorgelegt.

Im Jahr 2000 war für den deutschen Teil die ISSP-Umfrage mit der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) gekoppelt. Dabei sollte im Anschluss an die mündliche ALLBUS-Umfrage der schriftliche ISSP-Fragebogen von den Zielpersonen bearbeitet werden. Neben dem Umwelt-Modul wurde dabei auch das Modul „Soziale Ungleichheit“ eingesetzt. In der hier vorliegenden Studie wird nur das Modul „Umwelt“ berücksichtigt, das bei der Hälfte der Befragten zum Einsatz kam.

Das schriftliche Fragenprogramm bestand aus 25 Fragen mit insgesamt 61 Items, darunter 30 Items, in denen (im weitesten Sinn) Einstellungen zum Thema Umwelt erhoben wurden. Ein Großteil der Fragen repliziert solche, die schon 1993 im ISSP bzw. der Studie *Research into Environmental Attitudes and Perceptions* (REAP) gestellt wurden (vgl. Witherspoon/Mohler/Harkness 1995).

2.2 Stichproben und Feldverlauf

Wie bereits in den vorherigen ALLBUS-Studien bestand die Grundgesamtheit des ALLBUS/ISSP 2000 aus den in der Bundesrepublik Deutschland in Privathaushalten lebenden erwachsenen Personen ab 18 Jahren. Die Stichprobe wurde mit einem zweistufigen Ziehungsverfahren gebildet: 1) Auswahl von Gemeinden, 2) Auswahl von Personenadressen in den Gemeinden über die Einwohnermelderegister. Die Stichprobenziehung und die Feldarbeit wurden von Infratest Burke Sozialforschung

8 Ausführliche Informationen zum ISSP sind unter <http://www.issp.org> zu finden.

durchgeführt. Ausführliche Informationen finden sich im Methodenbericht zum ALLBUS 2000 (Koch/Wasmer/Harkness/Scholz 2001: 43ff.). Bei einer Ausschöpfung von 46,9 Prozent in West- bzw. 53,7 Prozent in Ostdeutschland wurden insgesamt 3138 persönlich-mündliche Interviews mit dem ALLBUS-Fragenprogramm durchgeführt. An der im Anschluss an diese mündlichen Interviews mittels schriftlichem Fragebogen durchgeführten ISSP-Befragung nahmen insgesamt 2933 Personen teil, darunter 1485 Personen an dem hier interessierenden Umwelt-Fragebogen des ISSP. Eine nahezu gleich große Gruppe von Befragten beurteilte das Fragenmodul zur sozialen Ungleichheit, das, wie schon ausgeführt, im Rahmen dieser Studie nicht berücksichtigt wurde. Die Erhebung fand zwischen Januar und Juli 2000 statt.

Die Grundgesamtheit der Teilnehmer und Teilnehmerinnen der Online-Befragung bildeten Personen ab 14 Jahren, die in privaten Haushalten der Bundesrepublik leben und die das Internet zumindest „selten“ nutzen sowie über eine E-Mail-Adresse zu Hause verfügen. Es handelte sich hierbei um Mitglieder des forsa-Online-Panels, d.h. eines Access-Panels des Berliner forsa-Institutes mit Internetnutzern, die im Rahmen des täglichen CATI-Omnibusses auf der Basis einer Bevölkerungsstichprobe rekrutiert wurden. Die Bereitschaft zur Teilnahme an dem Panel der telefonisch rekrutierten Mitglieder lag bei 39 Prozent, bezogen auf die soeben beschriebene Grundgesamtheit. Entsprechend der kürzlich erschienenen „Standards zur Qualitätssicherung für Online-Befragungen“ (ADM et al. a.a.O.) erfolgt bei diesem Access-Panel eine Validierung der Mitglieder, indem im Anschluss an die telefonische Rekrutierung den teilnahmebereiten Probanden ein Brief - samt Informationsblatt zum Datenschutz - zugesandt wird. Hierin wird eine Bestätigung der Mitwirkungsbereitschaft angefordert. Probanden, die ihre Mitwirkung bestätigen möchten, müssen auf der Panel-Website eine Grunderhebung zur Soziodemografie ausfüllen. Eine Incentivierung der Mitglieder dieses Panels erfolgt über Bonuspunkte, die einem wahren Währungsbetrag entsprechen. Dabei orientiert sich der Umfang der vergebenen Bonuspunkte - als kleine Aufwandsentschädigung für ihre Onlinezeit und die entstehenden Telefon- und Providerkosten konzipiert - an der Länge einer Befragung (Dauer der Befragung).

Per E-mail wurden zur Teilnahme an der ISSP Umfrage 901 Mitglieder dieses Online-Access-Panels aufgefordert. Von diesen hatten nach einer Feldzeit von sieben Wochen 500 Personen den Online Fragebogen vollständig ausgefüllt (55 Prozent Ausschöp-

fung)⁹. Ähnlich wie bei schriftlichen Befragungen zeigt sich auch bei der Online-Befragung, dass kurz nach Ankündigung der Befragung die Teilnahme am höchsten ist und danach kontinuierlich abnimmt. Auch am Wochenende und abends wurde eine höhere Teilnahme registriert als in den übrigen Zeiten, was zum einen mit den Kosten der Internetnutzung zusammenhängen dürfte und zum anderen aber auch daran liegt, dass viele berufstätige Teilnehmer erst nach Feierabend das Internet nutzen.

Bezogen auf Art und Umfang der Internetnutzung im Befragungszeitraum (8. Juni bis 26. Juli 2000) waren beim Vergleich der soziodemografischen Merkmale zwischen der in den CATI-Interviews festgestellten Internet-Grundgesamtheit, den am Panel teilnehmenden Personen sowie den tatsächlichen Teilnehmern des Online-Panels keine bedeutsamen Unterschiede festzustellen.

2.3 Befragungstool

Der schriftliche (Papier-) Fragebogen sollte den Befragten der Online-Stichprobe inhaltlich und prozedural identisch präsentiert werden wie den Befragten der Bevölkerungstichprobe, die den schriftlichen Fragebogen zu bearbeiten hatten. Dazu wurde das Umfragetool „Rogator“ des gleichnamigen Softwareherstellers verwendet. Es verfügt über eine Reihe von Möglichkeiten für eine an das Medium Computer angepasste Gestaltung von Umfrageseiten. Die erhobenen Daten können über ein Konvertierungsprogramm unmittelbar einer statistischen Analyse (z.B. in SPSS) unterzogen werden. Ausführliche Informationen zu diesem Tool finden sich unter dem URL <http://www.rogator.de>. Mit Hilfe dieses Tools war eine im Hinblick auf die Befragungsinhalte und -reihenfolge identische Umsetzung des schriftlichen ISSP Fragebogens in eine Web-basierte Variante möglich¹⁰. Es wurde darauf geachtet, dass bereits bei einer Auflösung von 800*600 Bildpunkten alle Items einer Frage auf einer einzigen Web-Seite darstellbar waren. Per Mausklick wurden dann die Ant-

⁹ Die Feldzeit dieser Online-Befragung war untypisch lang und die Ausschöpfung vergleichsweise niedrig, da forsa lediglich die Probanden des Online-Panels ZUMA zur Verfügung gestellt hatte. Im Regelfall nehmen die forsa-Online-Panel-Probanden (das forsa-Online-Panel umfasst im Juli 2001 rund 3.000 Haushalte) nur an Umfragen teil, die auf dem forsa eigenen Server und mit Hilfe spezieller von forsa entwickelter Befragungs- und Panelverwaltungssoftware durchgeführt werden. Eine Feldzeit von lediglich einer Woche wird benötigt, um eine Befragung der repräsentativ ausgewählten Internetnutzer im forsa-Online-Panel zu realisieren; die Ausschöpfung liegt aktuell bei ca. 75 Prozent.

¹⁰ Wir danken Wolfgang Neubarth, der während seines Praktikums am ZUMA mit der Umsetzung des Fragenprogramms mit Hilfe der Rogatorsoftware betraut war.

worten an den Befragungsserver übermittelt und das bzw. die nachfolgende(n) Item(s) eingeblendet.

Die beiden nachfolgenden Abbildungen verdeutlichen die Art der Bildschirmpräsentation. Abbildung 1 zeigt eine Matrixfrage des Web-Fragebogens, also eine Frage mit mehreren tabellarisch angeordneten Items. Wie in der schriftlichen (Papier-) Version wird auch die Antwortalternative „Kann ich nicht sagen“ angeboten. Per Mausklick waren die jeweiligen Antwortoptionen auszuwählen, ein Mausklick auf „Weiter“ führte zur nächsten Frage. Fortschrittsanzeigen informieren die Befragten auf jeder Bildschirmseite über den Stand bzw. den Umfang der noch folgenden Fragen.

Abbildung 1: Beispiel für die Präsentation einer Matrixfrage im Web-Fragebogen

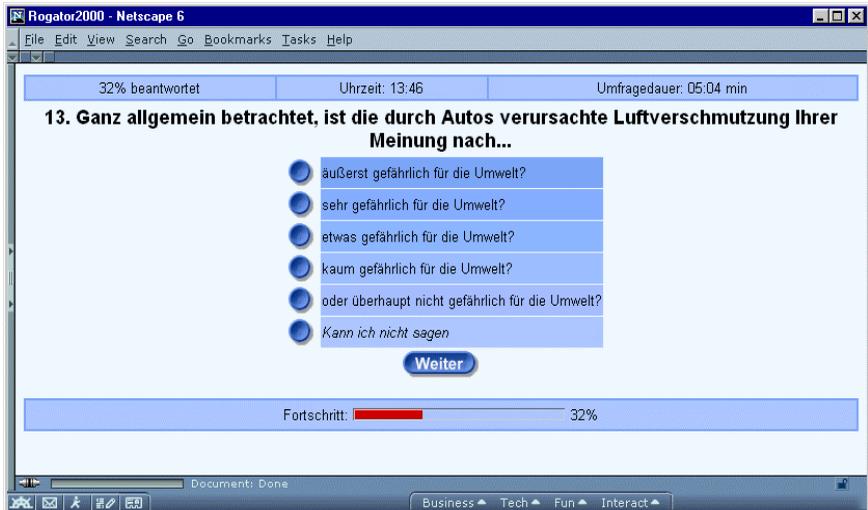
The screenshot shows a web browser window titled "Rogator2000 - Netscape 6". The survey progress bar at the top indicates "26% beantwortet", "Uhrzeit: 13:42", and "Umfragedauer: 01:36 min". The main question is: "11. Bitte geben Sie zu jeder der folgenden Aussagen an, inwieweit Sie zustimmen oder nicht zustimmen".

	Stimme stark zu	Stimme eher zu	Weder noch	Stimme eher nicht zu	Stimme überhaupt nicht zu	Kann ich nicht sagen
Für jemand wie mich ist es einfach zu schwierig, viel für die Umwelt zu tun	<input type="radio"/>	<input type="radio"/>				
Ich tue das, was für die Umwelt richtig ist, auch wenn mich das mehr Geld oder Zeit kostet	<input type="radio"/>	<input type="radio"/>				
Es gibt wichtigeres im Leben zu tun, als die Umwelt zu schützen	<input type="radio"/>	<input type="radio"/>				
Es ist zwecklos, meinen Beitrag für die Umwelt zu leisten, solange andere sich nicht genauso verhalten	<input type="radio"/>	<input type="radio"/>				
Viele Behauptungen über die Gefährdung der Umwelt sind übertrieben	<input type="radio"/>	<input type="radio"/>				

Below the table is a "Weiter" button. At the bottom, a progress bar shows "Fortschritt: 26%".

Beim zweiten Beispiel handelt es sich um eine Einzelfrage, die in gleicher Weise per Mausklick zu beantworten war (vgl. Abbildung 2).

Abbildung 2: Beispiel für die Präsentation einer Einzelfrage im Web-Fragebogen



3. Ergebnisse

Wie bereits erwähnt, basierte die schriftliche ISSP Umfrage auf einer bevölkerungsrepräsentativen Stichprobe von Personen, die 18 Jahre und älter waren. Zur Grundgesamtheit der Internetnutzer zählten demgegenüber auch jüngere Internetnutzer im Alter zwischen 14 und 17 Jahren, deren Daten aus Gründen der Vergleichbarkeit bei allen nachfolgenden Analysen ausgeschlossen wurden¹¹.

Beim Vergleich der demografischen Merkmale der Befragungsteilnehmer zeigen sich die zu erwartenden Unterschiede entsprechend der jeweils zugrundeliegenden Auswahlgesamtheiten: Unter den Teilnehmerinnen und Teilnehmern der Online-Befragung sind im Vergleich zur schriftlichen Befragung deutlich höhere Prozentanteile bei den Männern zu verzeichnen (66,1 Prozent versus 48,2 Prozent). Gravierende Unterschiede zeigen sich auch beim Alter und der Verteilung der Bildungs-

¹¹ Hierdurch reduzierte sich die Datenbasis der Online-Befragten von ursprünglich 500 auf 475 Personen.

abschlüsse: Die Mitglieder des Online- Panels sind im Durchschnitt jünger und verfügen größtenteils über einen höheren Bildungsabschluss¹² (vgl. Tabelle 1).

Tabelle 1: Vergleich soziodemografischer Merkmale der traditionell durchgeführten Befragung mit dem Web Survey (Angaben in Prozent)

	Schriftliche Befragung mittels Zufallsstichprobe (Basis: Allgemeinbevölkerung; n = 1485)	Web-basierte Befragung mittels repräsentativem Online-Access-Panel (Basis: Internetnutzer; n = 475)
Geschlecht		
Männer	48,2	66,1
Frauen	51,8	33,9
Alter		
18-29	16,4	40,4
30-44	31,1	44,8
45-60	27,0	13,6
über 60	25,5	2,1
Bildungsabschluss		
einfacher	45,2	8,6
mittlerer	31,8	21,4
höherer	23,0	70,0

Angesichts dieser Unterschiede stellt sich die Frage, inwieweit überhaupt ein Vergleich der Ergebnisse möglich und sinnvoll ist. In einem ersten Schritt wurden - analog der Ost-West-Gewichtung beim ALLBUS - gemäß der bei Gabler (1994: 78) beschriebenen Formel für die Gruppe der Online-Befragten Anpassungsgewichte berechnet.

Unter Berücksichtigung der Merkmale „Geschlecht“, „Alter“ und „Bildungsabschluss“ ergeben sich - insbesondere für die Befragten des Online-Panels mit einem einfachen Bildungsabschluss - teilweise Anpassungsgewichte, die über dem Wert Fünf liegen und somit als problematisch angesehen werden müssen. Eine solche Anpassungsgewichtung führt zu inkonsistenten Befunden, was beispielhaft an drei

¹² Die drei hier verwendeten Bildungskategorien basieren auf folgenden Festlegungen: A) Ein einfacher Bildungsabschluss lag vor, wenn kein Schulabschluss bzw. ein Volks- oder Hauptschulabschluss vorhanden waren. B) Der mittlere Bildungsabschluss wurde definiert über einen Realschulabschluss bzw. die Mittlere Reife. C) Zu den höheren Bildungsabschlüssen zählten die Fachhochschul- und Hochschulreife sowie das Abitur.

Items illustriert werden kann. Im ersten Beispiel („Tierversuche“) ist mit und ohne Anpassungsgewichtung der Online-Daten eine starke Übereinstimmung mit den schriftlich erhobenen Daten festzustellen (Tabelle 2).

Tabelle 2: Beispiel für eine hohe Übereinstimmung zwischen den nicht-gewichteten und gewichteten Daten der Web-Befragung und der schriftlichen Erhebung für das Item: „Tierversuche sind gerechtfertigt, wenn dies Menschenleben retten könnte“¹³

	Online		schriftlich (Offline)
	nicht-gewichtet	gewichtet	
Stimme stark zu	22,5	18,4	19,5
Stimme eher zu	45,7	46,2	44,5
Weder noch	8,2	7,6	11,1
Stimme eher nicht zu	14,6	15,3	15,3
Stimme überhaupt nicht zu	9,0	12,5	9,5
	100% (n=466)	100% (n=463)	100% (n=1428)

Im Vergleich zu den schriftlich erhobenen Daten zeigen sich sowohl für die nicht-gewichteten als auch für die gewichteten Daten der Web-Befragung keine signifikanten Unterschiede. Für die nicht-gewichteten Daten der Web-Befragung ergab sich: $\chi^2=4.90$; $df=4$; $p=.298$; 2-seitig. Gewichtete Daten: $\chi^2=7.92$; $df=4$; $p=.094$; 2-seitig.

Im zweiten Beispiel („Wahrscheinlichkeit eines Atomunfalls“) führt die Anpassungsgewichtung zu einer höheren Übereinstimmung mit den schriftlich erhobenen Daten (Tabelle 3).

Für den Vergleich der nicht-gewichteten Daten der Web-Befragung mit den schriftlich erhobenen Daten ergab sich im zweiten Beispiel: $\chi^2=20.70$; $df=3$; $p=.000$; 2-seitig. Bei den gewichteten Daten zeigten sich keine signifikanten Unterschiede: $\chi^2=3.017$; $df=3$; $p=.389$; 2-seitig.

¹³ Abweichungen zur Brutto-Stichprobe sind bei dieser, wie auch bei den folgenden Tabellen, auf fehlende Angaben bei einzelnen Variablen zurückzuführen.

Tabelle 3: Beispiel für eine 'verbesserte' Übereinstimmung zwischen den Daten der Web-Befragung und der schriftlichen Erhebung durch Gewichtung für das Item: „Für wie wahrscheinlich halten Sie es, dass in den nächsten fünf Jahren ein Unfall in einem Atomkraftwerk zu langfristigen Umweltschäden in vielen Ländern führen wird?“

	Online		schriftlich (Offline)
	nicht-gewichtet	gewichtet	
Sehr wahrscheinlich	12,1	11,0	12,8
Wahrscheinlich	33,8	40,2	42,9
Unwahrscheinlich	38,7	38,7	35,3
Sehr unwahrscheinlich	15,4	10,1	8,9
	100% (n=455)	100% (n=455)	100% (n=1207)

Das dritte Beispiel („Wirtschaftliches Wachstum nötig für Umweltschutz“) dokumentiert demgegenüber eine geringere Übereinstimmung für die gewichteten Daten (Tabelle 4).

Tabelle 4: Beispiel für eine 'verschlechterte' Übereinstimmung zwischen den Daten der Web-Befragung und der schriftlichen Erhebung durch Gewichtung für das Item: „Um die Umwelt schützen zu können, braucht Deutschland wirtschaftliches Wachstum.“

	Online		schriftlich (Offline)
	nicht-gewichtet	gewichtet	
Stimme stark zu	11,6	11,7	16,9
Stimme eher zu	43,1	41,3	40,2
Weder noch	15,4	10,8	16,6
Stimme eher nicht zu	25,3	32,2	21,8
Stimme überhaupt nicht zu	4,6	4,0	4,4
	100% (n=455)	100% (n=453)	100% (n=1332)

Der Unterschied zwischen den nicht-gewichteten Daten der Web-Befragung und den schriftlich erhobenen Daten ist in diesem Beispiel auf dem 5-Prozent-Niveau nicht signifikant ($\chi^2=8.744$; $df=4$; $p=.068$; 2-seitig) im Gegensatz zu den gewichteten Daten: $\chi^2=28.442$; $df=4$; $p=.000$; 2-seitig.

Vor dem Hintergrund dieses insgesamt inkonsistenten Bildes in den Ergebnissen wurden in einem weiteren Analyseschritt für alle Items, in denen im weitesten Sinn Umwelteinstellungen erhoben wurden (insgesamt 30 Items¹⁴) multiple Mittelwertvergleiche durchgeführt. Die Vergleiche beziehen sich auf den Erhebungsmodus (online versus schriftlich), und zwar sowohl für ungewichtete Online-Daten als auch für gewichtete Online-Daten. Ein dritter Vergleich berücksichtigt nur solche Befragte, die einen höheren Bildungsabschluss aufweisen. Da die Stichprobe der Online-Befragten im Vergleich zur Stichprobe der Allgemeinbevölkerung in weitaus stärkerem Maß höhere Bildungsabschlüsse aufweist (70 Prozent gegenüber 23 Prozent; vgl. Tabelle 1 weiter oben), liegen diesem Vergleich weitgehend parallelisierte Befragtengruppen zugrunde, wobei sich die Fallzahlen in beiden Gruppen auf ähnlichem Niveau bewegen (Online: N = 324; schriftlich: N = 337).

Tabelle 5: Anzahl und prozentualer Anteil statistisch bedeutsamer Unterschiede für drei Gruppen abhängiger Mittelwertvergleiche (t-Tests) über jeweils 30 Umwelteinstellungs-Items

	Online (ohne Gewichtung) vs. schriftlich	Online (mit Gewichtung) vs. schriftlich	nur Personen mit höherem Bildungsabschluss: online vs. schriftlich
Anzahl Items mit sign. Mittelwertunterschieden	22 Items (= 73 %)	20 Items (= 67 %)	8 Items (= 27 %)

Anmerkung: Pro Vergleichsgruppe wurden jeweils 30 zweiseitige t-Tests mit Bonferroni-Adjustierung des alpha-Niveaus der Einzeltests zur Erzielung einer insgesamten Irrtumswahrscheinlichkeit von 5% berechnet.

Die in Tabelle 5 zusammengefassten Ergebnisse zeigen einen deutlichen Trend: Nach t-Tests unterscheiden sich 73% aller Items (22 von 30 Items; vgl. Tabelle 5) zwischen der online und schriftlichen Version signifikant, wenn alle Befragten die Analysegrundlage bilden. Ähnlich fällt das Ergebnis aus, wenn für die Daten der Online-Befragten eine Anpassungsgewichtung vorgenommen wird (signifikante Unterschiede bei 20 von 30 Items). Erst dann, wenn unter beiden Erhebungsmodi

¹⁴ Es handelt sich hierbei um folgende Fragen aus dem ISSP-Fragebogen: Fragen 4, 5, 7, 8, 10, 11, 16, 17, 23. Eine Kurzbeschreibung der Fragen und Items findet sich im ALLBUS Methodenbericht (Koch et al. 2001: 34ff.), der auch unter der URL http://www.gesis.org/Dauerbeobachtung/Allbus/service_guide.htm einsehbar ist.

nur die Befragten berücksichtigt werden, die über einen vergleichbaren (= höheren) Bildungsabschluss verfügen, ändert sich das Bild: Nur noch bei acht Items sind zwischen online und schriftlich erhobenen Daten signifikante Unterschiede zu erkennen.

Bei der Anwendung von Signifikanztests kommt erst durch die Bestimmung von Effektstärken die Bedeutsamkeit der ermittelten Unterschiede zum Ausdruck (vgl. hierzu Bortz/Döring 1995: 563ff.). Deshalb wurde für alle t-Tests das entsprechende Effektmaß 'd' berechnet. Es entspricht der Mittelwertsdifferenz zwischen zwei Gruppen geteilt durch die gepoolte Varianz der Vergleichsgruppen. Eine Effektstärke in Höhe von $d = 0,2$ gilt als 'kleiner' Effekt, eine von $d = 0,5$ als 'mittlerer' und von $d = 0,8$ als 'großer' Effekt (vgl. auch Bortz/Döring 1995: 568).

Beschränkt sich der Vergleich online versus schriftlich nur auf Befragte, die über einen höheren Bildungsabschluss verfügen, zeigen sich für den größten Teil der Items nur wenig bedeutsame Mittelwertdifferenzen (8 von 30 Items; vgl. Tabelle 6, 3. Spalte). Werden dagegen die Antworten aller Befragten in die Analysen einbezogen, sind sowohl mit als auch ohne Anpassungsgewichtung die beschriebenen Unterschiede in den Ergebnissen zu erkennen.

Für den Vergleich online versus schriftlich ist es also nach diesen Ergebnissen wenig sinnvoll, eine Anpassungsgewichtung unter Zugrundelegung klassischer soziodemografischer Merkmale vorzunehmen. Die Unterschiede zwischen der Gruppe der Internetnutzer und der Allgemeinbevölkerung sind gegenwärtig noch zu stark ausgeprägt. Ein anderes Bild ergibt sich jedoch, wenn für den Vergleich der Erhebungsmodi die Befragtengruppen weitgehend parallelisiert werden. In die nachfolgenden Analysen wurden deshalb nur solche Befragte einbezogen, die über einen vergleichbaren (hier: höheren) Bildungsabschluss verfügten.

Tabelle 6: Anzahl und prozentualer Anteil nicht-signifikanter und signifikanter Mittelwertsunterschiede (t-Tests) und zugehörige Effektstärken

	Online (ohne Gewichtung) vs. schriftlich	Online (mit Gewichtung) vs. schriftlich	nur Personen mit höherem Bildungsabschluss: online vs. schriftlich
Nicht-signifikante Unterschiede	8 Items (= 27 %)	10 Items (=33%)	22 Items (=73%)
Signifikante Unterschiede bei einer kleinen Effektstärke	12 Items (= 40 %)	8 Items (= 27 %)	3 Items (= 10 %)
Signifikante Unterschiede bei einer mittleren Effektstärke	9 Items (= 30 %)	11 Items (= 37 %)	5 Items (= 17 %)
Signifikante Unterschiede bei einer großen Effektstärke	1 Item (= 3 %)	1 Item (= 3 %)	0 Items

Anmerkungen: Effektstärke $d = (\text{Mittelwertsdifferenz zweier Vergleichsgruppen}) / (\text{gepoolte Varianz})$. 'kleiner Effekt': $d < .30$; 'mittlerer Effekt': $d \geq .30 < .55$; 'großer Effekt' $d \geq .55$.

Bei näherer Betrachtung derjenigen Itempaare, die nur einen kleinen Mittelwertsunterschied aufweisen, zeigt sich, dass es sich vor allem um Fragen handelt, in denen es inhaltlich um die Einschätzung von Gefahrenpotentialen für die Umwelt geht. So z.B. in den Fragen 11 und 23 des ISSP-Fragenprogramms, die folgendermaßen formuliert waren:

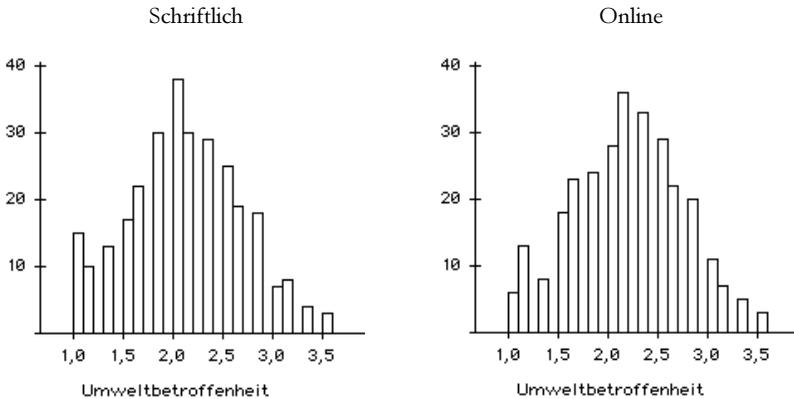
„Ganz allgemein betrachtet, wie gefährlich sind Ihrer Meinung nach die folgenden Dinge?“

- Die durch die Industrie verursachte Luftverschmutzung (*Fr. 11*)
- Die in der Landwirtschaft eingesetzten Pflanzenschutzmittel und Chemikalien (*Fr. 11*)
- Das Verschmutzen der Bäche, Flüsse und Seen in Deutschland (*Fr. 11*)
- Ein durch den „Treibhauseffekt“ verursachter weltweiter Temperaturanstieg (*Fr. 11*)
- Die Genmanipulation von Nutzpflanzen (*Fr. 11*)
- Kernkraftwerke (*Fr. 23*)“

Diese Items wurden ebenfalls in der 1993 durchgeführten REAP-Studie (vgl. Witherspoon et al. a.a.O.: 35ff.) zur Erfassung eines Einstellungskonstruktes eingesetzt, das mit „allgemeine Umweltbetroffenheit“ umschrieben werden kann. Die Bildung der neuen Skala erfolgte analog zu der in der REAP-Studie vorgenommenen Berechnung¹⁵.

Beim Vergleich der Mittelwerte der Online-Befragten mit denen, die auf schriftlichem Weg erhoben wurden, zeigen sich nahezu identische Werte (Online: 2,18; Schriftlich: 2,12). Der mittels t-Test durchgeführte Mittelwertvergleich erbrachte keine signifikanten Unterschiede. Zur Illustration des nahezu identischen Antwortverhalten unter beiden Erhebungsbedingungen dienen die beiden in Abbildung 3 wiedergegebenen Histogramme, die nahezu identische Verteilungen zeigen.

Abbildung 3: Verteilungen und Verteilungskennwerte der gebildeten Skala 'Umweltbetroffenheit' unter beiden Erhebungsmodi



M= 2,12 ; SD= 0,60; Median= 2,16; n= 291 M= 2,18; SD= 0,58; Median= 2,16; n= 287

¹⁵ Basis waren die 5-stufigen Antwortskalen von „äußerst gefährlich für die Umwelt“ (Wert: 1) bis „überhaupt nicht gefährlich für die Umwelt“ (Wert: 5). Der Skalenwert wurde gebildet durch Aufsummierung der Antworten der Befragten auf jedes Item dividiert durch die Anzahl der Items (mean score). Der Maximalwert der Skala liegt demnach beim Wert 6 (= starke Ablehnung aller Items), der Minimalwert beim Wert 1 (= starke Zustimmung zu allen Items). Ein geringer Wert ist also gleichbedeutend mit einer starken Betroffenheit hinsichtlich der Umweltgefahren. Cronbach's Alpha betrug .71.

D.h. bei weitgehend gleichen demografischen Hintergrundmerkmalen der Befragten zeigen sich auch auf der Konstruktebene „Umweltbetroffenheit“ so gut wie keine Unterschiede zwischen der Web-Variante und der in Form eines Papier-Fragebogens durchgeführten Erhebung.

Ein ähnliches Ergebnis ist festzustellen, wenn die Interkorrelationen thematisch ähnlicher Items unter den beiden Erhebungsmodi miteinander verglichen werden. In Anlehnung an die REAP-Studie wurden für diesen Vergleich solche Items gewählt, die einem 'romantic environmental imagery' oder einem 'modern materialistic' Bild entsprechen (Witherspoon a.a.O.: 19ff.). Die zum ersten Block gehörenden Items waren folgendermaßen formuliert:

- Wir vertrauen zu sehr der Wissenschaft und nicht genug unseren Gefühlen und dem Glauben.
- Alles in allem schadet die moderne Wissenschaft mehr als sie nützt.
- Fast alles, was wir in unserer modernen Welt tun, schadet der Umwelt.
- Wirtschaftswachstum schadet immer der Umwelt.

Die Items des zweiten Blocks waren:

- Die moderne Wissenschaft wird unsere Umweltprobleme bei nur geringer Veränderung unserer Lebensweise lösen.
- Wir machen uns zu viele Sorgen über die Zukunft der Umwelt und zu wenig um Preise und Arbeitsplätze heutzutage.
- Die Leute machen sich zu viele Sorgen, dass der menschliche Fortschritt der Umwelt schadet.
- Um die Umwelt schützen zu können, braucht Deutschland wirtschaftliches Wachstum.

Für die Items des ersten Blocks ergaben sich die folgenden Interkorrelationen (Spearman's rho) unter der Online- und der schriftlichen Variante (vgl. Tabelle 7):

Tabelle 7: Vergleich der Interkorrelationen bei online (grau unterlegt) und schriftlich erhobenen Daten von 'romantic environmental imagery' Items

	Wissenschaft: mehr Schaden als Nutzen		fast alles schadet der Umwelt		Wirtschaftswachstum schadet Umwelt	
zu viel Vertrauen in Wissenschaft	<u>.43</u>	p = 0.33	<u>.17</u>	p = 0.18	<u>.10</u>	p < 0.05
	.46		.24		.25	
Wissenschaft: mehr Schaden als Nutzen			<u>.26</u>	p = 0.49	<u>.22</u>	p = 0.24
			.27		.27	
fast alles schadet der Umwelt					<u>.38</u>	p = 0.51
					.38	

Die für die Online-Variante geltenden Korrelationen sind grau unterlegt, die für die schriftliche Variante errechneten Korrelationen sind in den jeweiligen Zellen darunter aufgeführt. Die p-Werte¹⁶ geben an, mit welcher Wahrscheinlichkeit beide Korrelationskoeffizienten gleich sind. Bei Betrachtung der Ergebnisse ist festzustellen, dass nur in einem von sechs Korrelationsvergleichen ein signifikanter Unterschied zu ermitteln ist.

Ein ähnliches Ergebnis zeigt sich bei den Einstellungs-Items des zweiten Blocks (vgl. Tabelle 8).

Tabelle 8: Vergleich der Interkorrelationen bei online (grau unterlegt) und schriftlich erhobenen Daten von 'modern materialistic' Items

	zu viele Sorgen um Umwelt		zu viele Sorgen um Fortschrittsschäden		wirtschaftl. Wachstum nötig für Umweltschutz	
Wissenschaft: wird Umweltprobl. lösen	<u>.39</u>	p < 0.05	<u>.23</u>	p = 0.42	<u>.24</u>	p = 0.27
	.20		.22		.20	
zu viele Sorgen um Umwelt			<u>.42</u>	p = 0.18	<u>.31</u>	p = 0.25
			.48		.26	
zu viele Sorgen um Fortschrittsschäden					<u>.29</u>	p = 0.11
					.20	

¹⁶ Berechnet mittels eines Online-Tools, zu erreichen unter: <http://home.clara.net/sisa/correl.htm>.

Auch hier zeigt sich nur in einem von sechs Korrelationsvergleichen ein signifikanter Unterschied. D.h. die online erhobenen Daten sind nahezu identisch mit denen, die auf schriftlichem Weg erhoben wurden.

Fazit

Verglichen wurden die online und schriftlich erhobenen Daten einer Zufallsstichprobe von Internetnutzern mit denen einer Zufallsstichprobe der Allgemeinbevölkerung. Zwischen beiden Gruppen - und damit einhergehend zwischen den beiden Erhebungsmodi - sind in der überwiegenden Zahl der Ergebnisse deutliche Unterschiede zu erkennen. Eine Anpassungsgewichtung der Stichprobe der Internetnutzer an die Bevölkerungsstichprobe auf der Basis soziodemografischer Merkmale (Alter, Bildung, Geschlecht) ist angesichts der gegenwärtigen Internetpenetration und der besonderen Strukturmerkmale der Internetnutzer wenig erfolgversprechend. Werden dagegen die Befragtengruppen weitgehend parallelisiert (wie z.B. im vorliegenden Fall über einen höheren Bildungsabschluss), so zeigen die Ergebnisse unter beiden Erhebungsmodi eine überraschend hohe Übereinstimmung. Mit anderen Worten: Sofern sich das Erkenntnisinteresse bei der ISSP Umfrage auf die Einstellungen derjenigen beschränkt hätte, die über einen höheren Bildungsabschluss verfügen, so hätte eine solche Zielsetzung auch unter Rückgriff auf ein zufallsrekrutiertes Online-Panel in einem Web Survey weitgehend realisiert werden können.

Deshalb sehen wir in der vorliegenden Studie einen Hinweis darauf, dass für per Web befragte Populationen mit hoher Internetpenetration ähnliche Ergebnisse zu erwarten sind wie für schriftlich Befragte der selben Grundgesamtheit. Ein Beispiel wären z.B. Befragungen unter Jugendlichen. Für diese Gruppe, und hierbei vor allem Gymnasiasten, gehört die Nutzung des Internet schon jetzt mehrheitlich zum Alltag. So weist eine Studie des Medienpädagogischen Forschungsverbundes Südwest (2000: 37) aus, daß zur Jahresmitte 2000 der Anteil der Gymnasiasten mit Internet-Erfahrung schon bei 66 Prozent lag (bei den Haupt- und Realschülern waren es 45 bzw. 53 Prozent). Es ist davon auszugehen, daß die Internet-Nutzung bei bestimmten Subgruppen, insbesondere Jugendlichen und Personen mit höheren Bildungsabschlüssen, weiterhin stark ansteigen wird - diese und auch andere Gruppen der Gesellschaft also zunehmend leichter erreicht und damit auch als Befragungsteilnehmer angesprochen werden können. Befragungen der Allgemeinbevölkerung auf der Basis von Zufallsstichproben der Internetnutzer sind dagegen derzeit sicherlich noch verfrüht.

Literatur

Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM)/Arbeitsgemeinschaft Sozialwissenschaftlicher Institute (ASI)/Berufsverband Deutscher Markt- und Sozialforscher (BVM)/Deutsche Gesellschaft für Online Forschung (D.G.O.F.) (Hrsg.), 2001: Standards zur Qualitätssicherung für Online-Befragungen. Frankfurt: ADM.

Bandilla, W., 1999: WWW-Umfragen. Eine alternative Datenerhebungstechnik für die empirische Sozialforschung? S. 6-17 in: B. Batinic/A. Werner/L. Gräf/W. Bandilla (Hrsg.): Online-Research. Methoden, Anwendungen und Ergebnisse. Göttingen: Hogrefe.

Bandilla, W./Bosnjak, M., 2000: Online Surveys als Herausforderung für die Umfrageforschung: Chancen und Probleme. S. 9-28 in: P. Ph. Mohler/P. Lüttinger (Hrsg.): Querschnitt. Festschrift für Max Kaase. Mannheim: ZUMA.

Batinic, B./Bosnjak, M., 2000: Fragebogenuntersuchungen im Internet. S. 287-318 in: B. Batinic (Hrsg.): Internet für Psychologen. Göttingen: Hogrefe (2. Aufl.).

Birnbaum, M.H., 2001: Introduction to Behavioral Research on the Internet. Upper Saddle River, NJ: Prentice-Hall.

Bortz, J./Döring, N., 1995: Forschungsmethoden und Evaluation für Sozialwissenschaftler. Berlin, Heidelberg: Springer (2. Auflage).

Couper, M.P., 2000: Web Surveys. A Review of Issues and Approaches. Public Opinion Quarterly 64: 464-494.

Dillman, D.A., 2000: Mail and Internet Surveys – The Tailored Design Method. New York: Wiley.

Feierabend, S./Klingler, W., 2001: Kinder und Medien 2000: PC/Internet gewinnen an Bedeutung. Media Perspektiven 7: 345-357.

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

forsa, 2001: forsa omniNet - Next Generation Research, Berlin: forsa Gesellschaft für Sozialforschung und statistische Analyse mbH.

Gabler, S., 1994: ALLBUS-Baseline-Studie 1991 und ALLBUS 1992: Ost-West-Gewichtung der Daten. ZUMA Nachrichten 35: 77-81.

Hoppe, M., 2000: Aufbau und Organisation eines Access-Panels. S. 145-165 in: Statistisches Bundesamt (Hrsg.): Neue Erhebungsinstrumente und Methodeneffekte. Band 15 der Schriftenreihe Spektrum Bundesstatistik. Stuttgart: Metzler-Poeschel.

Koch, A./Wasmer, M./Harkness, J./Scholz, E., 2001: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS 2000). ZUMA Methodenbericht 2001/05.

Medienpädagogischer Forschungsverbund Südwest, 2000: JIM 99/2000 – Jugend, Information, (Multi-)Media. Baden-Baden: SWR Medienforschung.

Reips, U.-D., 2001: The Web Experimental Psychology Lab: Five years of data collection on the Internet. *Behavioral Research Methods, Instruments, and Computers* 33: 201-211.

Reips, U.-D./Bosnjak, M. (Eds.), 2001: *Dimensions of Internet Science*. Lengerich: Papst.

Schnell, R./Hill, P.B./Esser, E., 1995: *Methoden der empirischen Sozialforschung*, München/Wien: Oldenbourg (5. Auflage).

Theobald, A./Dreyer, M./Starsetzki, T. (Hrsg.), 2001: *Online-Marktforschung – Beiträge aus Wissenschaft und Praxis*. Wiesbaden: Gabler.

Van Eimeren, B./Gerhard, H./Frees, B., 2001: ARD/ZDF-Online-Studie 2001: Internetnutzung stark zweckgebunden. *Media Perspektiven* 8: 382-397.

Witherspoon, S./Mohler, P./Harkness, J., 1995: Report on Research into Environmental Attitudes and Perceptions, “European Consortium for Comparative Social Surveys, March, 1995.

RELIABILITÄT VON RETROSPEKTIVEN BERUFSVERLAUFSDATEN*

EIN VERGLEICH ZWISCHEN DER PRIVATWIRTSCHAFT UND DEM ÖFFENTLICHEN DIENST ANHAND VON PANELDATEN

ROLF BECKER

Im vorliegenden Beitrag wird die Zuverlässigkeit retrospektiver Angaben der Befragten zu beruflichen Tätigkeiten in ihrem Berufsverlauf untersucht. Insbesondere ist ausgehend von einem Befund von De Graaf und Wegener (1989) zu klären, ob generell Beschäftigte des öffentlichen Dienstes ihren Berufsverlauf unzuverlässiger erinnern als Staatsbeschäftigte. Es wird davon ausgegangen, dass gerade Beamte aufgrund der institutionellen Besonderheiten ihrer Beschäftigung größere Schwierigkeiten haben, ihren Berufsverlauf konsistent zu rekonstruieren als andere Befragte. Empirische Analysen von Panel-Daten erhärten diese Vermutung, was die Anzahl der beruflichen Tätigkeiten anbelangt. Jedoch machen Beamte zu anderen Attributen ihres Berufsverlaufs ebenso zuverlässige Angaben wie andere Befragte auch. Diese beamtenspezifischen Erinnerungsprobleme sind bei zukünftigen Erhebungen mittels ereignisorientierter Befragungsinstrumente zu berücksichtigen.

The study investigates the reliability of retrospective data respondents provided about their occupational careers. De Graaf and Wegener (1989) found that respondents employed in the public sector provided less reliable answers about their career than did employees in the private sector. The institutional context in which civil servants operate may make it more difficult for them to reconstruct the course of their career and the number of different jobs they have held. Empirical analyses of panel data confirm this hypothesis with respect to the number of jobs held. At the same time, data from respondents in the civil service about other aspects of their occupational career are as accurate and reliable as data from respondents in the private

* Für hilfreiche Kommentare danke ich Erika Brückner, Nan Dirk de Graaf, Karl Ulrich Mayer, Klaus Schömann und den anonymen Gutachtern. Besonderer Dank gilt Bernd Wegener, der mir seine Daten für eine Reanalyse zur Verfügung stellte.

sector. The problem of recall with civil servants thus seems to relate only to the number of jobs held in the course of their occupational career. This should be taken into account by researchers collecting event-history data about on the occupational careers.

1. Einleitung

Seit Ende der 1960er Jahre haben in der Sozialstrukturanalyse und insbesondere in der soziologischen Mobilitätsforschung die Längsschnittstudien mit Retrospektivdaten zugenommen (Blum et. al 1969; Ramsøy 1973; Mayer/ Brückner 1989; Mayer/Tuma 1990; Brückner/Mayer 1995). Vor allem seit den 1980er Jahren ist ein deutlich gewachsenes Interesse an Bildungs- und Berufsverläufen zu verzeichnen, die anhand von retrospektiven Ereignisdaten untersucht werden (Mayer/Papastefanou 1983; Tuma 1985; Carroll/Mayer 1986; Mayer/ Carroll 1987; Carroll/Mosakowski 1988; Blossfeld/Mayer 1988; Allmendinger 1989; Blossfeld 1989; Blossfeld/Mayer 1991; Becker 1993; Schömann 1994; Lauterbach 1994; Becker/Schömann 1996; Konietzka 1999). Um den Berufsverlauf adäquat im Längsschnitt untersuchen zu können, ist es notwendig, diesen als eine Sequenz von beruflichen Tätigkeiten (*Jobs*) zu erfassen. In der technischen Erhebung und der Datenauswertung werden diese Sequenzen in einer ereignisorientierten Datenstruktur abgespeichert (Blossfeld et al. 1986: 20-21). Eine Voraussetzung hierfür ist nicht nur die vollständige retrospektive Rekonstruktion des gesamten Berufsverlaufs, sondern auch die exakte Erinnerung der zeitlichen Datierung und logischen Abfolge von Tätigkeitsepisoden mit all ihren interessierenden Merkmalen (wie etwa Einkommen, Wirtschaftszweig, Betriebsgröße, berufliche Stellung, Tätigkeits- und Berufsbezeichnung, Arbeitszeit, etc.). Um unverzerrte und aussagekräftige Ergebnisse mit solchen Retrospektivdaten erzielen zu können, muss neben Objektivität und Validität auch die *Reliabilität* dieser Ereignisdaten sichergestellt sein (vgl. Klein/Fischer-Kerli 2000; De Graaf und Wegener 1989; Papastefanou 1980).¹

1 Neben der Objektivität und Gültigkeit ist die Zuverlässigkeit (Reliabilität) ein wichtiges Gütekriterium für die Messung von interessierenden Sachverhalten und der mit bestimmten Messinstrumentarien erhobenen Daten. Im allgemeinen wird mit der Reliabilität eines Messinstrumentariums ein Maß für die Reproduzierbarkeit von Messergebnissen — das Zusammenwirken von Stabilität und Genauigkeit von Messungen bei wiederholten Messungen — definiert (vgl. Diekmann 1996: 217) und stellt eine Voraussetzung für die Gültigkeit von Daten dar (vgl. Friedrichs 1973). D.h. mit dem selben Instrument sollten sich die Befunde mit anderen Daten reproduzieren lassen, sofern sie nicht durch systematische Messfehler verzerrt werden. Gerade bei Ereignisdaten ist die Zuverlässigkeit der Angaben

Jedoch stellt die Erhebung solcher quantitativen Ereignisdaten erhebliche Anforderungen an Erhebungsinstrumentarien, an die Interviewer und schließlich an die Befragten selbst, wenn sie ihren gesamten Lebenslauf oder Teilbereiche ihrer Lebensspanne lückenlos mit detaillierten Zeitangaben und exakten Sequenzabfolgen rekonstruieren sollen (vgl. Brückner 1990: 388-391; Pearson et al. 1992: 67; Brückner 1995: 68). (1) Um zuverlässige Angaben über Berufsverläufe zu erhalten, so eine methodisch bewährte Vorgehensweise der deutschen Lebensverlaufsstudie, muss das *Erhebungsinstrumentarium* wie etwa der Fragebogen von vornherein die Befragten in die Situation zu Beginn des Prozesses versetzen (vgl. Tölke 1980). So hat es sich aus kognitionspsychologischen Gründen als vorteilhaft erwiesen, den gesamten Berufsverlauf separat von anderen Teilbereichen des individuellen Lebenslaufs zu messen, und den Berufsverlauf selbst wiederum — unterteilt nach den einzelnen Sequenzen der Erwerbsgeschichte — in seiner logischen wie chronologischen Reihenfolge zu erheben (vgl. Dex/Culloch 1998: 498; Sudman et al. 1996: 168; Pearson et al. 1992: 67; Papastefanou 1980; Linton 1978). Diese Erhebungsweise, die beispielsweise bei der deutschen Lebensverlaufsstudie vorgenommen wird, erleichtert nicht nur die Wiedergabe einzelner Episoden, sondern entspricht auch dem kognitiven Prozess der Kodierung und Speicherung solcher Zustände und Ereignisse im Berufsverlauf (vgl. Sudman et al. 1996; Huttenlocher et al. 1988). So präferieren Personen die Erinnerung von Ereignissen in der chronologischen Reihenfolge. Befunde von Experimenten belegen aber auch, dass die Abfrage in umgekehrter Reihenfolge von Gegenwart zu Vergangenheit (backward recall) zu besseren Ergebnissen führen kann. Ebenso scheinen offene Fragen ungünstiger als geschlossene, einfach formulierte Fragen zu sein, die Hinweise mit Signalcharakter (*cues*) für den interessierenden und zu erinnernden Sachverhalt beinhalten (vgl. Schwarz et al. 1994: 191-192).² Zudem sollten „die Informationen in einem Kontext abgefragt werden, der die nötigen Erinnerungsleistungen unterstützt“ (Mayer/Papastefanou 1983: 119). Dabei können entsprechende Hinweisreize bei der Fragestellung hilfreich sein. Offensichtlich sind Erinnerungsleistungen desto besser, je spezifischer die

bedeutsam, weil es „fatal wäre, wenn die rekonstruierten Prozesse und Zustandsveränderungen gedächtnispsychologische Artefakte darstellen würden“ (Papastefanou 1980: 41).

² Fehlende Angaben zu einzelnen Items, die auf dem Schwierigkeitsgrad oder Unverständlichkeit der Fragen beruhen, können zu systematischen Folgefehlern im Erinnerungsprozess führen (vgl. Peters 1988). Diese Teilausfälle, die in der Folge zu weiteren unzuverlässigen Angaben oder gar Ausfällen kompletter Ereignissequenzen und Verläufen führen können, sind nicht unerheblich, wenn es darum geht, Berufsverläufe möglichst genau und vollständig zu erfassen. Des weiteren sind dann systematische Verzerrungen bei den statistischen Auswertungen nicht auszuschließen (vgl. Little 1992).

Hinweisreize den Informationen nahe kommen, die beim Encoding eine Rolle gespielt haben. Wenn aber entsprechende Zeit- und Raumhinweise fehlen, dann entstehen Probleme bei der korrekten Erinnerung von Ereignissen (vgl. Sudman et al. 1996), so dass oftmals relevante Episoden kaum erinnert werden können (vgl. Schwarz 1996: 55; Brückner 1990). Darum sollte das Erhebungsinstrumentarium so gestaltet sein, dass die befragte Person nicht nur ausreichend Zeit zum Antworten, sondern auch zum Vergleichen hat.³ Hinweise auf herausragende historische Ereignisse und Perioden können dabei ebenso hilfreich sein (vgl. Dex/McCulloch 1998: 498) wie die Nutzung persönlicher Dokumente (vgl. Brückner 1990). Aber gleichzeitig darf das einzelne Interview nicht zu lange andauern, weil mit zunehmender Interviewdauer die Aufmerksamkeit der am Interview beteiligten Personen schwindet. Unter der nachlassenden Aufmerksamkeit leidet die Zuverlässigkeit der Antworten (Sudman/Bradburn 1974). Andererseits kann die Dauer des Interviews mit der Komplexität des Berufsverlaufs zusammenhängen: je komplexer der Berufsverlauf ist, desto mehr Zeit wird für das Interview aufzuwenden sein, desto mehr Anforderungen werden an die Erinnerungskapazität der befragten Person gestellt und damit werden Erinnerungsprobleme und daraus resultierende Fehler wahrscheinlicher. Die Festlegung mehrerer zeitlich begrenzter Interviewsituationen kann dieses Problem minimieren (vgl. Brückner 1990).

(2) Des Weiteren kann der *Interviewer* selbst (un)bewusst zur Zuverlässigkeit der erinnerten Angaben beitragen. Brückner (1995) zufolge können Interviewer bei der Erinnerung von Berufsverläufen hilfreich sein, aber auch die befragte Person durch die Ereignissequenzen jagen, so dass mit dem Verpassen von Antworten oder Auslassen von Fragen die vermeintliche Zeitersparnis mit unzuverlässigeren Angaben erkaufte wird. Auch beeinflussen Merkmale des Interviews (etwa Tonbandaufzeichnung oder Anzahl vorhergehender Interviews), des Interviewers (etwa Dauer der Arbeit von Interviewern, ihr Bildungsniveau oder Alter) die vom Befragten berichtete und vom Interviewer festgehaltene Anzahl von Tätigkeitsepisoden im Berufsverlauf. Jedoch können diese Fehler durch eine intensive Datenedition ausgeglichen werden. Dabei auftretende Fehler können die Zuverlässigkeit von Daten ebenso beeinträchtigen (Brückner 1995: 68).

³ Möglicherweise geben Individuen, die bereits über ihren Berufsverlauf nachgedacht haben oder während der Befragung reflektieren konnten, genauere Angaben. So berichten Bradburn et al. (1987), dass „increasing the amount of time for a response can affect the strategy a respondent uses, as well as the accuracy of the resulting answer“ (S. 158).

(3) Schließlich ist die intellektuelle und altersabhängige Fähigkeit befragter Personen, ihren Berufsverlauf erinnern zu können, eine wichtige Voraussetzung für die Datenerhebung in der Lebensverlaufsforschung. Systematische Erinnerungsfehler wie etwa der „telescoping effect“ (Sudman et al. 1996) oder das „Vergessen“ (Bradburn et al. 1987), selektive oder heuristische Rekonstruktion von Verläufen in der Retrospektive (Tversky/Kahneman 1986) oder die Anpassung zurückliegender Ereignisse an den gegenwärtigen Zustand sind solche kognitiven Einflüsse, die als Zweifel an der Zuverlässigkeit von Retrospektivangaben angeführt werden (vgl. Friedrichs 1973: 206).⁴ So seien die Daten ungenauer, je länger die Ereignisse zurückliegen (Jahoda et al. 1951: 161; Goldfarb 1960: 45), und sie würden nur bei sozial bedeutsamen Ereignissen zuverlässig wiedergegeben (vgl. Papastefanou 1980: 14), während sie eher vergessen werden, je bedrohlicher das Ereignis war (Friedrichs 1973: 206).

Vor allem wegen dieser Vorbehalte gegenüber den Erinnerungsleistungen der befragten Personen wird die Zuverlässigkeit von Lebensverlaufsdaten oftmals bezweifelt (vgl. Rendtel 1995: 12) und den Panel- oder Querschnittsinformationen über gegenwärtige Merkmale oder Situationen eine größere Zuverlässigkeit zugeschrieben (Brüderl 1990). Aber „es gibt insgesamt nur sehr wenige methodische Studien zur Zuverlässigkeit retrospektiv erhobener Daten. Die allgemeinen Vorbehalte in Bezug auf die höhere Fehlerhaftigkeit solcher Daten sind daher ebenso plausibel wie empirisch ungesichert“ (Mayer/Papastefanou 1983: 108). Diese Aussage von Mayer und Papastefanou aus dem Jahre 1983 kann heute noch aufrechterhalten werden, obgleich in der Zwischenzeit unterschiedliche Studien vorgelegt wurden, die für die Zuverlässigkeit retrospektiv erhobener Längsschnittdaten und damit für das ereignisorientierte Erhebungsdesign der Lebensverlaufsstudie am Berliner Max-Planck-Institut für Bildungsforschung sprechen (vgl. Blossfeld 1989; Brückner 1990; Schömann 1994).

⁴ So können Personen — wenn sie annehmen, es hätte sich in ihrem Lebenslauf kaum etwas getan — ihre Vergangenheit durch Attribute der Gegenwart rekonstruieren. Ansonsten können sie Theorien über den Wandel heranziehen, um sich an ihre Vergangenheit zu erinnern. So sind die Erinnerungen verzerrt, wenn sich Merkmale verändert haben, aber der Wandel von der befragten Person nicht wahrgenommen wurde. Die Befragten können auch die Natur des Wandels fehleinschätzen oder von einem Wandel berichten, obwohl keiner stattgefunden hat (Pearson et al. 1992: 53).

Ende der achtziger Jahre untersuchten De Graaf und Wegener (1989) anhand eines Drei-Wellen-Panels des Projektes „Prozesse der Berufsfindung“ in systematischer Weise die Reliabilität von retrospektiv erhobenen Berufsverlaufsdaten.⁵ Dabei stellten sie mit Hilfe logistischer Regressionen fest, dass vor allem Beschäftigte im öffentlichen Dienst inkonsistentere Angaben über ihre Tätigkeitsepisoden machen. Diese Inkonsistenz äußert sich darin, dass bei den Beschäftigten im öffentlichen Dienst die *Anzahl der erinnerten Tätigkeiten im Berufsverlauf* von Befragung zu Befragung stärker variiert, als bei den Beschäftigten in der Privatwirtschaft (De Graaf/Wegener 1989: 21). Demnach erinnern Beschäftigte in der Privatwirtschaft die einzelnen Tätigkeitsepisoden weitaus zuverlässiger als Staatsbeschäftigte. „Although one could speculate that individuals working in the public sector are able to give more secure answers, this does definitively not hold for the number of job spells. A post-hoc explanation for our result is that it is for individuals working in the public sector less clear whether they get a new job or just make a small move within the firm.“ (De Graaf/Wegener 1989: 22).

Dieser selektive Befund von De Graaf und Wegener (1989) ist nicht nur erklärungsbedürftig, sondern hat möglicherweise Konsequenzen für die Konstruktion der Erhebungsinstrumente und für die Datenanalyse und -interpretation. Unklar ist zunächst, warum die Angaben von Staatsbeschäftigten zu ihrer beruflichen Karriere generell unzuverlässiger sein sollen. Wenn es hierfür keine erhebungstechnischen Gründe gibt, dann würde sich die Frage stellen, ob es überhaupt Sinn macht, Berufsverläufe im öffentlichen Dienst zu untersuchen (vgl. Becker 1993). Wenn Staatsbeschäftigte jedoch aufgrund besonderer Beschäftigungsverhältnisse im öffentlichen Dienst weniger in der Lage sind, Tätigkeitswechsel und unbedeutende berufliche Veränderungen zu diskriminieren, dann — so die grundlegende Annahme — dürfte dies eher für Beamte als für Arbeiter und Angestellte im öffentlichen Dienst zutreffen (vgl. Becker 1993). Entweder sind die bislang verwendeten

5 Trotz methodischer Vorbehalte gegenüber Panel (wie etwa unnatürliche Panelmortalität bzw. Abnutzung der Ausgangsstichprobe, Design- und Datenerhebungsprobleme, Probleme der wiederholten Messung und Selektivität (Rendtel 1995)) bietet solch ein Datensatz die Möglichkeit erklärende Faktoren für mögliche mangelnde Zuverlässigkeit solcher Daten zu isolieren. In gewisser Hinsicht sind Paneldaten geeignet, um die Zuverlässigkeit retrospektiver Lebensverlaufsdaten zu untersuchen, wenn die gleichen Fragen wiederum nach einiger Zeit präsentiert werden. Allerdings bergen häufige Erinnerungen die Gefahr in sich, dass zurückliegende Ereignisse modifiziert, angepasst und reinterpretiert werden. Die „Glättung zu einem normalen Lebensverlauf“ oder „Rekonstruktion zu logisch konsistenten Ereignissequenzen und Episoden“ sind Beispiele dafür (Sudman et al. 1996). Andererseits werden Ereignisse und Sequenzen durch wiederholte Erzählung (rehearsal) eher im Langzeitgedächtnis gespeichert und später zuverlässiger erinnert (Sudman et al. 1996: 174).

Erhebungsinstrumente für Beschäftigte im öffentlichen Dienst unzulänglich, um zuverlässige Informationen über ihren Berufsverlauf zu erheben, wenn die Erhebungseinheit die berufliche Tätigkeit (Zustand) und ihre Veränderung (Ereignis) sind. Aufgrund der Organisation ihrer spezifischen Ausbildung und Beschäftigung in einem Funktionsbereich, wo sie ein Bündel diverser Tätigkeiten am Arbeitsplatz ausfüllen sollen, dürften demnach Beamte größere Schwierigkeiten haben in der Retrospektive, *Tätigkeitswechsel* — also Veränderungen beruflicher Tätigkeiten, die auch mit Änderungen im sogenannten ISCO-Code (International Standard Classification of Occupations) einhergehen — voneinander abzugrenzen. Dies findet dann seinen Niederschlag in der geringeren Reliabilität für die Anzahl ihrer Tätigkeiten im Berufsverlauf, wenn sie diese wieder erinnern und rekonstruieren müssen.

Oder der Befund von De Graaf und Wegener (1989) beruht auf einem statistischen Artefakt infolge unbeobachteter Heterogenität, die bei den Regressions-schätzungen unberücksichtigt blieb. So muss möglicherweise bei den Beschäftigten im öffentlichen Dienst wegen unterschiedlicher Laufbahnstrukturen, Tätigkeiten und Beschäftigungsbereiche in besonderer Weise zwischen den einzelnen Beschäftigtengruppen differenziert werden, wenn Ausmaß und Bestimmungsgründe für die Reliabilität von Berufsverlaufsdaten zu beurteilen sind (vgl. Becker 1993). Es ist wiederum zu vermuten, dass es aufgrund institutioneller und struktureller Besonderheiten des öffentlichen Dienstes gerade die *Beamten* sind, die in einem höheren Maße inkonsistente Angaben bei der Erinnerung von Tätigkeitsepisoden machen als andere Staatsbeschäftigte oder Beschäftigte in der Privatwirtschaft.

Um die Annahmen empirisch zu überprüfen, wird in diesem Beitrag mit den von De Graaf und Wegener (1989) verwendeten Panel-Daten eine Replikation und Reanalyse mit entsprechenden Erweiterungen der von ihnen spezifizierten Modellschätzungen vorgenommen. Im folgenden Abschnitt werden sowohl institutionelle Besonderheiten der Beschäftigung und beruflichen Karriere im öffentlichen Dienst als auch kognitionspsychologische Grundlagen der Erinnerung von Berufsverläufen diskutiert sowie Verbindungen zwischen diesen beiden Aspekten hergestellt. Im dritten Abschnitt wird die Datenbasis diskutiert und im darauffolgenden Abschnitt die verwendeten Reliabilitätsmaße erläutert. Die Darstellung deskriptiver Befunde erfolgt im fünften Teil und im sechsten Abschnitt die Interpretation zentraler Ergebnisse der multivariaten Analyse. Zusammenfassung und Schlussfolgerung bilden den Abschluss des vorliegenden Beitrages.

2. Theoretische Implikationen für die Reliabilität von Retrospektivdaten mit besonderer Berücksichtigung des öffentlichen Dienstes

Bei der Erklärung von fehlerhafter Erinnerung (recall bias) werden diverse Gründe angeführt, die mitunter durch kognitionspsychologische Modelle und empirische Befunde abgestützt sind (vgl. Sudman et al. 1996). Es besteht folgende Ansicht: Je länger ein anvisiertes Ereignis zeitlich zurückliegt, desto wahrscheinlicher sind unzuverlässige Antworten bei Retrospektivbefragungen (Friedrichs 1973: 206; Tölke 1980: 35; vgl. Bradburn et al. 1987). Selbst wenn den Befragten die einzelnen Episoden über Ereignisse zur Verfügung stehen, werden ihre Angaben durch andere Einflüsse verzerrt, wie z.B. durch „telescoping bias“ oder durch Versäumnisse („Vergessen“). Je länger die Referenzperiode ist, desto eher werden Episoden vergessen oder desto wahrscheinlicher sind systematische Fehlplatzierungen von früheren Ereignissen in die jüngste Vergangenheit (forward telescoping). Beim „Vergessen“ werden weniger Ereignisse berichtet als tatsächlich aufgetreten sind. Infolge von „telescoping“ werden dagegen eher zu viele Ereignisse vom Respondenten berichtet (vgl. Strube 1987). Im hierarchisch organisierten episodischen Langzeitgedächtnis, das ereignisbezogene und autobiographische Wissensbestände enthält, nehmen die Feinheiten der zeitlichen Wissensseinheiten bzgl. eines Ereignisses ab, je mehr Zeit verstreicht. Weiterhin bedeutet der Ersatz von Monatsangaben durch Jahres(zeit)angaben einen Verlust exakter Informationen über Zeitpunkte (vgl. Bradburn et al. 1994): Je größer die zurückliegende Zeitperiode ist, desto größer wird die Standardabweichung der berichteten Zeitpunkte. Personen speichern die Datumsangaben zwar in der Regel akkurat im Gedächtnis, aber mit fortschreitender Zeit überwiegen die gröberen Zeitreferenzen über die feineren. Schließlich werden die berichteten Angaben ungenauer, aber sie sind nicht systematisch verzerrt (Huttenlocher et al. 1988: 472). Auch in der Erinnerung und Wiedergabe werden die Zeitangaben bei längeren Referenzperioden zunehmend gröber, und es besteht oftmals auch die Gefahr, dass dann nicht alle Informationen zu zurückliegenden Ereignissen vom Gedächtnis freigegeben werden. Und schließlich kann die Länge des zu erinnernden Zeitintervalls den Grad der Spezifität der Informationen und damit ihre Genauigkeit und Zuverlässigkeit beeinträchtigen.

Darüber hinaus werden offensichtlich mehr Ereignisse im ersten Interview als in späteren Interviews genannt (Sudman et al. 1996: 175; Huttenlocher et al. 1988: 472). Gerade solche „telescoping effects“ beeinträchtigen die Zuverlässigkeit von

Retrospektivdaten (Strube/Weinert 1987: 161; Pearson et al. 1992: 83). Diese Gründe für Erinnerungsfehler alleine erklären die Unterschiede in der Reliabilität von Berufsverlaufsdaten zwischen den Staatsbeschäftigten und privatwirtschaftlich Beschäftigten nicht. Dagegen sprechen zunächst empirische Befunde über die Berufsverläufe im öffentlichen Dienst und in der Privatwirtschaft (vgl. Becker 1993). Staatsbeschäftigte der um 1930, 1940 und 1950 geborenen Kohorten haben im Durchschnitt weniger Tätigkeitswechsel als Beschäftigte in der Privatwirtschaft und verweilten auch länger in den einzelnen Jobs. Bei vergleichbaren Referenzperioden weisen Beschäftigte im öffentlichen Dienst, und dabei vor allem auch Beamte, weitaus stetigere und weniger komplexe Berufsverläufe als Beschäftigte im privatwirtschaftlichen Sektor auf. Wenn aber dynamische, also sich ändernde Variablen die Stabilität der Informationen im Gedächtnis erschweren, dann erschweren sie sowohl ihre Speicherung als auch die exakte Erinnerung. Sind diese Informationen aber stabil, dann ist der Zugriff auf diese Informationen besser, d.h. das retrospektive Datum ist gültiger und auch zuverlässiger (vgl. Papastefanou 1980: 15-16). Daher müssten für Staatsbeschäftigte eher reliable Angaben zu ihrem Berufsverlauf und der erinnerten Anzahl einzelner Episoden zu erwarten sein als für die in der Privatwirtschaft Beschäftigten (vgl. De Graaf/Wegener 1989).⁶

Als Einwand könnte die geringere Zuverlässigkeit der Berufsverlaufsdaten von Staatsbeschäftigten selbst angeführt werden. Jedoch ergaben zusätzliche Analysen mit den Daten der deutschen Lebensverlaufsstudie, dass die Interviewer die Angaben von Staatsbeschäftigten und von diesen wiederum die der Beamten als zuverlässiger einstufen. Und schließlich könnte man als weiteres kognitionspsychologisches Argument anführen, dass dies deswegen so sei, weil Personen für selten auftretende, sowie positiv affektiv besetzte Ereignisse gute Erinnerungsleistungen aufweisen (Sudman/Bradbury 1974), während negativ besetzte Ereignisse rasch wieder vergessen werden (Friedrichs 1973: 206). Andererseits wird die Häufigkeit seltener Ereignisse überschätzt, während die Zahl häufiger Ereignisse unterschätzt wird (vgl. Strube 1987: 94). Die diesbezüglichen empirischen Befunde aus der Lebensverlaufsstudie sind ambivalent. Zwar treten berufliche Ereignisse im öffentlichen Dienst, wie etwa Tätigkeits- oder Arbeitgeberwechsel,

⁶ Dafür spricht auch, dass im öffentlichen Dienst eher Personen mit höheren Bildungsniveaus beschäftigt werden als in der Privatwirtschaft (vgl. Blossfeld/Becker 1989). Ob dies jedoch mit sektorspezifischen Unterschieden in der intellektuellen Fähigkeit niederschlägt, die berufliche Karriere vollständig erinnern zu können, wäre empirisch zu überprüfen.

seltener auf als in der Privatwirtschaft, aber gleichzeitig erfahren Staatsbeschäftigte (und dabei vor allem Beamte) seltener berufliche Auf- und Abstiege als privatwirtschaftlich Beschäftigte (vgl. Becker 1993; Mayer/Carroll 1987). Da wir nicht wissen, ob sich die Beschäftigten in den beiden Sektoren in der subjektiven Bewertung solcher Ereignisse unterscheiden, die für einzelne Personen möglicherweise zentrale Lebensereignisse darstellen, können wir auch keine schlüssigen Vermutungen über ihre Auswirkungen auf die Reliabilität der Aussagen anstellen. Aber man könnte durchaus annehmen, dass bei den Staatsbeschäftigten im Vergleich zu den Beschäftigten in der Privatwirtschaft die einzelnen Tätigkeitswechsel im Berufsverlauf unbedeutender sind, die nur dann erinnert werden, wenn sie in einem kurzen Zeitintervall zwischen Ereignis und Interview liegen (Goldfarb 1960: 45).

Als einen weiteren Grund für die Unterschiede in der Reliabilität zwischen Beschäftigten im öffentlichen Dienst und in der Privatwirtschaft könnten die unterschiedlichen institutionellen Regelungen der Beschäftigungsverhältnisse und der damit verbundenen beruflichen Karriere angeführt werden. So betont beispielsweise Papastefanou (1980) den Institutionalierungsgrad des sozialen Kontextes durch rechtliche und soziale Normen, in denen Informationen bei Retrospektivbefragungen indiziert werden: „Je stärker institutionalisiert ein Interaktionskontext von Informationen ist, umso mehr werden diesbezügliche Informationen im Langzeitgedächtnis gespeichert. Die gute Speicherung im Langzeitgedächtnis bedeutet aber — ceteris paribus — eine gute Validität retrospektiv gegebener Informationen“ (Papastefanou 1980: 12). Dieser Zusammenhang dürfte auch für die Reliabilität retrospektiver Informationen zutreffen (vgl. Sudman et al. 1996; Sudman/Bradburn 1974). So besteht im öffentlichen Dienst ein gesetzlich regulierter Dualismus des Rechtsstatus der Arbeitnehmer: Beamte stehen nach den „hergebrachten Grundsätzen des Berufsbeamtentums“ (Art. 33, Abs. 2 GG) im öffentlich-rechtlichen Dienst- und Treueverhältnis zum Arbeitgeber, während die Arbeiter und Angestellten im öffentlichen Dienst einen privatrechtlichen Arbeitnehmerstatus innehaben, der durch das Tarifrecht geregelt wird. Im Laufe der Zeit kam es zu einer Annäherung und Durchdringung dieser unterschiedlichen Beschäftigungsverhältnisse. In der Privatwirtschaft besteht keine entsprechende rechtliche, gesetzlich fixierte Differenzierung zwischen den Beschäftigten, sondern es werden sozialversicherungs- und tarifvertragsrechtliche Unterschiede gemacht. Ein zentrales Merkmal des öffentlichen Dienstes ist das Laufbahngruppensystem, das durch die Laufbahnverordnung und das Laufbahnprinzip reguliert wird. Diese Laufbahngruppen stellen ein idealtypisches System von exakt definierten und begrenzten Mobilitätsketten

dar, die im Unterschied zur Privatwirtschaft nicht an einen bestimmten Betrieb gebunden sind. Diese Laufbahnen umfassen jeweils ein Eingangsamtsamt und vier darauffolgende unterschiedlich besoldete Beförderungssämter, die gleichzeitig mit hierarchisch angeordneten Besoldungsgruppen verbunden sind. Diese Ämter stehen in keiner eindeutigen Beziehung zu den einzelnen Dienstposten bzw. Arbeitsplätzen und deren Anforderungsstruktur (Brinkmann 1976). Innerhalb dieser Laufbahngruppen sind die einzelnen Arbeitsplätze in einem hierarchischen Qualifikationsprofil angeordnet. Die Abfolgen von Beförderungen — als Kette einer Abfolge von höheren Positionen, verbunden mit höherem Grundgehalt, zunehmendem Prestige sowie mehr Weisungsbefugnissen — lassen sich als Karrieren definieren. Die Laufbahnen des öffentlichen Dienstes sind rechtlich geordnete Berufswege und entsprechen bei näherer Betrachtung einem besonders stark formalisierten Karrieremuster. Der Eintritt in eine Laufbahn und die sich daran anknüpfende Karriere ist weitaus transparenter und in einem höheren Grade formalisiert als in der Privatwirtschaft. Daher müsste entgegen den Befunden von De Graaf und Wegener (1989) erwartet werden, dass die Staatsbeschäftigten und vor allem die Beamten aufgrund der Eindeutigkeit und Transparenz von Beschäftigungsverhältnissen und Karrierewegen ihren Berufsverlauf exakt erinnern können. Zumindest dürften die Angaben der Staatsbeschäftigten genauso zuverlässig sein wie die der Beschäftigten im öffentlichen Dienst.

Für Beamte stellen Beförderungen eine besondere Form vertikaler Mobilität dar. Im Unterschied zu den Angestellten werden sie nicht nur für einen bestimmten Dienstposten, sondern für eine Laufbahn und damit in der Erwartung gewisser Positionsänderungen eingestellt. Aus dem Laufbahnprinzip des öffentlichen Dienst folgt für Beamte, dass die Bewerber im Vorbereitungsdienst (Anwärter) nicht für einen bestimmten Arbeitsplatz, sondern für verschiedene Dienstposten in einem Funktionsbereich mit einem weitgefächerten Aufgaben- und Tätigkeitsbündel innerhalb einer Laufbahn ausgebildet werden. Demgegenüber werden Angestellte und Arbeiter „dem Prinzip nach zur Erfüllung der Arbeitsaufgaben an einen bestimmten Arbeitsplatz oder Dienstposten eingestellt“ (Behr/Schultz-Wild 1973: 20). Der berufliche Aufstieg von Beamten, die hochgradig formalisierte und rechtlich fixierte Beförderung innerhalb der Laufbahn, kann mit einer Versetzung auf einen anderen Arbeitsplatz verbunden sein (Funktionsaufstieg), aber es kann auch das alte Amt bzw. die vorhergehende Tätigkeit beibehalten werden. „Beförderungen sind dem Prinzip nach ohne Arbeitsplatzwechsel möglich, andererseits bedeutet ein Wechsel auf Arbeitsplätze mit höheren Anforderungen oder mit einem breiteren Verantwortungsbereich nicht unbedingt ein entsprechendes Aufsteigen in der Personalstruktur“ (Behr/Schultz-Wild 1973: 23). Vor diesem Hintergrund ist es

plausibel, dass Beamte im Gegensatz zu anderen Arbeitnehmern im öffentlichen Dienst oder privatwirtschaftlich Beschäftigten größere Schwierigkeiten haben dürften, zeitlich zurückliegende Tätigkeitswechsel voneinander abzugrenzen, wenn sie nicht den Aufstieg in ein höheres Amt gleichzeitig als einen Tätigkeitswechsel begreifen. Somit könnte die größere Unsicherheit im temporären Gedächtnis von Beamten, die einzelnen beruflichen Tätigkeiten und Arbeitsplatzwechsel zu diskriminieren, dazu führen, dass sie ihren Berufsverlauf nicht so präzise wie die anderen Beschäftigtengruppen rekonstruieren. Daraus dürfte eine vergleichsweise geringere Zuverlässigkeit ihrer Retrospektivangaben resultieren, die der Forscher an der Anzahl erinnerter Tätigkeitsepisoden bemessen kann.

3. Beschreibung des Datensatzes

Für die Analyse der Reliabilität von Berufsverlaufsdaten wird auf den Datensatz des Drei-Wellen-Panels zurückgegriffen, den De Graaf und Wegener (1989) für eine Reanalyse zur Verfügung gestellt hatten. Diese Daten wurden im Rahmen des am Max-Planck-Institut für Bildungsforschung (Berlin) angesiedelten Projekts „Prozesse der Berufsfindung“ erhoben (vgl. Schneid 1986). Diese Studie wurde von Bernd Wegener initiiert und in Zusammenarbeit mit ZUMA (Zentrum für Umfragen, Methoden und Analysen) durchgeführt. Die Felderhebung erfolgte im Jahre 1987 in der Region Mannheim, Grünstadt und Schriesheim. Die Wiederbefragung fand zwischen September 1987 und November 1987 im Abstand von jeweils einem Monat statt. Weil es das Ziel des Forschungsdesigns war, solche Personen zu befragen, die mindestens einen Beruf hatten, wurden Personen im Alter zwischen 30 bis 34 (Kohorte 1953-57) bzw. 40 bis 44 (Kohorte 1943-47) Jahren interviewt (zu weiteren Details siehe Schneid 1989). Es wurden keine Maßnahmen ergriffen, um Erinnerungsfehler zu minimieren. Für den Vergleich der Daten der einzelnen Wellen miteinander wurden drei einzelne Datensätze gebildet. Daraus resultierte eine Reduktion der Fallzahlen: 1.) für den Vergleich der ersten mit der zweiten Welle um 18 Prozent (von 604 auf 497 Fälle) und 2.) für den Vergleich der ersten mit der dritten Welle bzw. für den Vergleich aller Wellen um 27 Prozent (von 604 auf 440 Fälle).

Die Panel-Mortalität ist durchaus beträchtlich, und könnte möglicherweise auch einen Einfluss auf das Gesamtergebnis haben, wenn sie unnatürlich und daher systematisch ist (vgl. Rendtel 1989). So können aus systematischen Panelabnutzungen verzerrte Ergebnisse für die Reliabilitätsanalyse entstehen, wenn davon Personengruppen betroffen sind, die ohnehin eher unzuverlässige Angaben zu ihrem Berufsverlauf machen. Explorativen Analysen zufolge ist dies für die

Differenzierung nach Sektorzugehörigkeit oder beruflicher Stellung nicht der Fall. Demnach liegen für keine der Berufsgruppen systematische Ausfälle in der Abfolge der einzelnen Panelwellen vor. Zweifelsohne können selektive Ausfälle mit Erinnerungsfehlern zusammenhängen und die Vergleichbarkeit der Wellen miteinander erschweren. Jedoch ist von einer „complete case analysis“ von Befragten abzusehen, die an allen Wellen teilgenommen haben. Denn Retrospektivdaten werden in der Regeln nicht wiederholt im Paneldesign erhoben, sondern in einer einmaligen Befragung. Deswegen ist es von Interesse, empirisch zu eruieren, ob die Abweichungen bei der Erinnerung von Berufsverläufen nicht nur vom Erhebungsinstrumentarium, sondern auch von der realisierten Stichprobe abhängen.

Bei dieser Befragung wurde eine berufliche Tätigkeit nur dann als Episode berücksichtigt, wenn sie mindestens sechs Monate andauerte. Diese Einschränkung wurde im Interview verdeutlicht. So lautet der exakte Wortlaut der Frage: „Ich möchte mich jetzt mit Ihnen darüber unterhalten, was Sie beruflich seit Ihrem Schulabgang alles gemacht haben. Es geht dabei nur um Beschäftigungen, die Sie mindestens sechs Monate ohne einen Wechsel des Unternehmens oder Arbeitgebers oder als Selbständiger hatten“. Um Verzerrungen durch Interviewer- und Vercodungsfehler zu minimieren, werden bei den folgenden Analysen diejenigen Episoden ausgeschlossen, die weniger als sechs Monate andauern. Daraus resultiert eine weitere Reduzierung der Fallzahl, welche schon durch fehlende Werte infolge von Interviewer- oder Vercodungsfehler verursacht wird. Des weiteren wurden auch mithelfende Familienangehörige von den Analysen ausgeschlossen. Schließlich haben De Graaf und Wegener (1989) aus Gründen der Komplexitätsreduzierung bei den einzelnen Wellenvergleichen nur maximal sechs Tätigkeitsepisoden pro Befragten berücksichtigt. Vermutlich sollen auf diesem Wege auch zufällige Schwankungen und „Ausreißer“ bei der Erinnerung der Tätigkeitsepisoden ausgeschlossen werden. Um die Vergleichbarkeit mit den Analysen von De Graaf und Wegener sicherzustellen, wurden diese Stichprobenauswahlen übernommen.

4. Reliabilitätsmaße

Als erstes wird die *Konsistenz der Angaben der Befragten* (CONSIS) gemessen. Damit wird der Prozentanteil der Entsprechung bzw. Nichtentsprechung der Antwort zu einer Variablen im Vergleich zu den anderen Wellen wiedergegeben.

Zweitens wird die durchschnittliche Abweichung von Angaben zur Tätigkeitsanzahl zwischen zwei Wellen gemessen:

$$\text{DEV} = \Sigma(X_{\text{Welle } k} - X_{\text{Welle } k+1}) / N$$

Dieser *Abweichungswert* DEV gibt einen Hinweis dafür, ob — durchschnittlich gesehen — Befragte in der darauffolgenden Welle mehr oder weniger Tätigkeitsepisoden erinnern. Zwar gibt dieser Koeffizient die strukturelle Abweichung zwischen zwei Wellen wieder, aber strukturelle Heterogenität der Stichprobe wird dadurch nicht kontrolliert. So könnten beispielsweise Frauen in systematischer Weise mehr Tätigkeiten als Männer angeben.

Ergänzend dazu wird als drittes Maß der durchschnittliche *absolute Abweichungswert* ABSDEV berechnet:

$$\text{ABSDEV} = \Sigma(\text{ABS}(X_{\text{Welle } k} - X_{\text{Welle } k+1})) / N$$

Damit sind Aussagen möglich, in welchem Umfang die Angaben zu den Tätigkeitsepisoden zwischen den Wellen voneinander abweichen.

Viertens wird die *durchschnittliche absolute Abweichung als ein prozentualer Anteil an der vorhergehenden Welle* berechnet:

$$\% \text{-ABDE} = ((\Sigma \text{ABS}(X_{\text{Welle } k} - X_{\text{Welle } k+1}) / X_{\text{Welle } k}) / N) \cdot 100 \text{ Prozent}$$

Hiermit wird die durchschnittliche absolute Abweichung der Angabe zu Variable X — die Tätigkeitsanzahl — von der darauffolgenden Welle berechnet und als prozentualer Anteil zur vorhergehenden Welle wiedergegeben. D.h. man erhält mit %-ABDE die gesamte absolute Abweichung einer Angabe in der darauffolgenden Welle in Berücksichtigung zu den Angaben der vorhergehenden Welle. Somit ist dieses Maß ein Indikator für eine „Fehlerrate“ der Angaben zu den Tätigkeitsepisoden in darauffolgenden Wellen. Schließlich wird das *Test-Retest-Maß* herangezogen. Es ist ein Produkt-Moment-Korrelationskoeffizient, der aus der Korrelation zwischen der Tätigkeitsanzahl aus der aktuellen Welle k mit der gleichen Variablen aus der darauffolgenden Welle k+1 resultiert. Über den Vergleich dieses Koeffizienten über die einzelnen Wellen hinweg kann aufgezeigt werden, ob sich die Zuverlässigkeit der Angaben zu den Tätigkeitsepisoden verbessert oder verschlechtert (vgl. Carmines/Zeller 1979). Da die Zeitabstände zwischen den Befragungen kurz sind, dürften Verzerrungen des Test-Retest-Maßes durch Veränderungen in der Stabilität der Jobanzahl weitestgehend ausgeschlossen sein und daher zuverlässige Informationen über die Reliabilität der Angaben der Befragten liefern (vgl. Diekmann 1995: 219). Allerdings sind Einflüsse des Lernens infolge wiederholter Befragungen

innerhalb kürzester Zeit auf die Test-Retest-Messung nicht ausgeschlossen. Denn während in der Regel die Test-Retest-Korrelationen die Zuverlässigkeit empirischer Messungen zumeist unterschätzt, können bei geringen Zeitabständen zwischen den einzelnen Messungen durchaus Überschätzungen der Zuverlässigkeit durch gelernte, konsistente Erinnerungen entstehen (Carmines/Zeller 1979: 40). Daher sind diese Messungen auch wegen möglicher Auswirkungen von Sensibilisierung und Reaktivität infolge wiederholter Messungen, sowohl in Verbindung mit den anderen Maßzahlen zu betrachten als auch vor dem Hintergrund der theoretischen Annahmen zu beurteilen.

5. Deskriptive Befunde zur Anzahl erinnerter Tätigkeiten im Berufsverlauf

Die zentrale *abhängige Variable* ist die Anzahl erinnerter Tätigkeiten im Berufsverlauf. Sie ist eine nachträglich generierte Variable, die sich aus der Zählung festgehaltener Tätigkeitsepisoden ergibt. Die Tätigkeitsepisoden werden durch Änderungen des ISCO voneinander abgegrenzt. Diese Vorgehensweise wurde auch für die Staatsbeschäftigten realisiert, ohne dass die Besonderheiten des Laufbahngruppensystems berücksichtigt wurden. Leider liegen im Datensatz die Klartextangaben der Befragten nicht vor, so dass weiterführende Untersuchungen zur Vercodungs-Reliabilität nicht vorgenommen werden konnten.

In *Tabelle 1* sind die deskriptiven Befunde über die Reliabilität der erinnerten Tätigkeiten im Berufsverlauf dokumentiert. Im Vergleich zur ersten Welle gaben 70,8 Prozent aller Befragten in der zweiten Welle die gleiche Anzahl von beruflichen Tätigkeiten (*Jobs*) an, und in der dritten Welle sinkt die Konsistenz im Vergleich zu den Angaben in der ersten Welle auf 67,4 Prozent. Die sinkenden Test-Retest-Koeffizienten weisen auf eine abnehmende Zuverlässigkeit der retrospektiven Angaben zu den Tätigkeitsepisoden hin. Die Abweichungswerte (DEV und ABSDEV) steigen von Befragung zu Befragung, wobei — so die negativen Vorzeichen der Abweichungswerte — die Befragten dazu neigen, jeweils mehr Tätigkeitsepisoden als in der vorherigen Befragung anzugeben (vgl. Huttenlocher et al. 1988: 473). Die entsprechenden prozentualen Abweichungswerte („Fehlerraten“) liegen bei 22,7 und 27,5 Prozent. Soweit wurden die Befunde von De Graaf und Wegener (1989: 9) repliziert. Zudem bemerken die beiden Autoren, dass in der zweiten Welle rund 23 Prozent der Befragten mehr Tätigkeitsepisoden nannten, während für rund 7 Prozent die Jobanzahl bei der Wiederbefragung geringer war. Im Vergleich zur ersten Welle berichteten in der dritten Welle ein Viertel der Befragten mehr Jobs und fast 8 Prozent weniger Jobs (vgl. De Graaf/Wegener 1989: 10).

Tabelle 1: Abweichungswerte, Konsistenz der Antworten und Test-Retest-Korrelation für die drei Wellen

Variablen	CONSIS	DEV	ABSDEV	%-ABDE	Test-Retest	N
<i>Erste-Zweite Welle</i>						
Insgesamt	70.8	-0.26	0.41	22.7	0.87	497
Privatwirtschaft	72.9	-0.26	0.40	22.0	0.85	354
Öffentlicher Dienst	67.5	-0.21	0.39	23.6	0.89	120
Beamte	50.0	-0.33	0.75	52.2	0.44	24
<i>Erste-Dritte Welle</i>						
Insgesamt	67.4	-0.32	0.60	27.5	0.81	439
Privatwirtschaft	70.6	-0.27	0.43	23.2	0.82	390
Öffentlicher Dienst	63.2	-0.39	0.61	37.3	0.77	109
Beamte	50.0	-0.36	0.73	51.5	0.28	22
<i>Alle Wellen</i>						
Insgesamt	67.3	-0.32	0.68	36.4	PW1-2/ PW1-3 0.86/ 0.81	437
Privatwirtschaft	70.4	-0.27	0.63	33.3	0.84/ 0.82	307
Öffentlicher Dienst	63.3	-0.39	0.75	44.2	0.89/ 0.77	109
Beamte	50.0	-0.36	0.82	60.6	0.25/ 0.28	22

Quelle: Projekt: „Prozesse der Berufsfindung“ (De Graaf/Wegener 1989) und eigene Berechnungen.

Beim Vergleich aller drei Wellen gleichzeitig gehen nur diejenigen Befragten in die Analyse ein, die an allen drei Befragungen teilgenommen hatten. Ein auffälliges Ergebnis ist einerseits, dass der Anteil derjenigen Befragten mit konsistenten Angaben in allen drei Wellen auf rund 67 Prozent sinkt. Andererseits steigt die prozentuale absolute durchschnittliche Abweichung, also die „Fehlerrate“ im Verhältnis zur ersten Welle, auf 36,4 Prozent an. Plausibelerweise nehmen auch die Werte für die Abweichungsmaße DEV (-0,32) und ABSDEV (0,68) zu, da sich die abweichenden Angaben zwischen den einzelnen Wellen kumulieren. So werden in der Tendenz über alle drei Wellen zu 32 Prozent mehr Jobs als in vorhergehenden Befragungen angegeben. Detaillierte Analysen ergaben, dass – kumuliert über alle Wellen – ein Viertel der Befragten jeweils mehr Jobs und gerade 7 Prozent der Befragten weniger Jobs berichteten.

Welche Unterschiede ergeben sich allerdings, wenn man zwischen den beiden Beschäftigungsbereichen *Öffentlicher Dienst* und *Privatwirtschaft* unterscheidet? Beim Vergleich der Angaben in der ersten Welle mit denen in den darauffolgenden Wellen zeigt sich bei den Beschäftigten in der Privatwirtschaft ein höherer Prozentanteil

konsistenter Angaben zur Jobanzahl, als bei den Beschäftigten im öffentlichen Dienst. Im Vergleich zur ersten Welle geben fast ein Viertel der Beschäftigten im öffentlichen Dienst in der zweiten Welle mehr Jobs an und 9,2 Prozent der staatlich Beschäftigten weniger Tätigkeiten an. Genau 21,5 Prozent der Beschäftigten in der Privatwirtschaft erinnert mehr Jobs in der zweiten Welle und 5,7 Prozent der privatwirtschaftlich Beschäftigten berichten in der zweiten Welle weniger Jobs als in der ersten Welle. In der dritten Welle gaben rund 28 Prozent der Beschäftigten im öffentlichen Dienst mehr Jobs an als in der ersten Welle, während dies fast 23 Prozent der privatwirtschaftlich Beschäftigten tun. Während 8 Prozent der Beschäftigten im öffentlichen Dienst in der letzten Welle weniger Jobs als in der ersten Welle angeben, tun dies dagegen fast 7 Prozent der Beschäftigten in der Privatwirtschaft. Somit sind die Unterschiede zwischen den beiden Sektoren recht gering.

Zieht man die Abweichungswerte (DEV) des Vergleichs der ersten beiden Wellen in Betracht, so ist für den öffentlichen Dienst festzustellen, dass die Befragten dort in einem geringeren Maße eine größere Jobanzahl in der zweiten Welle angeben als die Befragten in der Privatwirtschaft. Beim Vergleich zwischen der ersten und der dritten Welle zeigt sich allerdings, dass die staatlich Beschäftigten in der letzten Welle mehr Jobs angaben als die Beschäftigten in der Privatwirtschaft. Auch der absolute Abweichungswert von 0,61 für die staatlich Beschäftigten zeigt, dass sie im Vergleich zu den Beschäftigten in der Privatwirtschaft überproportional inkonsistente Angaben zur Jobanzahl machen. *Insgesamt ist festzustellen, dass Beschäftigte im öffentlichen Dienst inkonsistentere Angaben zur Jobanzahl als die Beschäftigten in der Privatwirtschaft machen. Diese Inkonsistenz erhöht sich mit dem zeitlichen Abstand der einzelnen Wellen.* Aber die geringen Abweichungswerte und hohen Test-Retest-Koeffizienten weisen dennoch auf eine hohe Reliabilität dieser Angaben hin. Diese Schlussfolgerung gilt auch, wenn man alle drei Wellen miteinander vergleicht. Der Anteil der Befragten mit konsistenten Angaben entspricht den Werten beim Vergleich der ersten mit der zweiten Welle bzw. der ersten Welle mit der dritten Welle. So machen 63 Prozent der Beschäftigten im öffentlichen Dienst über alle drei Wellen konsistente Angaben zu ihrer Jobanzahl, während der Prozentsatz bei Beschäftigten im privatwirtschaftlichen Sektor bei 70 Prozent liegt. Staatsbeschäftigte erinnern durchschnittlich bis zu 39 Prozent mehr Jobs und Beschäftigte in der Privatwirtschaft bis zu 27 Prozent mehr Jobs. Auch die absolute Abweichung ist für den öffentlichen Dienst höher als für die Privatwirtschaft (0,75 vs. 0,63). Die „Fehlerrate“ im Vergleich zur ersten Welle beträgt für den öffentlichen Dienst rund 44 Prozent und für die Privatwirtschaft rund 33 Prozent. Vergleicht man die Test-Retest-Werte für den Vergleich zwischen der zweiten und dritten Welle, so zeigt sich für den öffentlichen Dienst, dass dieser

Wert bei einem Vergleich über alle Wellen sinkt, während er für die Privatwirtschaft relativ stabil bleibt. Dies stützt wiederum den Befund von De Graaf und Wegener (1989), dass Beschäftigte im öffentlichen Dienst in der Retrospektive inkonsistentere Angaben zur Tätigkeitsanzahl machen.

Welche Ergebnisse findet man schließlich für *Beamte*? Bei den ersten beiden Wellenvergleichen machten jeweils nur 50 Prozent der Beamten konsistente Antworten zur Anzahl ihrer Tätigkeiten. Beim Vergleich der beiden letzten Wellen steigt dieser Anteil auf rund 90 Prozent. Detaillierte Analysen ergaben, dass 29 Prozent der Beamten in der zweiten Welle mehr Jobs angaben als in der ersten Welle und rund 21 Prozent der Beamten gaben weniger Jobs an. Gegenüber der ersten Welle gaben 32 Prozent der Beamten in der dritten Welle mehr Jobs an. Der Anteil der Beamten, die weniger Jobs erinnerten, beträgt 18 Prozent. Betrachtet man die Abweichungswerte, so findet man für die ersten beiden Wellen, dass die Beamten zu 33 Prozent mehr Jobs angaben als in der ersten Welle. Die absolute Abweichung beträgt 0,75 und der prozentuale Anteil der absoluten Abweichung an der ersten Welle hat einen Wert von 52,2. So liegt der absolute Abweichungswert für Beamte bei 0,73, und diese Beamten geben bis zu 36 Prozent mehr Jobs an. Der niedrige Test-Retest-Koeffizient von 0,44 für die Welle 1-2 sinkt auf 0,28 für die Welle 1-3. Vergleicht man jedoch alle Wellen miteinander, so entspricht der Anteil derjenigen Beamten mit konsistenten Angaben (50 Prozent) den Anteilen in den vorhergehenden Wellen. Bei diesem Wellenvergleich zeigt sich, dass Beamte in den jeweils darauffolgenden Befragungen zu 36 Prozent mehr Jobs angeben. Die durchschnittliche absolute Abweichung erhöht sich auf 61 Prozent. Berücksichtigt man die Test-Retest-Werte, so machen Beamte in der ersten nachfolgenden Welle in einem hohen Grade inkonsistente Retrospektivangaben zur Jobanzahl (Welle 1-2: 0,25; Welle 1-3: 0,28). *Diese Befunde weisen darauf hin, dass für Beamte eine äußerst niedrige Reliabilität der erinnerten Anzahl von Tätigkeitsepisoden im Berufsverlauf vorliegt.* Jedoch sind die Befunde für die Beamten aufgrund ihrer geringen Fallzahl mit einiger Vorsicht zu beurteilen. Deswegen ist es auch nicht möglich, zu überprüfen, ob sich hinter diesen Ergebnissen qualifikatorische Unterschiede zwischen den einzelnen Laufbahnen verbergen. So könnte man durchaus vermuten, dass die Beamten im unteren und mittleren Dienst unzuverlässigere Angaben machen als die qualifizierteren Beamten im gehobenen und höheren Dienst (vgl. Becker 1993).

6. Erklärende Faktoren für die Reliabilität der Jobanzahl

Im Vordergrund steht nunmehr die Überprüfung der These, dass nicht generell die Staatsbeschäftigten unzuverlässigere Angaben zu ihrem Berufsverlauf machen als Beschäftigte im privatwirtschaftlichen Sektor, sondern dass gerade Beamte inkonsistentere Angaben zur Anzahl ihrer Tätigkeiten machen. Sie dürften größere Schwierigkeiten haben, die einzelnen Tätigkeiten und Tätigkeitswechsel in der Retrospektive voneinander abzugrenzen, wenn nicht die einzelnen Ämter in den Beamtenlaufbahnen, sondern die beruflichen Tätigkeiten die Erhebungseinheiten sind. Eine Möglichkeit, die Verteilung der absoluten Abweichungswerte für die Tätigkeitsepisoden anhand von verschiedenen Merkmalen der Befragten differenzierter zu beschreiben, ist die Anwendung multivariater Regressionsanalysen. Jedoch entspricht die Verteilung dieser Abweichungswerte keineswegs einer Normalverteilung, so dass die Fehlerwerte nicht normalverteilt und somit die Voraussetzungen für eine OLS-Regression nicht gegeben sind. Stattdessen bietet sich die logistische Regression an, um die Verteilung der Wahrscheinlichkeiten für inkonsistente Angaben zu schätzen. Die *abhängige Variable* ist die *Inkonsistenz der erinnerten Anzahl beruflicher Tätigkeiten* (0 = konsistente Antwort und 1 = inkonsistente Antwort). Die Modell schätzung erfolgt in zwei Schritten. Zuerst wird die Inkonsistenz für die erste Tätigkeit (Tabelle 2) und dann für die letzte Tätigkeit (Tabelle 3) geschätzt, wobei die drei Wellen miteinander verglichen werden.

Tabelle 2: Logistische Regressionsschätzung für die Inkonsistenz von Angaben zur Jobanzahl – Faktoren des ersten Jobs (in Klammern: Standardfehler der Koeffizienten)

	De Graaf/ Wegner PW1-PW2	De Graaf/ Wegner PW1-PW3	PW1-PW2 Modell 1	PW1-PW3 Modell 2	PW1-PW2- PW3
Anzahl der Jobs	0.26* (0.04)	0.35* (0.05)	0.27* (0.04)	0.35* (0.04)	0.31* (0.05)
Arbeitszeit			$-9 \cdot 10^{-3}$ (0.01)	$-2 \cdot 10^{-3}$ (0.01)	$3 \cdot 10^{-3}$ (0.01)
Bildung			$-1 \cdot 10^{-3}$ (0.06)	$8 \cdot 10^{-3}$ (0.06)	0.08 (0.07)
Geschlecht			0.02 (0.12)	-0.10 (0.12)	$-2 \cdot 10^{-3}$ (0.14)
Anfangszeitpunkt des ersten Jobs			$1 \cdot 10^{-3}$ ($1 \cdot 10^{-3}$)	$1 \cdot 10^{-3}$ ($1 \cdot 10^{-3}$)	$1 \cdot 10^{-3}$ ($1 \cdot 10^{-3}$)
Öffentlicher Dienst	-0.32* (0.13)	-0.37* (0.14)			
Beamte			0.70* (0.27)	0.65* (0.28)	0.37 (0.34)
Angestellte im öffentlichen Dienst			0.03 (0.23)	-0.02 (0.26)	-0.46 (0.36)
Angestellte in der Privatwirtschaft			0.14 (0.17)	-0.10 (0.20)	-0.08 (0.22)
Arbeiter im öffentlichen Dienst			0.14 (0.57)	0.32 (0.52)	0.23 (0.51)
Arbeiter in der Privatwirtschaft			0.27 (0.18)	0.09 (0.20)	0.27 (0.22)
N			486	428	336
χ^2			495.9	502.6*	260.6
d.f.			10	10	10

* $p \leq 0.05$

Quelle: Projekt: „Prozesse der Berufsfindung“ – eigene Berechnungen.

In den ersten beiden Spalten werden die Ergebnisse von De Graaf und Wegener (1989) wiedergegeben.⁷ Demnach werden inkonsistente Angaben zu den Tätigkeitsepisoden umso wahrscheinlicher, je größer die Anzahl der Jobs ist. Auch mit unseren Modellspezifikationen bleibt dieser Effekt für alle Wellenvergleiche erhalten (vgl. *Tabelle 2*).⁸ Die Anzahl der Tätigkeiten kann als Indikator für die Dauer des Interviews mit den entsprechenden Auswirkungen auf die Aufmerksamkeit interpretiert werden. Des weiteren scheinen größere Erinnerungsfehler aufzutreten, je komplexer der Berufsverlauf ist und je weiter die einzelnen Ereignisse zurückliegen.⁹ Jedoch stellten De Graaf und Wegener (1989) fest, dass die Zeitvariablen wie etwa Beginn oder Ende einer Tätigkeitsepisode keinen Einfluss auf die Inkonsistenz der entsprechenden Angaben zum Berufsverlauf haben. Dies gilt auch für den Anfangszeitpunkt (gemessen in Monaten seit 1900) der ersten beruflichen Tätigkeit, die wir in unseren Modellschätzungen berücksichtigt haben. Daher ist es nicht zwingend, dass inkonsistente Angaben dann auftreten, wenn die Ereignisse oder Episoden weit zurückliegen.

Wie eingangs vermutet, verschwindet der Inkonsistenzeffekt für den öffentlichen Dienst, wenn die berufliche Stellung der Befragten kontrolliert wird. Die Referenzkategorie sind die Selbständigen und Landwirte. Die Schlussfolgerung der beiden Autoren, dass privatwirtschaftlich Beschäftigte in einem höheren Maße konsistente Angaben bei der Wiedererinnerung der einzelnen Tätigkeiten machen als staatlich Beschäftigte, ist in dieser Form nicht ohne weiteres haltbar. Vielmehr sind es gerade die Beamten, die inkonsistentere Angaben zur ihrer Berufsgeschichte machen. In anderen hier nicht dokumentierten Modell-Schätzungen wurde der „Beamteneffekt“

⁷ Leider haben die Autoren lediglich die signifikanten Effekte dokumentiert. Auch liegen keine weiteren Kennwerte zu den Modellschätzungen vor.

⁸ Da es nicht um die Güte der Modellanpassung geht, sondern lediglich um die Isolierung bedeutsamer erklärender Variablen für die Inkonsistenz der Jobanzahl, wird auf diese nicht gesondert eingegangen.

⁹ Aufgrund von Identifikationsfehlern kann die These über den Zusammenhang von Alter und Gedächtnisfehler nicht überprüft werden. Bei der Kontrolle des Alters zu einem Befragungszeitpunktes ist unklar, ob sich dahinter nicht nur Lebenszykluseffekte, sondern auch Kohorten- und Periodeneffekte verbergen. Verwendet man in multivariaten Schätzmodellen gleichzeitig die mit dem gleichen Maßstab gemessenen Zeitdimensionen Alter, Periode und Kohorte, dann entsteht das Identifikationsproblem, weil bei Kenntnis von zwei Zeitdimensionen die dritte Zeitdimension eindeutig identifiziert werden kann (z.B. Alter = Periode-Kohorte) (vgl. Mason et al. 1973). Um dieses Identifikationsproblem zu lösen, kann zum einen unterstellt werden, dass eine der drei Zeitdimensionen mangels Effekt auf die abhängige Variable vernachlässigt werden kann. Diese Lösung ist in unserem Falle theoretisch inplausibel (vgl. Mayer/Huinink 1990: 454-455).

auch durch die Differenzierung der einzelnen Laufbahnen detailliert untersucht. Hierbei wurde kein „Laufbahn-Effekt“ gefunden. Jedoch ist dieser Befund wegen der geringen Fallzahlen für die Beamten mit äußerster Vorsicht zu interpretieren. Zwar verschwindet der „Beamteneffekt“ aufgrund von Lern- oder Stichprobeneffekten, wenn alle Wellen miteinander verglichen werden. In methodischer Hinsicht wäre ein Aufatmen verfrüht, da Retrospektivdaten — etwa bei der Lebensverlaufsstudie oder beim Sozio-ökonomischen Panel — einmalig und nicht im Paneldesign wiederholt erhoben werden. Daher ist — vorausgesetzt, dass bei der Erhebung des Berufsverlaufs die einzelnen beruflichen Tätigkeiten anhand von Änderungen des ISCO abgegrenzt werden — anzunehmen, dass dann Beamte mit einer größeren Wahrscheinlichkeit inkonsistentere Angaben machen als andere Beschäftigtengruppen. Mit der gebotenen Vorsicht ist daher zu vermuten, dass der „Beamteneffekt“ zum einen ein Kontexteffekt ist, wenn man alleine auf die beruflichen Tätigkeiten abstellt. Für Beamte sind Tätigkeitswechsel weniger eindeutig als für andere Beschäftigten. Zum anderen dürfte dieser Effekt auch durch das Erhebungsinstrumentarium hervorgerufen werden, das die institutionellen Besonderheiten der beruflichen Karriere von Beamten nicht berücksichtigt (vgl. Becker 1993).

Letztlich werden die Effekte für die Reliabilität der Jobangaben für die letzte Tätigkeit analysiert (vgl. *Tabelle 3*). Für alle Wellenvergleiche gilt: Je mehr Tätigkeiten eine befragte Person in ihrem Berufsverlauf hatte, umso geringer ist die Konsistenz der Antworten bei Wiederbefragungen.

Das Muster der „Beamteneffekte“ entspricht den vorhergehenden Analysen mit den Merkmalen der ersten beruflichen Tätigkeit. Im Unterschied zu diesen Analysen weisen die signifikant negativen Effekte für *Angestellte in der Privatwirtschaft* darauf hin, dass vor allem diese Beschäftigtengruppe in der Retrospektive am ehesten zuverlässige, hochgradig konsistente Angaben zu ihrem gesamten Berufsverlauf macht. Vermutlich entspricht für diese Angestellten das verwendete ereignisorientierte Erhebungsinstrumentarium am ehesten den Strukturen des im autobiographischen Gedächtnis gespeicherten Berufsverlaufs (vgl. Sudman et al. 1996: 168-169). Abgesehen von den Beamten und Landwirten dürften sie im Vergleich zu den Arbeitern und den weniger qualifizierten Selbständigen auch intellektuell eher in der Lage sein, ihre Erwerbsgeschichte in konsistenter Weise zu rekonstruieren. Aufgrund geringer Fallzahlen können diese Vermutungen mit den herangezogenen Paneldaten nicht überprüft werden.

Tabelle 3: Logistische Regressionsschätzung für die Inkonsistenz von Angaben zur Jobanzahl – Faktoren des letzten Jobs (in Klammern: Standardfehler der Koeffizienten)

	PW1-PW2	PW1-PW3	PW1-PW2-PW3
Anzahl der Jobs	0.27* (0.04)	0.36* (0.05)	0.34* (0.05)
Arbeitszeit	$7 \cdot 10^{-3}$ ($6 \cdot 10^{-3}$)	$5 \cdot 10^{-3}$ ($6 \cdot 10^{-3}$)	$7 \cdot 10^{-2}$ ($9 \cdot 10^{-3}$)
Bildung	$3 \cdot 10^{-3}$ (0.06)	0.07 (0.06)	0.09 (0.06)
Geschlecht	0.19 (0.13)	0.09 (0.14)	0.17 (0.13)
Anfangszeitpunkt des letzten Jobs	$2 \cdot 10^{-4}$ ($2 \cdot 10^{-3}$)	$-3 \cdot 10^{-4}$ ($1 \cdot 10^{-3}$)	$1 \cdot 10^{-4}$ ($1 \cdot 10^{-3}$)
Beamte	0.57* (0.25)	0.56* (0.27)	0.38 (0.26)
Angestellte im öffentlichen Dienst	-0.14 (0.20)	-0.31 (0.22)	(0.23 (0.21)
Angestellte in der Privatwirtschaft	-0.38* (0.15)	-0.48* (0.17)	-0.42* (0.16)
Arbeiter im öffentlichen Dienst	0.02 (0.34)	0.14 (0.35)	0.24 (0.33)
Arbeiter in der Privatwirtschaft	-0.18 (0.16)	-0.01 (0.17)	0.02 (0.17)
N	485	428	426
χ^2	513.0*	630.0*	558.6*
d.f.	10	10	10

* $p \leq 0.05$

Quelle: Projekt: „Prozesse der Berufsfindung“ – eigene Berechnungen

7. Zusammenfassung und Schlussfolgerung

In der vorliegenden Untersuchung wurde der Frage nachgegangen, ob bei Retrospektivbefragungen Beschäftigte im öffentlichen Dienst unzuverlässigere Antworten zu Fragen ihres Berufsverlaufs geben als Beschäftigte in der Privatwirtschaft. Der Ausgangspunkt der Analysen ist ein Befund von De Graaf und Wegener (1989), wonach Staatsbeschäftigte die einzelnen beruflichen Tätigkeiten in ihrem Berufsverlauf unzuverlässiger erinnern. Mit Hilfe der Panel-Daten des Projektes „Prozesse der Berufsfindung“ sollte dieser partielle Aspekt aufgeheitelt werden. Es wurde die These aufgestellt, dass gerade Beamte aufgrund ihrer Platzierung im Laufbahngruppensystem mit seinen immanenten spezifischen Regelungen von Ausbildung, Platzierung und Allokation von Beschäftigten, größere Schwierigkeiten haben, Tätigkeitswechsel in ihrem Berufsverlauf abzugrenzen. Dies schlägt sich in der Retrospektive in vergleichsweise weniger zuverlässigen Angaben zur Anzahl von Tätigkeiten nieder.

Den empirischen Befunden zufolge kann insgesamt von einer hohen *Reliabilität retrospektiver Berufsverlaufsdaten* und damit von einer *guten Datenqualität solcher Längsschnittdaten* ausgegangen werden. Jedoch zeigte sich in der Tat für berufliche Tätigkeitswechsel, dass Beamte offensichtlich Schwierigkeiten haben, die einzelnen Tätigkeiten voneinander abzugrenzen. Nicht endgültig geklärt ist die Frage, ob dies wiederum generell für alle Beamten gilt oder ob es hier qualifikatorisch bedingte Unterschiede zwischen den einzelnen Laufbahnen gibt. Vorliegende Befunde sprechen eher gegen Unterschiede zwischen den Laufbahnen und sprechen für unsere Kontext-Hypothese über die institutionellen Besonderheiten des Berufsstatus und der beruflichen Tätigkeiten von Beamten. Allerdings gab es auch deutliche Hinweise, dass die unzuverlässigeren Angaben der Beamten auch erhebungstechnisch bedingt sind, weil die verwendeten Erhebungsinstrumentarien die institutionellen Besonderheiten der Tätigkeitsmerkmale von Beamten, des Laufbahnsystems im öffentlichen Dienst und der daran geknüpften Berufswege der Beamten nicht in adäquater Weise berücksichtigen. Weiterführende Analysen mit größeren Datensätzen könnten diese offenen Fragen abschließend klären. Sie könnten auch dazu dienen, in differenzierter Weise weitere Gründe für die Zuverlässigkeit retrospektiver Berufsverlaufdaten zu untersuchen.

Welche Konsequenzen für die Retrospektivbefragungen von Berufsverläufen mit *Tätigkeitswechsel als Untersuchungseinheiten* sind daraus zu ziehen? Das Hauptgewicht einer Reorganisation vorhandener Erhebungsinstrumente — wie sie etwa bei der deutschen Lebensverlaufsstudie eingesetzt wurden — dürfte nur bei der Befragung von Beamten liegen. Bei dieser Beschäftigtengruppe müssten auch andere Merkmale ihrer

Tätigkeiten, wie z.B. Dienstposten bzw. Amt, Gehaltsstufe, Beförderungen, usw. als offene Fragen abgefragt werden. Denn diese Merkmale sind für Beamte Bestandteil des institutionellen Kontextes, und könnten die Zuverlässigkeit der Wiedergabe der Tätigkeiten und ihrer Änderungen erhöhen, da sie einen Orientierungsrahmen für die Wiedererinnerung bieten könnten. Zusätzlich sollten bei der Beschreibung der Tätigkeiten durch die beamteten Befragten neben der Einräumung einer längeren Zeitspanne für die Erinnerung ein größerer Beantwortungsfreiraum für diese Beschäftigtengruppe gegeben werden.

Korrespondenzadresse

Dr. Rolf Becker

Technische Universität Dresden, Institut für Soziologie

Mommsenstraße 13

D - 01069 Dresden

Tel. 0351/463-4112; Fax. 0351/463-7085

E-Mail: becker-r@rcs.urz.tu-dresden.de

Literaturverzeichnis

Allmendinger, J., 1989: Career Mobility Dynamics. A Comparative Analysis of the United States, Norway, and West Germany, Berlin: edition sigma (Reihe Studien und Berichte des Max-Planck-Instituts für Bildungsforschung Nr. 49).

Becker, R., 1993: Staatsexpansion und Karrierechancen. Berufsverläufe im öffentlichen Dienst und in der Privatwirtschaft, Frankfurt am Main: Campus.

Becker, R./K. Schömann, 1996: Berufliche Weiterbildung und Einkommensdynamik. Eine Längsschnittstudie mit besonderer Berücksichtigung von Selektionsprozessen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48: 426-461.

Behr, M./R. Schultz-Wild, 1973: Arbeitsplatzstruktur und Laufbahnreform im öffentlichen Dienst (Studienkommission für die Reform des öffentlichen Dienstrechts, Bd. 9). Baden-Baden: Nomos.

Blossfeld, H.-P., 1989: Kohortendifferenzierung und Karriereprozess. Frankfurt am Main: Campus.

Blossfeld, H.-P./A. Hamerle/K.U. Mayer, 1986: Ereignisanalyse. Frankfurt: Campus.

Blossfeld, H.-P./K.U. Mayer, 1988: Arbeitsmarktsegmentation in der Bundesrepublik Deutschland. Eine empirische Überprüfung von Segmentationstheorien aus der Perspektive des Lebenslaufs. Kölner Zeitschrift für Soziologie und Sozialpsychologie 40: 262-283.

Blossfeld, H.-P./K.U. Mayer, 1991: Berufsstruktureller Wandel und soziale Ungleichheit. Entsteht in der Bundesrepublik Deutschland ein neues Dienstleistungsproletariat? Kölner Zeitschrift für Soziologie und Sozialpsychologie 43: 671-696.

- Blossfeld, H.-P./R. Becker, 1989: Arbeitsmarktprozesse zwischen öffentlichem und privatwirtschaftlichem Sektor. Kohortenspezifische Auswirkungen der Expansion des Staates als Arbeitgeber. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 22: 233-247.
- Blum, Z./D., Nancy/L. Karweit/Aage B. Sørensen, 1969: *Collection and Analysis of Retrospective Life Histories*. Baltimore: The John Hopkins University.
- Bradburn, N./M., Janelen Huttenlocher/L. Hedges, 1994: Telescoping and Temporal Memory. S. 203-215 in: Norbert Schwarz und Seymour Sudman (Hg.): *Autobiographical Memory and the Validity of Retrospective Reports*. Berlin/New York: Springer.
- Bradburn, N. M./Lance J. Rips/S.K. Shewell, 1987: Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys. *Science* 236: 157-161.
- Brinkmann, Gerhard, 1976: *Aufgaben und Qualifikation der öffentlichen Verwaltung*, Göttingen: Schwartz.
- Brückner, E., 1990: Die retrospektive Erhebung von Lebensverläufen. S. 374-403 in: Karl Ulrich Mayer (Hg.): *Lebensverläufe und sozialer Wandel (Sonderheft 31 der Kölner Zeitschrift für Soziologie und Sozialpsychologie)*. Opladen: Westdeutscher Verlag.
- Brückner, H., 1995: *Surveys Don't Lie, People Do? An Analysis of Data Quality in a Retrospective Life Course Study*. Materialien aus der Bildungsforschung Nr. 50, Berlin: Max-Planck-Institut für Bildungsforschung.
- Brückner, H./K.U. Mayer, 1995: *Lebensverläufe und gesellschaftlicher Wandel. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954-1956 und 1959-61*. Materialien aus der Bildungsforschung Nr. 48, Berlin: Max-Planck-Institut für Bildungsforschung.
- Brüderl, J., 1990: Zur Analyse von Einkommensverläufen mit Längsschnittdaten. *Allgemeines Statistisches Archiv* 74: 213-222.
- Carmines, E. G./R. A. Zeller, 1979: *Reliability and Validity Assessment*. Newbury Park: Sage.
- Carroll, G. R./E. M. Mosakowski, 1988: The Career Dynamics of Self-Employment. *Administrative Science Quarterly* 32: 570-589.
- Carroll, G. R./K.U. Mayer, 1986: Job-Shift Patterns in the Federal Republic of Germany: The Effects of Social Class, Industrial Sector, and Organizational Size. *American Sociological Review* 51: 323-341.
- De Graaf, N. D./B. Wegener, 1989: Reliability of Job Career and Social Resource Data. Paper presented at the Utrecht Meeting of the ISA Research Committee on Social Stratification, 26.-29. April 1989. Berlin: Max-Planck-Institut für Bildungsforschung (unveröffentlichtes Manuskript).
- Dex, S./A. McCulloch, 1998: The Reliability of Retrospective Unemployment History Data. *Work, Employment & Society* 12: 497-509.
- Diekmann, A., 1995: *Empirische Sozialforschung*. Reinbek: Rowohlt.
- Friedrichs, J., 1973: *Methoden der empirischen Sozialforschung*. Reinbek: Rowohlt.
- Goldfarb, N., 1960: *An Introduction to Longitudinal Statistical Analysis*. Glencoe: University Press.
- Huttenlocher, J./L. Hedges/V. Prohaska, 1988: Hierarchical Organization in Ordered Domains: Estimating the Dates of Events. *Psychological Review* 95: 471-484.
- Jahoda, M./M. Deutsch/S. Cook, und andere, 1951: *Research Methods in Social Relations (two volumes)*, New York: Sage.

- Klein, T./D. Fischer-Kerli, 2000: Die Zuverlässigkeit retrospektiv erhobener Lebensverlaufsdaten. Analysen zur Partnerschaftsbiographie des Familiensurvey. *Zeitschrift für Soziologie* 29: 294-312.
- Konietzka, D., 1999: Ausbildung und Beruf. Die Geburtsjahrgänge 1919-1961 auf dem Weg von der Schule in das Erwerbsleben. Opladen: Westdeutscher Verlag.
- Lauterbach, W., 1994: Berufsverläufe von Frauen. Frankfurt am Main: Campus.
- Linton, M., 1978: Gedächtnis für Ereignisse der Realwelt. S. 371-396 in: Donald A. Norman und David E. Rumelhart (Hg.): *Strukturen des Wissens. Wege der Kognitionsforschung*. Stuttgart: Klett-Cotta.
- Little, R. J. A., 1992: Incomplete Data in Event History Analysis. S. 209-230 in: J. Trussel/R. Hankinson/J. Tilton (Hg.): *Demographic Application of Event History Analysis*. Oxford: Clarendon Press.
- Mason, Karen Oppenheim/H.H. Winsborough/W. M. Mason/W. K. Poole, 1973: Some Methodological Issues In Cohort Analysis Of Archival Data, *American Sociological Review* 38: 242-258.
- Mayer, K.U./Brückner E., 1989: Lebensverläufe und Wohlfahrtsentwicklung (Materialien aus der Bildungsforschung Nr. 35), Berlin: Max Planck-Institut für Bildungsforschung.
- Mayer, K.U./G. Papastefanou, 1983: Arbeitseinkommen im Lebensverlauf. Probleme der retrospektiven Erfassung und empirische Materialien. S. 101-122 in: W. Schmähl (Hg.): *Ansätze der Lebenseinkommensanalyse*. Tübingen: Mohr (Siebeck).
- Mayer, K.U./G. R. Carroll, 1987: Jobs and Classes: Structural Constraints on Career mobility. *European Sociological Review* 3: 14-38.
- Mayer, K.U./J. Huinink, 1990: Alters-, Perioden- und Kohorteneffekte in der Analyse von Lebensverläufen oder: Lexis Ade? S. 442-459 in: K.U. Mayer (Hg.): *Lebensverläufe und sozialer Wandel (Sonderheft 31 der Kölner Zeitschrift für Soziologie und Sozialpsychologie)*. Opladen: Westdeutscher Verlag.
- Mayer, K.U./Nancy B. Tuma, 1990: Life Course Research and Event History Analysis: An Overview. S. 3-20 in: K.U. Mayer/Nancy B. Tuma (Hg.): *Event History Analysis in Life Course Research*. Madison, Wiss.: The University of Wisconsin Press.
- Papastefanou, G., 1980: Zur Güte von retrospektiven Daten. Eine Anwendung gedächtnispsychologischer Theorie und Ergebnisse einer Nachbefragung. Frankfurt/Mannheim: Arbeitspapier Nr. 29 des SFB 3 "Mikroanalytische Grundlagen der Gesellschaftspolitik".
- Pearson, R.W./M. Ross/R. M. Dawes, 1992: Personal Recall and the Limits of Retrospective Questions in Surveys. S. 65-94 in: J. M. Tanur (Hg.): *Questions About Questions. Inquiries into the Cognitive Bases of Surveys*. New York: Sage.
- Peters, H. E., 1988: Retrospective Versus Panel Data in Analyzing Lifecycle Events. *The Journal of Human Resources* 23: 488-513.
- Ramsøy, N. R., 1973: The Norwegian Occupational Life History Study: Design, Purpose, and a few preliminary Results. Oslo: Institute of Applied Social Research.
- Rendtel, U., 1995: Panelausfälle und Panelrepräsentativität. Frankfurt am Main: Campus.
- Schneid, M., 1986: Feldbericht - Soziale Beziehungen im Prozess der Berufsfindung (Technischer Bericht Nr. 88/06). Mannheim: ZUMA.

- Schömann, K., 1994: The Dynamics of Labor Earnings over the Life Course. A Comparative and Longitudinal Analysis of Germany and Poland, Berlin: Max-Planck-Institut für Bildungsforschung/edition sigma.
- Schwarz, N., 1996: Cognition and Communication. Judgmental Biases, Research Methods, and the Logic of Conversation, Mahwah, N.J.: Lawrence Erlbaum.
- Schwarz, N./H.-J. Hippler/E. Noelle-Neumann, 1994: Retrospective Reports: The Impact of Response Formats. S. 187-199 in: N. Schwarz/S. Sudman (Hg.): *Autobiographical Memory and the Validity of Retrospective Reports*. New York: Springer.
- Strube, G., 1987: Answering Survey Questions: The Role of Memory. S. 86-101 in: H.-J. Hippler/N. Schwarz/S. Sudman (Hg.): *Social Information Processing and Survey Methodology*. New York: Springer.
- Strube, G./F. E. Weinert, 1987: Autobiographisches Gedächtnis: Mentale Repräsentation der individuellen Biographie. S. 151-167 in: G. Jüttemann/H. Thomae (Hg.): *Biographie und Psychologie*. Berlin: Springer.
- Sudman, S./N. M. Bradburn/N. Schwarz, 1996: *Thinking about Answers. The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers.
- Sudman, S./N. M. Bradburn, 1974: *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Tölke, A., 1980: Zuverlässigkeit retrospektiver Verlaufsdaten — Qualitative Ergebnisse einer Nachbefragung. Frankfurt/Mannheim: Arbeitspapier Nr. 30 des SFB 3 "Mikroanalytische Grundlagen der Gesellschaftspolitik".
- Tuma, N. B., 1985: Effects of labor market structures on job shift patterns. S. 327-363 in: J. J. Heckman/B. Singer (Hg.): *Longitudinal analysis of labor market data*. Cambridge: Cambridge University Press.
- Tversky, A./D. Kahnemann, 1986: Rational Choice and the Framing of Decisions. S. 67-94 in: R. M. Hogarth/M. W. Reder (Hg.): *Rational Choice. The Contrast between Economics and Psychology*. Chicago: University of Chicago Press.

DER EINFLUSS DER EXPERTISE AUF DAS ANTWORTVERHALTEN IN DELPHI-STUDIEN: EIN HYPOTHESENTEST¹

KNUT BLIND UND KERSTIN CUHLS

Basierend auf den Ergebnissen von Delphi'98, einer Delphi-Studie zur globalen Entwicklung von Wissenschaft und Technik im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF), wurde der Einfluss der Expertise der beteiligten Experten auf das Antwortverhalten getestet. Eine erste Hypothese, dass Experten mit geringem Expertenwissen mit geringerer Wahrscheinlichkeit an der zweiten Befragungsrunde teilnehmen, konnte für die Mehrheit der zwölf Technikfelder bestätigt werden. Auf Basis der individuellen Einschätzungen der möglichen Realisierungszeiträume von über 1.000 zur Diskussion gestellten Thesen im Delphi'98 wurde eine Reihe weiterer Hypothesen getestet. Für die Bewertung und Weiterentwicklung der Delphi-Methode besonders interessant ist die weit gehende Bestätigung der Hypothese, dass bei höherer Expertise die Experten weniger bereit sind, sich der Durchschnittsmeinung anzuschließen, und damit den Konvergenzprozess nicht unterstützen. Diesem Zusammenhang muss sowohl bei der Durchführung von Delphi-Untersuchungen als auch bei der Analyse der Ergebnisse Rechnung getragen werden.

The second wave of a Delphi study on the global development of science and technology (Delphi '98) commissioned by the German Federal Ministry of Education, Research, Science and Technology (BMBF) was used to investigate relationships between the level of expertise and response and non-response behaviours. For the majority of the 12 technological fields covered in the study, the level of expertise correlates positively with participation in the second wave of the Delphi panel. Respondents' assessments of the expected realisation time frame for more than 1,000 topics were used to investigate a number of hypotheses. Our findings support the hypothesis that the greater their expertise, the less likely respondents are to agree with the average assessment of a topic. Such experts therefore do not support the process of

1 Die Autoren danken zwei anonymen Gutachtern für wertvolle Hinweise.

convergence. This relationship has to be taken into account in the conduct of Delphi studies and the analysis of the results.

1. Einleitung

Delphi-Studien werden bereits seit den fünfziger Jahren durchgeführt. Anfangs als reines Prognose-Instrument betrachtet, traten später andere Ziele wie Konsensfindung oder Anregung von Kommunikation in den Vordergrund. Erst mit den umfassenderen nationalen Studien der neunziger Jahre ist es möglich geworden, ausreichend große Samples für die Überprüfung von Hypothesen bezüglich des individuellen Antwortverhaltens der Experten zu erhalten. Die Studie Delphi '98, die das Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI) im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) durchgeführt hat, ist eine derartige Studie und liegt der folgenden Hypothesenprüfung zu Grunde.

Der Kern des Delphi-Verfahren besteht aus mindestens zwei so genannten „Runden“. Von Fachkommissionen erarbeitete Thesen werden einer großen Anzahl an Experten zur Bewertung vorgelegt. Deren Antworten werden ausgewertet und denselben Personenkreisen noch einmal zugeschickt. In dieser zweiten Runde sollen die Experten ihre Antworten unter dem Einfluss der Einschätzungen ihrer Fachkollegen noch einmal überdenken und ihre Meinung ändern - oder nicht. Anonymität ist gewährleistet, so dass bei einer Meinungsänderung niemand sein Gesicht verliert oder sich rechtfertigen muss.

In den letzten Jahren wuchsen Zweifel, ob es richtig sei, für eine Delphi-Befragung nur „Experten“ zu Rate zu ziehen und nur diese an den nationalen Foresight-Übungen teilnehmen zu lassen, denn somit wird ihnen eine grundlegende Entscheidungsvorbereitung überlassen (Cuhls 2000). Im Hintergrund steht die Annahme, dass Experten dazu neigen, ihr eigenes Arbeitsgebiet zu überschätzen (Bias). Dies konnte in den vom ISI durchgeführten Delphi-Studien bisher jedoch nur in ausgewählten Themenfeldern bestätigt werden, z. B. für die Energie oder die Biotechnologie (Grupp 1995), für die Mehrzahl der Felder fanden sich dagegen keine grundsätzlichen Auffälligkeiten (Cuhls/Blind/Grupp 2001).

Der vorliegende Artikel greift vier Hypothesen zur Bedeutung des Expertenstatus für das Verhalten in Delphi-Studien auf, die anhand der Delphi '98-Daten differenziert nach Themenfeldern getestet werden. Jedoch wird zunächst kurz beschrieben, wie das Delphi '98 durchgeführt wurde. Der Beitrag schließt mit Folgerungen für die zukünftige Durchführung von Delphi-Studien und die Analyse der Ergebnisse.

2. Hintergrund: Delphi´98

In Deutschland werden seit Anfang der neunziger Jahre auf nationaler Ebene Vorausschau-Studien durchgeführt. Verschärfter Wettbewerb auf den Weltmärkten, Globalisierung der Wirtschaftsunternehmen und ein immer unüberschaubar werdender technischer Fortschritt zwingen dazu, nach Orientierungswissen zu suchen. Die wesentliche Fragestellung lautet: Welchen technischen Möglichkeiten kommt in Zukunft die größte Bedeutung zu, um die ökonomische Wettbewerbsfähigkeit, die Erhaltung oder Wiederherstellung einer lebenswerten Umwelt oder die soziale Wohlfahrt zu gewährleisten?

Mit Delphi-Studien kann ein fundierter Blick in die Zukunft gewagt werden. Wie die Zukunft tatsächlich aussehen wird, weiß niemand. Es ist jedoch möglich, bereits jetzt bestimmte Entwicklungen einzuschätzen, sie gemeinsam zu überprüfen, zu diskutieren und dann Maßnahmen zu überlegen, um sie Wirklichkeit werden zu lassen - oder sie gegebenenfalls zu verhindern. Damit wird Zeit gewonnen, evidente Fehlentwicklungen zu bremsen oder Innovationen, die dringend benötigt werden, anzustoßen oder rascher voranzutreiben. Delphi-Studien liefern also nicht einfach ein Bild von der Zukunft, sondern eine Informationsgrundlage für die Entscheidung, was heute zu tun oder zu lassen ist. Dabei wird nicht mehr davon ausgegangen, dass grundsätzlich ein Konsens erreicht wird oder erreicht werden kann, sondern bereits die Feststellung, dass bezüglich eines bestimmten Zukunftsthemas Konsens herrscht (oder eben nicht), ist eine für die teilnehmenden Akteure wichtige Information.

Die erste deutsche Delphi-Studie zur Entwicklung von Wissenschaft und Technik (BMFT 1993) wurde 1993 im Auftrag des BMFT vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI), das mit dem japanischen National Institute of Science and Technology Policy (NISTEP) zusammenwirkt, durchgeführt. Da sich die erste deutsche Delphi-Studie in der Umsetzung und strategischen Nutzung für die Wirtschaft, aber auch für staatliche Entscheidungsträger als sehr hilfreich erwiesen hat, wurde mit den Mini-Delphi-Studien (Cuhls/Breiner/Grupp 1995) von 1995 in acht wissenschaftlich-technischen Problemlösungsbereichen die japanisch-deutsche Kooperation fortgesetzt und die Methode weiterentwickelt. Frankreich, Südkorea und Großbritannien sowie seit kurzem auch andere Länder wie Österreich, Ungarn oder Südafrika sind inzwischen der Initiative, Delphi-Befragungen für eine umfassende Technikvorausschau zu nutzen, gefolgt. In Japan ist inzwischen bereits die 7. Delphi-Studie veröffentlicht worden (NISTEP 2001). Auf europäischer Ebene existiert mittlerweile ein Netzwerk der beteiligten Institutionen, um Informationen zu neuer Technologie, Technologiepolitik und auch Methodentheorie auszutauschen.

1996 startete das Delphi '98. Diese neuerliche Studie sollte eine Aktualisierung der Daten ermöglichen, denn in den letzten Jahren sind in Deutschland und in anderen hoch entwickelten Industrieländern einige Weichen für die Zukunft neu gestellt worden, so dass die zukünftigen Entwicklungsmöglichkeiten aus der heutigen Sicht womöglich anders zu beurteilen sind.

Zuerst wurde ein Lenkungsausschuss am BMBF gegründet, der zukunftsrelevante Themenfelder definiert hat. Die detaillierte inhaltliche Vorbereitung der Studie (Themenfindung usw.) fand in Fachausschüssen statt, die mit mehr als 100 fachkundigen Personen aus Industrie, Hochschulen und anderen Einrichtungen besetzt waren. Koordiniert wurde der gesamte Prozess vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI) in Karlsruhe. Von dort wurde auch die weitere Erarbeitung der Fragebögen, der Versand, der Aufbau einer Adress-Datenbank und die analytische Arbeit betreut. Der Originalbericht (Cuhls/Blind/Grupp 1998) steht für Interessierte zur Verfügung (www.isi.fhg.de/ti/projektbeschreibungen).

2.1 Die Themenfelder und ihre Inhalte

Die in Tabelle 1 aufgeführten Themenfelder wurden für so wichtig erachtet, dass ein Blick in ihre Zukunft lohnt. Für jedes Themenfeld wurde ein separater Fragebogen mit bis zu 114 Thesen erstellt. Insgesamt haben die zwölf Fragebögen 1.070 Thesen enthalten. Stärker als beim ersten Mal wurden beim Delphi '98 die deutschen Belange berücksichtigt. Um internationale Vergleiche zu ermöglichen, werden aber auch gemeinsame Themen mit dem sechsten japanischen Delphi, das parallel zu der deutschen Studie durchgeführt wurde, betrachtet. Eine genauere Analyse findet sich in Cuhls/Blind/Grupp (2001). Die jeweilige Anzahl der in Deutschland und in Japan identischen Thesen ist in der Tabelle 1 ebenfalls angegeben. Die Thesen sind in Japan allerdings zum Teil in anderen Themenfeldern zur Diskussion gestellt worden. Um auch beobachten zu können, ob sich bestimmte Einschätzungen in den letzten fünf Jahren geändert haben, wurden einige Thesen aus dem ersten deutschen Delphi-Report unverändert übernommen und noch einmal in die verschiedenen Fragebögen integriert.

Tabelle 1: Anzahl der Thesen in den Delphi'98 - Themenfeldern

Delphi'98 Themenfelder	Anzahl der Thesen im Delphi'98	gleiche Thesen in Japan	gleiche Thesen im Delphi 93
1. Information & Kommunikation	111	68	16
2. Dienstleistung & Konsum	78	21	6
3. Management & Produktion	71	14	7
4. Chemie & Werkstoffe	104	33	24
5. Gesundheit & Lebensprozesse	104	36	7
6. Landwirtschaft & Ernährung	101	20	7
7. Umwelt & Natur	76	15	2
8. Energie & Rohstoffe	114	32	14
9. Bauen & Wohnen	75	32	9
10. Mobilität & Transport	107	18	5
11. Raumfahrt	78	24	9
12. Großexperimente	51	10	7
Summe	1.070	323	113

2.2 Wie sahen die Fragebögen aus?

Die erarbeiteten Fragebögen umfassten Thesen, die durch Ankreuzen nach verschiedenen Kriterien zu beurteilen waren. Die erste Spalte, die im Fragebogen anzukreuzen war, betraf die Selbsteinschätzung der *Fachkenntnis* als „groß“ (bezeichnet mit Fk1), „mittel“ (Fk2), „gering“ (Fk3) oder „fachfremd“ (Fk4). Aufgrund dieser Selbstbewertung wurde der Wissensstand auf Basis der Antwortenden pro These berechnet. Personen, die sich als „fachfremd“ einstufen, wurden bei dieser Berechnung nicht einbezogen, da sie die Fragen nicht weiter zu beantworten hatten.

Die Thesen sollten dann nach den folgenden Kriterien von den Experten beurteilt werden. Zum Ersten wurde nach der Wichtigkeit der Lösungsbeiträge gefragt. Dabei wurde unterschieden nach der Bedeutung für die Erweiterung menschlichen Wissens, für die wirtschaftliche Entwicklung, für die gesellschaftliche Entwicklung, für die Lösung der ökologischen Probleme und für Arbeit und Beschäftigung. Als zweiter Punkt sollte der mögliche Zeitraum der Verwirklichung eingeschätzt werden, um festzulegen, ob es sich um einen langen oder eher kürzeren Zeitraum handelt, auf den sich die weitere Beurteilung bezieht, und um den zeitlichen Rahmen für Handlungsspielräume festzulegen. Der Zeithorizont war in Fünf-Jahres-Intervalle unterteilt, ergänzt um den Zeitraum bis zum Jahre 2000 und ein offenes Intervall nach 2025.

Die nächste Frage bezog sich auf den Stand der Forschung und Entwicklung. Welches Land wird als führend angesehen, die USA, Japan, Deutschland, ein anderes Land der EU oder ein anderes Land außerhalb der EU? Außerdem wurde danach gefragt, wie die Rahmenbedingungen eingeschätzt werden bzw. welche Maßnahmen zu ihrer Verbesserung angebracht sind. Als Optionen standen verbesserte Ausbildung, der Austausch von Personal zwischen Wissenschaft und Wirtschaft, die internationale Kooperation, die Verbesserung der Forschungsinfrastruktur, eine verbesserte staatliche Projektförderung oder Regulationsänderungen zur Auswahl. Schließlich wurde nach möglichen Folgeproblemen bei einer Realisierung der Visionen gefragt. Alternative Lösungsmöglichkeiten konnten ebenfalls genannt werden.

2.3 Wer wurde befragt?

Wer waren nun die in der Studie befragten Personen? Wer ist Experte für die Zukunft? Da niemand genau wissen kann, wie die Zukunft werden wird, sollten möglichst viele Personen beteiligt sein. Denn es ist erwiesen, dass bei einer großen Anzahl von Antworten individuelle Schätzfehler ausgemittelt werden können und damit die Wahrscheinlichkeit einer „treffsicheren Prognose“ größer ist.² Die „Treffsicherheit“ darf daher keinesfalls als Alleinkriterium für den „Erfolg“ einer Studie angesehen werden, sondern die Tatsache, dass Aktivitäten angestoßen werden (Kooperationen, neue Projekte etc.) ist gemessen an den Zielen der Studie als ebenso wichtig anzusehen.

Die Befragten mussten sich mit der These auseinander setzen können, um überhaupt ein Urteil abgeben zu können. Deshalb wurde versucht, eine Adress-Datenbank aufzubauen, die Personenkreise umfasst, welche sich auf den unterschiedlichen Gebieten auskennen. Diese Personen sollten möglichst zu jeweils einem Drittel aus der Industrie, aus Hochschulen sowie aus anderen Forschungseinrichtungen, dem öffentlichen Dienst und Verbänden stammen.

Um pro Themenfeld etwa 100 auswertbare Bögen zu erhalten bzw. sogar möglichst mindestens 100 Antworten pro Fragestellung, wurden zunächst ca. 7.000 Personen angeschrieben. Ihnen wurde angeboten, einen zusätzlichen Fragebogen anzufordern, falls sie sich auch für ein anderes Gebiet kompetent fühlten, oder ihren Bogen gegen einen anderen auszutauschen. Insgesamt haben über 2.000 Personen 2.453 Bögen ausgefüllt (Tabelle 2). Dies ist ein erfreuliches Ergebnis, da ein auswertbarer Rücklauf von mehr als 30 % (in einigen Gebieten des Delphi '98 sind es sogar fast 50 %) die

² Dies wurde bereits früh von Dalkey, Helmer und anderen in den USA getestet (siehe z.B. Dalkey 1969a und b; Dalkey/Brown/Cochran 1969; Dalkey 1968; Dalkey/Helmer 1963).

Erfahrungswerte in Deutschland übertrifft. In einigen Feldern, z.B. „Chemie und Werkstoffe“, „Gesundheit und Lebensprozesse“, „Energie und Rohstoffe“ sowie „Information und Kommunikation“, sind deutlich mehr als 200 Antworten eingegangen. In Feldern wie „Großexperimente“ oder „Raumfahrt“ dagegen konnten so viele Antworten nicht erwartet werden, denn die Zahl der Forscher liegt auf diesen Gebieten in Deutschland niedriger als auf anderen. In der folgenden Tabelle findet sich die Anzahl der in der ersten Runde eingegangenen Bögen und damit in der zweiten Runde beteiligten Experten. Diejenigen Personen, die anonym geantwortet haben, konnten leider an der zweiten Runde nicht mehr beteiligt werden. Der Rücklauf am Ende der zweiten Runde findet sich in der rechten Spalte der Tabelle 2 und beträgt ca. 75 % der in der ersten Runde eingegangenen Bögen.

Tabelle 2: Anzahl der auswertbaren Bögen nach Themenfeldern

Delphi Themenfelder	Anzahl der auswertbaren Fragebögen	
	1. Runde	2. Runde
1. Information & Kommunikation	287	206
2. Dienstleistung & Konsum	215	163
3. Management & Produktion	229	179
4. Chemie & Werkstoffe	260	206
5. Gesundheit & Lebensprozesse	255	182
6. Landwirtschaft & Ernährung	206	140
7. Umwelt & Natur	282	209
8. Energie & Rohstoffe	246	187
9. Bauen & Wohnen	110	94
10. Mobilität & Transport	150	122
11. Raumfahrt	97	77
12. Großexperimente	116	91
Gesamt	2.453	1.856

2.4 Die Expertise der Beteiligten

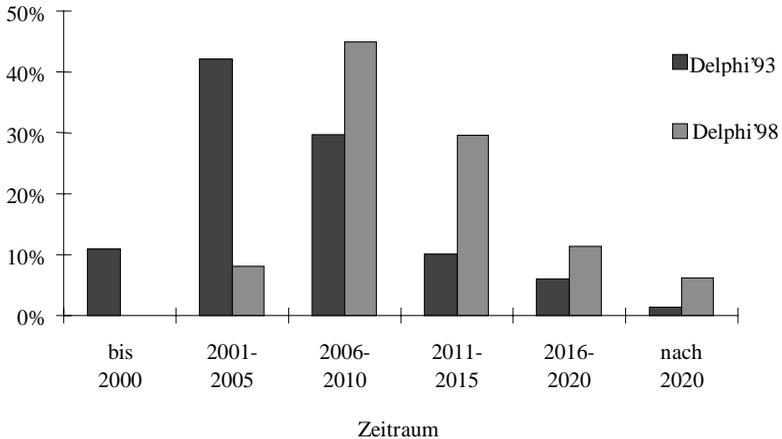
Bei der Auswahl der angeschriebenen Experten war das Ziel, viele Personen mit einer möglichst hohen Expertise in den jeweiligen Bereichen zu identifizieren. Jedoch sollten auch Personen mit so genannter „sekundärer“ Fachkenntnis in den Kreis der Befragten aufgenommen werden, um ein gewisses korrigierendes Gegengewicht zu den Wissenschaftlern mit höchster Expertise zu erhalten.

Im Endergebnis haben sich durchschnittlich 10 % der Antwortenden als Experten mit höchster Fachkenntnis eingestuft, d.h. sie arbeiten auf dem betreffenden Gebiet. Weitere 30 % haben mittlere (d.h. frühere Arbeit auf dem Gebiet und/oder Lesen der Primärliteratur) und 60 % geringe Fachkenntnis (Lesen der Sekundärliteratur bzw. im Gespräch mit Fachleuten stehend) angegeben. Dies mag auf den ersten Blick unbefriedigend erscheinen, jedoch gilt es zu bedenken, dass die zwölf Fragebögen eine so große inhaltliche Spannbreite umfassen, dass der einzelne Experte lediglich bei sehr wenigen Thesen glaubwürdig eine *hohe* Fachkenntnis (arbeitet auf dem Gebiet) bekunden kann. Ferner ist anzumerken, dass der Anteil der Experten mit höchster Fachkenntnis in der ersten Runde noch bei 13,5 % lag. Es haben sich wohl mehrere „Fachkenner“ in der zweiten Runde nicht mehr beteiligt bzw. in ihrer Selbsteinschätzung zurückgestuft. Eine grundsätzliche Tendenz zur Überschätzung der eigenen Fachkenntnis konnte auf der Basis einzelner Fragebögen nicht beobachtet werden und spiegelt sich daher auch im aggregierten Ergebnis nicht wider.

2.5 Erwarteter Realisierungszeitraum der Visionen

Die Verteilung der Realisierungszeiträume der Thesen unterscheidet sich gemäß Abbildung 1 in ihrer generellen Form nicht von der Verteilung bei der ersten Delphi-Studie. Einzige Ausnahme ist, dass sie sich um etwa fünf Jahre in die Zukunft verschoben hat. Nahezu die Hälfte der zur Einschätzung vorgelegten Thesen kann nach Meinung der befragten Experten im Zeitraum zwischen 2006 und 2010 realisiert werden, ein weiteres Drittel zwischen 2011 und 2015. Bis zum Jahr 2000 sollte es nach der durchschnittlichen Einschätzung der Experten noch *keine einzige* Realisierung geben. Aber auch nach 2025 werden nicht einmal 30 Visionen eingestuft. Damit ist der angestrebte Zeithorizont bis 2025 bei der Mehrzahl der Thesen eingehalten worden.

Während zwischen den Experten in Unternehmen und Hochschulen kein Unterschied in der Beurteilung des Realisierungszeitraums auszumachen ist, zeigen sich geringe Differenzen zwischen den verschiedenen „Expertentypen“. So sehen die auf Basis der Einschätzungen allgemeiner Megatrends bestimmten Umweltpessimisten und Bevölkerungsoptimisten die Innovationen zwischen 2007 und 2015 als realisierbar an, während die Fortschrittsskeptiker ihre Verwirklichung im Durchschnitt ein Jahr später erwarten - vor dem Hintergrund eines Zeithorizontes von 30 Jahren kein großer Unterschied (Blind/Cuhls/Grupp 2001). Ferner neigen die Antwortenden mit hoher Expertise dazu, die Realisierung der Thesen früher einzuschätzen.

Abbildung 1: Zeitraumverteilungen Delphi'93 und Delphi'98

Da die durchschnittlichen Realisierungszeiträume von der Zusammensetzung der Thesen hinsichtlich der Innovationsstufen nicht ganz unabhängig sind, soll an dieser Stelle auch hierauf eingegangen werden. Die Thesen in den Fragebögen enthalten stereotype Wörter, die zwischen wissenschaftlichem, technischem oder wirtschaftlichem Fortschritt unterscheiden. Es dominieren mit nahezu 40 % die Thesen, die sich auf bereits vermarktete, aber noch nicht weit verbreitete Produkte und Dienstleistungen beziehen. Weitere 30 % umfassen Visionen zur ersten wirtschaftlichen Anwendung bereits entwickelter und als Prototyp vorliegender Technologien. Weitere 18 % haben die Entwicklung von Prototypen zum Gegenstand, und lediglich etwas mehr als 11 % der Thesen sind der Grundlagenforschung zuzurechnen. Damit ist im Vergleich zur ersten deutschen Delphi-Studie eine höhere Anwendungsorientierung zu bemerken.

3. Test von Hypothesen zur Rolle des Expertenstatus

Bevor die Hypothesen zur Rolle des Expertenstatus getestet werden, muss eine Reihe methodischer Vorbemerkungen gemacht werden. Die befragten Experten im Bereich von Wissenschaft und Technologie wurden auf der Basis von öffentlich zugänglichen Datenbanken und durch Nennungen von anderen Experten ausgewählt. Damit sollte die Durchführung einer Expertenbefragung sichergestellt werden. Jedoch wurden

verschiedene Wissensgebiete in zwölf Fragebögen zusammengefasst, um der Interdisziplinarität nahe stehender Gebiete Rechnung zu tragen, um ausreichende Samples zu gewährleisten und um Kosten zu sparen. Solche, mehrere Themenfelder umfassende Fragebögen, bedeuten für die Experten, dass sie nur in ihren ausgewählten Teilbereichen eine hohe Expertise besitzen, während sie in den angrenzenden Randbereichen eine geringere und in noch entfernteren Feldern unter Umständen gar kein Expertenwissen mehr besitzen. Die Angaben der Experten zu ihrer Expertise bezüglich einer bestimmten Vision beruhen allein auf ihrer Selbsteinschätzung, die mit objektivierbaren Kriterien im Einzelfall nicht überprüfbar ist.³ Deshalb muss generell davon ausgegangen werden, dass diese Selbsteinschätzung korrekt vorgenommen wurde und zwischen den Experten eine gewisse Übereinstimmung besteht.

Methodische Konsequenz aus diesem Sachverhalt ist, dass den befragten Experten kein genereller Expertenstatus zugeordnet und deshalb auch nicht danach segmentiert werden kann. Die Analyse wird bei der Beobachtungseinheit deshalb nicht auf die ungefähr 2.000 Experten zurückgreifen, sondern auf deren knapp 100.000 Antworten aus den beiden Runden der Delphi-Befragung bzw. ca. 65.000 Antworten von Experten, die in der ersten und zweiten Runde teilgenommen haben.

In der Vergangenheit wurde die Panelmortalität bei Delphi-Untersuchungen in einem engen Zusammenhang mit der Expertise der Befragten gesehen. Schon Bardecki (1984) konnte nicht bestätigen, dass Experten mit hoher Expertise bzw. starkem ego-involvement mit einer geringeren Wahrscheinlichkeit aus einem Delphi-Prozess ausscheiden. Häder (2000) konnte in seiner Untersuchung dieser Fragestellung auch keine signifikanten Unterschiede zwischen der Abbruchwahrscheinlichkeit und dem Expertenstatus der Befragten ausmachen. Im Kontext dieser Untersuchungen ist die folgende erste Hypothese zu sehen.

Hypothese 1: Experten mit geringem Expertenwissen werden mit geringerer Wahrscheinlichkeit an der zweiten Befragungsrunde teilnehmen. Hypothese 1 lässt sich dadurch begründen, dass mit zunehmender Expertise das Interesse an einer Einflussnahme und den Endergebnissen selbst ausgeprägter ist, weil sich die Fachleute davon unter Umständen auch neue Erkenntnisse für die eigene zukünftige Forschung versprechen. Ferner dürften Teilnehmer bei zu geringer Expertise befürchten, eine „falsche“ Antwort zu geben, und deshalb eher in der zweiten Runde nicht mehr teilneh-

³ Es wurde jedoch geprüft, ob einzelne Experten in allen von ihnen ausgefüllten Fragen die höchste Expertise angegeben haben. Da dies aus den oben geschilderten Gründen nicht möglich ist, wären solche Antworten aus dem Datensatz entfernt worden. Dieser Fall kam aber nicht vor.

men. Daher wird im Folgenden geprüft, ob Experten mit geringerem Fachwissen tatsächlich in größerer Zahl von einer Teilnahme an der zweiten Runde absehen.

Das Gesamtsample wurde differenziert in diejenigen Antworten, die nur in der ersten Runde, und diejenigen, die auch in der zweiten Runde abgegeben wurden. Abbildung 2 stellt die Verteilung der Expertise nach Themenfeldern basierend auf den Antworten zu einzelnen Thesen in der ersten Runde dar. Als Test-Methode wurde der Chi-Quadrat nach Pearson verwendet. Tabelle 3 stellt das aggregierte Ergebnis über alle zwölf Themenfelder und das Ergebnis des Chi-Quadrat-Tests dar. Es wird deutlich, dass die Durchschnittsexpertise bei den Antworten der Abbrecher unter derjenigen liegt, die sich auch an der zweiten Runde beteiligt haben.

Eine getrennte Durchführung des Tests nach den zwölf Themenfeldern ergibt jedoch differenzierte Ergebnisse. Während in den Themenfeldern „Information & Kommunikation“, „Dienstleistung & Konsum“, „Management & Produktion“, „Gesundheit & Lebensprozesse“, „Umwelt & Natur“, „Bauen & Wohnen“, „Mobilität & Transport“ und „Großexperimente“ ein negativer Zusammenhang zwischen der Expertise und der Abbruchwahrscheinlichkeit besteht, haben in den Themenfeldern „Chemie & Werkstoffe“, „Energie & Rohstoffe“ und „Raumfahrt“ an der zweiten Runde weniger Personen mit hoher Expertise teilgenommen. Hier könnten sich die Experten mit hoher Expertise an den Inhalten, nach methodischen Aspekten oder nach Analyse der Ergebnisse der ersten Runde dazu entschieden haben, sich an der zweiten Runde nicht mehr zu beteiligen.⁴ Im Themenfeld „Landwirtschaft & Ernährung“ konnte keinerlei Zusammenhang zwischen Expertise und Abbruch-Wahrscheinlichkeit festgestellt werden.

⁴ Hier kann auch Zufall nicht ausgeschlossen werden, da viele Experten einfach aufgrund von Zeitmangel/Arbeitsüberlastung den zweiten Bogen nicht mehr ausfüllen konnten.

Abbildung 2: Verteilung der Experten-Anteile in den Themenfeldern differenziert nach Abbrechern (2. Zeile) und Nicht-Abbrechern (1. Zeile)

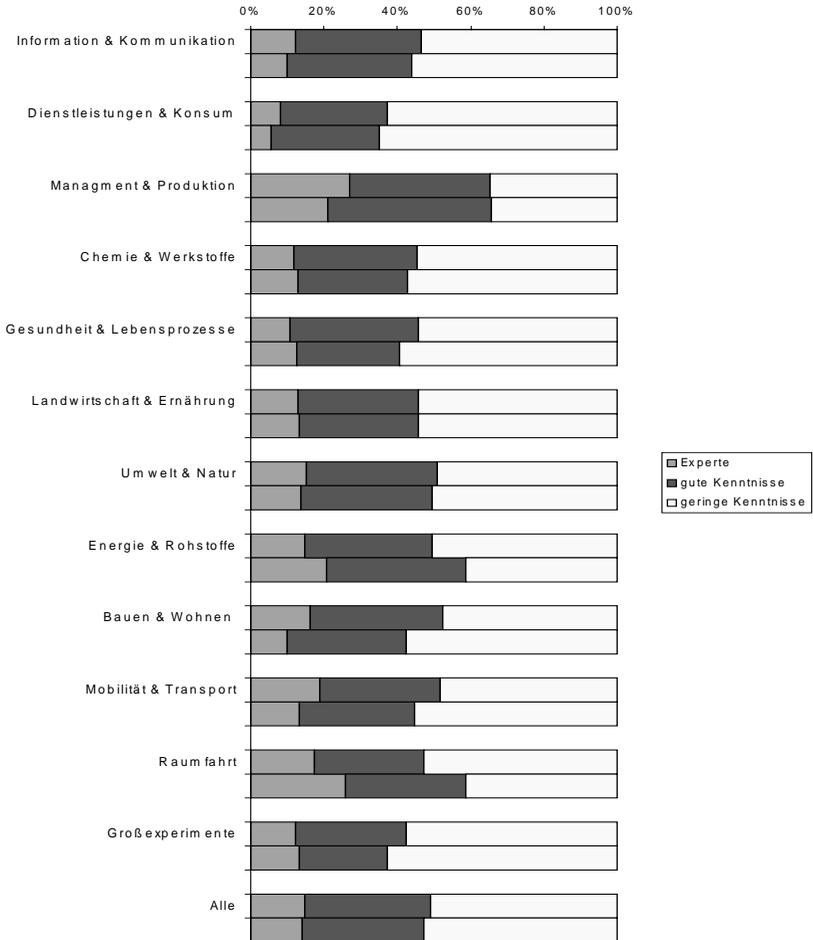


Tabelle 3: Zusammenhang zwischen Abbruch und Grad der Expertise in der ersten Runde⁵

Themenfelder	Ergebnis des Chi-Quadrat-Tests	Anzahl der Beobachtungen
1. Information & Kommunikation	---	11.201
2. Dienstleistung & Konsum	---	8.352
3. Management & Produktion	---	9.013
4. Chemie & Werkstoffe	++	7.901
5. Gesundheit & Lebensprozesse	---	11.809
6. Landwirtschaft & Ernährung	?	8.011
7. Umwelt & Natur	-	9.028
8. Energie & Rohstoffe	+++	12.239
9. Bauen & Wohnen	---	3.660
10. Mobilität & Transport	---	7.970
11. Raumfahrt	+++	3.348
12. Großexperimente	--	1.816
Gesamt	---	94.348

Obwohl in zwei Dritteln der Themenfelder die Abbruchwahrscheinlichkeit mit fallender Expertise steigt, müssen in diesem Zusammenhang noch zwei weitere Hypothesen für einen Abbruch genannt werden (Häder 2000). So besagt die Dissonanz-Hypothese (Festinger 1978), dass Aussteiger die Sachverhalte systematisch anders bewerten und diese kognitiven Dissonanzen ein Abbruchmotiv darstellen. In dieselbe Richtung, jedoch mit umgekehrter Kausalität, argumentieren die Befürworter der Nonkonformitäts-Hypothese, dass Aussteiger aufgrund ihrer Distanz zum Thema extremere Urteile abgeben als der Durchschnitt der Experten. Die Stärke der Dissonanz hängt allerdings von verschiedenen Faktoren ab, ebenso wie die Reaktionen der Einzelpersonen. So haben diese nach Bardecki (1984) die verschiedenen Möglichkeiten, 1. den psychologischen Anker des zurückgespielten Mittelwertes oder Medians zu ignorieren, 2. sich bewusst in Kontrast zum psychologischen Anker zu stellen oder 3. seine Meinung in Richtung des psychologischen Ankers zu verändern (Assimilation). Generell kann nur gesagt werden „agreeing with people decreases dissonance, disagreeing increases it“ (Bardecki 1984, S. 283). Wie die Delphi-Experten im Einzel-

⁵ Abbruch ist einfach dadurch definiert, dass der Experte an der zweiten Runde nicht mehr teilgenommen hat. Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; ++/-- < 0,05, +/- < 0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

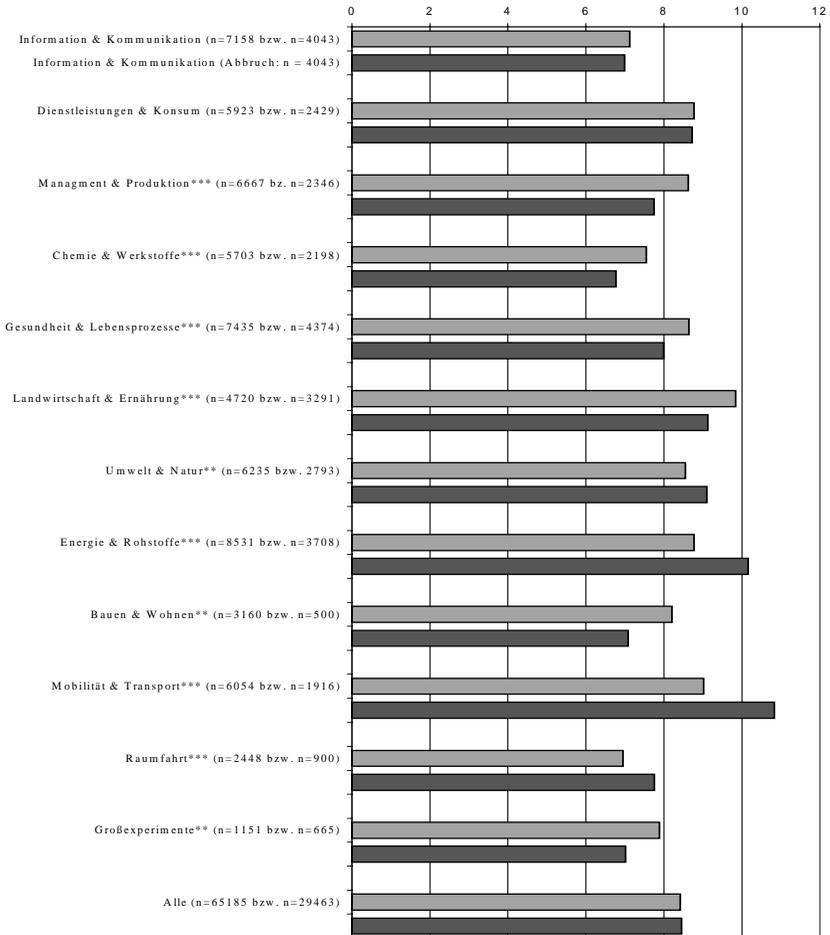
nen reagieren, kann daher nicht pauschalisiert werden, sondern scheint sich nach der Einzelperson zu richten.

Extreme Einschätzungen lassen sich im Delphi'98 am besten an den Antworten zum erwarteten Realisierungszeitraum ablesen, während die Skalen zu den anderen Fragen in der Regel nur Ja/Nein-Optionen erlauben. Die Differenz zwischen der Einzeleinschätzung bezüglich des Realisierungszeitraums und des Medianwertes wurde als Absolutwert berechnet. In Abbildung 3 sind die durchschnittlichen Differenzen der Abbrecher bzw. Nicht-Abbrecher von den jeweiligen Medianen differenziert nach Themenfeldern angegeben. Während im Aggregat aller Antworten kein Unterschied beim Mittelwert der Abweichungen festzustellen ist, sind in zehn der zwölf Themenfelder signifikante Unterschiede auszumachen. Jedoch halten sich die Themenfelder, in denen die Abbrecher extremere Urteile gefällt haben, und die Themenfelder, in denen die Nicht-Abbrecher stärker vom Median abweichen, die Waage. Dieses Ergebnis unterstreicht den Bedarf an weiteren Untersuchungen zur Abbruchwahrscheinlichkeit bei Delphi-Untersuchungen. Unmittelbar an den Zusammenhang zwischen Abbruchwahrscheinlichkeit und Extremurteilen schließt sich der in Hypothese 2 unterstellte Zusammenhang an.

Hypothese 2: Je höher die Expertise, desto größer sind die Unterschiede zwischen der eigenen Einschätzung und der mittleren Einschätzung. Bei dieser Hypothese wird davon ausgegangen, dass Experten vorgefasste Meinungen haben, die gegebenenfalls extrem sein können und von denen sie nicht abweichen wollen.⁶ Dies trifft auf ein allgemein verbreitetes Vorurteil gegenüber Wissenschaftlern (diese seien in ihr Thema „verliebt“ und ließen sich selten von ihrer Meinung abbringen) und bezieht sich bezüglich Delphi'98 auf die zweite, bei anderen Delphi-Untersuchungen eventuell auch auf spätere Runden. Wie oben schon erwähnt, lassen sich extreme Einschätzungen am besten an den Antworten zum erwarteten Realisierungszeitraum ablesen. Für die Überprüfung der Hypothese wurde die Differenz zwischen der Einzeleinschätzung und des Medianwertes als Absolutwert ermittelt und mit der jeweiligen Expertise korreliert.

⁶ Ein unfreiwilliger Test ergab sich diesbezüglich im Mini-Delphi von 1995, bei dem in einem Fragebogen durch einen technischen Fehler falsche Daten zu Thesen zurückgespielt wurden. Die meisten Experten korrigierten ihre Meinung in Richtung der falschen Angabe. Nur einige der Fk1-Experten bestanden vehement auf ihrer vorherigen Meinung und wunderten sich ob der seltsamen Ergebnisse (Kommentare, Cuhls/Breiner/Grupp 1995).

Abbildung 3: Mittelwerte der Abweichungen von den Medianen in den Themenfeldern differenziert nach Abbrechern (2. Zeile) und Nicht-Abbrechern (1. Zeile)⁷



⁷ Abbruch ist einfach dadurch definiert, dass der Experte an der zweiten Runde nicht mehr teilgenommen hat. Die Anzahl der Sternchen repräsentiert das Signifikanzniveau: ***<0,01; **<0,05.

Tabelle 4: Zusammenhang zwischen der Abweichung von der Durchschnittseinschätzung und dem Grad der Expertise⁸

Themenfelder	Signifikanzniveaus des Spearman-Rho-Tests						
	Abweichung in Runde 1 mit Expertise in Runde 1 (mit Abbrecher)	Abweichung in Runde 1 mit Expertise in Runde 1 (ohne Abbrecher)	Anzahl der Beobachtungen	Abweichung in Runde 2 mit Expertise in Runde 2	Abweichung in Runde 1 mit Durchschnitts-Expertise	Abweichung in Runde 2 mit Durchschnitts-Expertise	Anzahl der Beobachtungen
1. Information Kommunikation	---	?	11.201	+++	--	?	7.158
2. Dienstleistung Konsum	---	---	8.352	?	--	?	5.923
3. Management Produktion	---	?	9.013	?	?	+	6.667
4. Chemie Werkstoffe	---	--	7.901	+++	?	+++	5.703
5. Gesundheit Lebensprozesse	---	?	11.809	+++	---	+	7.435
6. Landwirt-Ernährung	---	---	8.011	?	---	?	4.720
7. Umwelt Natur	---	?	9.028	?	--	-	6.235
8. Energie Rohstoffe	---	+++	12.239	+++	+	?	8.531
9. Bauen Wohnen	?	++	3.660	?	?	?	3.160
10. Mobilität & Transport	---	--	7.970	?	?	?	6.054
11. Raumfahrt	---	?	3.348	+++	--	?	2.448
12. Großexp.	?	?	1.816	?	?	?	1.151
Gesamt (Pearson)⁹	+++	?	94.348	+++	---	?	65.185

⁸ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; +/--- <0,05, +/- <0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

⁹ Aufgrund der hohem Fallzahl erlaubt SPSS die Berechnung des den Daten angemessenen Spearman-Rho-Tests nicht.

In Tabelle 4 sind die Signifikanzniveaus der Korrelationsanalyse dokumentiert, wobei zum einen zwischen den Unterschieden in der ersten und zweiten Befragungsrunde differenziert wurde. Ferner wurden in den Analysen eine Durchschnittsexpertise basierend aus den Angaben aus der ersten und zweiten Runde berechnet und mit den Unterschieden korreliert, weil – wie oben schon ausgeführt – die Experten dazu tendierten, ihre in der ersten Runde postulierte Expertise nach unten zu revidieren. Wer sich als wahrer Experte (Fk1) betrachtet, hat demnach häufiger das Standing, auch bei einer anderen von der Mehrheitsmeinung abweichenden Zeiteinschätzung zu bleiben.

Die Ergebnisse zeigen auf den ersten Blick ein völlig heterogenes Bild. Während für die Antworten in der ersten Runde Hypothese 2 eindeutig widerlegt werden kann, weil die Abweichungen der individuellen Einschätzungen mit abnehmender Expertise in den meisten Themenfeldern signifikant zunehmen, gilt dies für die Antworten in der zweiten Runde nicht mehr. Hier kann aber auch für kein Technikfeld die Gegenhypothese zu Hypothese 2 bestätigt werden. Dagegen kann in der Hälfte der Themenfelder eine signifikante Bestätigung der Hypothese 2 nachgewiesen werden. Hier nimmt mit zunehmender Expertise der Abstand zur Median-Meinung zu. Zwei Effekte können hier wohl als Gründe für die Bestätigung der Hypothese herangezogen werden. Zum einen haben wir gesehen, dass viele Experten – angesichts des sehr niedrigen Anteils an ausgewiesenen Experten – ihre Selbsteinschätzung bezüglich ihrer Expertise revidiert und in vielen Fällen reduziert haben. Dies bedeutet, dass nur noch die „wahren“, auf dem Gebiet arbeitenden Experten das Selbstbewusstsein haben, diese Kategorie im Fragebogen auch anzukreuzen. Zum Zweiten haben die Experten nun die Durchschnittsmeinung ihrer Kollegen kennen gelernt. Auch hier bedarf es eines ausgeprägten Selbstbewusstseins, um sich weiterhin signifikant von der Durchschnittsmeinung zu unterscheiden. Das Zusammenspiel dieser beiden Effekte kann somit erklären, dass es in vielen Themenfeldern doch noch zu einer Bestätigung von Hypothese 2 kommt. Zur Information sind auch noch die Signifikanzniveaus der Korrelationsanalyse der Differenzen mit der Durchschnittsexpertise aus beiden Runden dokumentiert. Angesichts der oben beschriebenen Effekte ergeben sich deshalb lediglich für die Einschätzungen der zweiten Runde in manchen Themenfeldern die erwarteten Ergebnisse.

Hypothese 3: Je höher die Expertise, desto geringer sind die Unterschiede zwischen den Einschätzungen in der 1. und der 2. Runde. Aufgrund der bereits zu Hypothese 2 gemachten Erläuterungen kann davon ausgegangen werden, dass die Antwortenden mit einer hohen Fachkenntnis nur in einem geringeren Umfang ihre in der ersten Befragungsrunde abgegebenen Einschätzungen revidieren werden. Ähnlich wie bei

Häder (2000) liegt Hypothese 3 die Annahme zu Grunde, dass Personen, die auf dem Gebiet arbeiten, ein gesundes Selbstbewusstsein hinsichtlich ihrer (Eigen-) Einschätzung besitzen.

Dazu wurden die Ergebnisse der 1. und 2. Runde hinsichtlich einer Veränderung bei der Angabe zur „Fachkenntnis“ überprüft. Die in Tabelle 5 dargestellten Ergebnisse machen deutlich, dass Hypothese 3 in knapp der Hälfte der Themenfelder bestätigt werden kann. Es herrscht daher die Tendenz, dass diejenigen Personen mit geringer Expertise ihre Einschätzungen stark revidieren (d. h. der Mehrheitsmeinung anpassen), während die „wahren“ Experten nur in sehr geringem Ausmaß ihre Einschätzung der ersten Runde verändern.

Tabelle 5: Zusammenhang zwischen der Veränderung der eigenen Einschätzung und dem Grad der Expertise¹⁰

Themenfelder	Ergebnisse des Spearman-Rho-Tests mit Expertise der 2. Runde	Ergebnisse des Spearman-Rho-Tests mit Durchschnittsexpertise	Anzahl der Beobachtungen
1. Information & Kommunik.	++	+++	7.158
2. Dienstleistung & Konsum	++	++	5.923
3. Management & Produktion	?	?	6.667
4. Chemie & Werkstoffe	?	?	5.703
5. Gesundheit & Lebensprozesse	+++	+	7.435
6. Landwirtschaft & Ernährung	?	++	4.720
7. Umwelt & Natur	+++	++	6.235
8. Energie & Rohstoffe	?	-	8.531
9. Bauen & Wohnen	+++	+	3.160
10. Mobilität & Transport	?	?	6.054
11. Raumfahrt	?	?	2.448
12. Großexperimente	?	?	1.151
Gesamt	?	?	65.185

¹⁰ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/--- < 0,01; ++/-- < 0,05, +/- < 0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

Hypothese 4: Je höher die Expertise, desto geringer ist die Konvergenz (Differenz zwischen der Abweichung vom Median in der ersten Runde und der Abweichung vom Median in der zweiten Runde). In engem Zusammenhang mit Hypothese 3 steht Hypothese 4, wonach die Konvergenz, d. h. die Annäherung an die Median-Meinung in der zweiten Befragungsrunde, mit zunehmender Fachkenntnis schwächer ausgeprägt ist.

Tabelle 6: Zusammenhang zwischen der Konvergenz und dem Grad der Expertise¹¹

Themenfelder	Ergebnisse des Spearman-Rho-Tests mit Expertise der 2. Runde	Ergebnisse des Spearman-Rho-Tests mit Durchschnittsexpertise	Anzahl der Beobachtungen
1. Information & Kommunik.	---	?	7.158
2. Dienstleistung & Konsum	?	---	5.923
3. Management & Produktion	-	?	6.667
4. Chemie & Werkstoffe	?	?	5.703
5. Gesundheit & Lebensprozesse	---	---	7.435
6. Landwirtschaft & Ernährung	?	?	4.720
7. Umwelt & Natur	---	?	6.235
8. Energie & Rohstoffe	---	+	8.531
9. Bauen & Wohnen	--	?	3.160
10. Mobilität & Transport	?	?	6.054
11. Raumfahrt	---	--	2.448
12. Großexperimente	?	?	1.151
Gesamt	---	---	65.185

Auch hierzu wurden die Ergebnisse der ersten mit denen der zweiten Runde verglichen. Tabelle 6 zeigt eindrücklich, dass Hypothese 4 für zwei Drittel der Themenfelder bestätigt werden kann. Dies bedeutet, dass die Befragten in ihren Feldern mit hoher Expertise sich dem der Delphi-Methode typischen Konvergenz-Prozess nicht in dem Ausmaß unterwerfen, wie es für die Experten mit mittlerer oder geringer Kenntnis der Fall ist. Dieses Ergebnis ruft unmittelbar einen der Methode immanenten Widerspruch hervor. Zum einen will man für Delphi-Befragungen im Bereich von Wissenschaft und Technik möglichst viele Antwortenden gewinnen, die sich durch eine hohe Expertise

¹¹ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; ++/-- < 0,05, +/- < 0,1.

auszeichnen. Andererseits deuten die Ergebnisse bezüglich des Tests von Hypothese 4 darauf hin, dass sich gerade die „wahren“ Experten in ihren speziellen Fachgebieten durch eine gewisse Resistenz gegenüber dem Konvergenzprozess auszeichnen und dadurch das Ziel einer mehrstufigen Delphi-Befragung unterlaufen.¹² Dies unterstreicht die neueren Ziele, Delphi-Befragungen nicht mehr anzuwenden, um Konsens herbeizuführen, sondern um festzustellen, ob bereits ein Konsens besteht oder nicht (Cuhls/Blind/Grupp 1998). Dies war insbesondere im Delphi'98 eine der wichtigen Informationen für diejenigen Akteure, welche die Studie für sich selbst ausgewertet haben.

Abbildung 4 bestätigt die Ergebnisse des Hypothesentest. Denn mit zunehmender Fachkenntnis¹³ für die einzelnen Themenfelder nimmt die Varianz (Differenz zwischen dem 3. und dem 1. Quartil) bei der Einschätzung der Realisierungszeiträume zu. Ausnahme ist hier das Feld „Management & Produktion“.

Angesichts dieser Ergebnisse ist es noch wichtiger, die Gesamtergebnisse von Delphi-Befragungen differenziert nach der Expertise der Beteiligten auszuweisen. Es wäre natürlich auch überprüfenswert, ob und wie die Experten ihre Meinungen bei Einzelthesen geändert haben, zum Beispiel, inwieweit die Fk1-Experten im Einzelnen bzw. gesamt konvergieren. Dies kann und darf aufgrund der geringen Fallzahlen nicht ausreichend nachgewiesen werden.

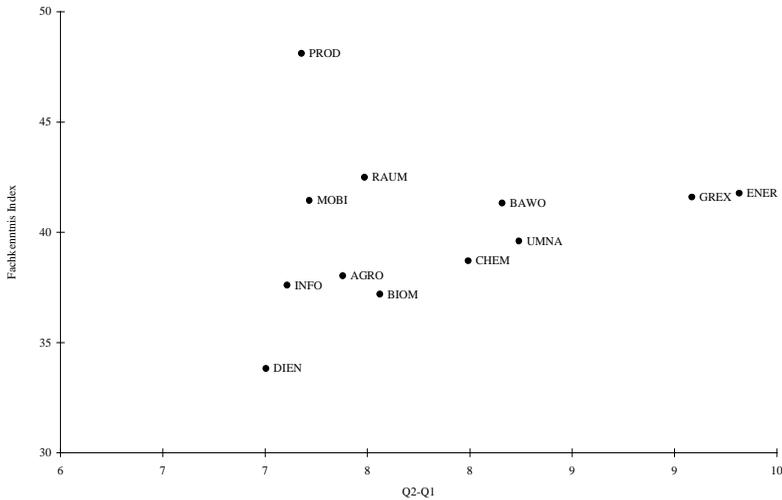
4. Zusammenfassung und Schlussfolgerungen

Delphi '98 hatte nicht zum Ziel, methodisch und statistisch genaue Trennungen der Fachexperten sozialwissenschaftlich zu untersuchen. Da sich aber die einmalige Gelegenheit eines großen Samples bietet (bisher sind nur sehr kleine Samples für Begründungen und Tests herangezogen worden), wurde der Versuch einer detaillierten Prüfung des Expertenverhaltens unternommen.

¹² Theoretisch könnte man auch das Konvergenzverhalten der Experten bezüglich deren mittlerer Einschätzung untersuchen. Jedoch ist die Anzahl der abgegebenen Einschätzungen insbesondere in der 2. Runde bezüglich einzelner Thesen so gering, dass die Berechnung des Medians oder Mittelwertes nicht gerechtfertigt ist.

¹³ Der Fachkenntnisindex berechnet sich aus den Anteilen der Fk1-, Fk2- und Fk3-Experten, wobei die Fk1-Experten mit dem Gewicht 4 und die Fk2-Experten mit dem Gewicht 2 in die Indexberechnung eingehen, so dass Wert 100 bei einem 100%-Anteil der Fk1-Experten und der Wert 25 bei einem 100%-Anteil der Fk3-Experten erreicht wird.

Abbildung 4: Zusammenhang zwischen Fachkenntnis und Varianz (Differenz zwischen dem ersten und dritten Quartil)



Die vier für die Delphi-Methodik grundlegenden Hypothesen konnten nur teilweise bestätigt werden. Das heißt, die der Methodik zu Grunde liegenden Annahmen müssen im Einzelfall sehr sorgfältig überprüft werden. Hypothese 1 konnte nur in einigen Themenfeldern nachgewiesen werden. Es ist also nur bedingt anzunehmen, dass die Experten mit geringem Wissensstand häufiger zu sogenannten Drop-outs werden und sich an der zweiten Runde nicht mehr beteiligen. Es scheinen also diverse Gründe eine Rolle zu spielen, sich weiter zu beteiligen oder nicht. Die Expertise ist nur eine Möglichkeit unter mehreren. Damit unterstreicht dieses Ergebnis generell das Funktionieren der Delphi-Methodik, weil es nicht zu einem systematischen Ausscheiden der Teilnehmer mit geringer Expertise kommt.

Auch die zweite Hypothese kann nur bedingt belegt werden. Während in der ersten Runde kein Beleg erfolgt, kann für die zweite Runde konstatiert werden, dass die Experten mit hohem Kenntnisstand stärker das Selbstbewusstsein haben, zu ihrer Meinung zu stehen, wenn diese vom mittleren Wert abweicht. In diesem Zusammenhang steht auch die Hypothese 3, die in knapp der Hälfte der Themenfelder belegt werden konnte: In diesen Feldern revidierten die Personen mit geringer und mittlerer

Expertise ihre Meinungen stark, während die „wahren“ Experten ihre Einschätzungen kaum veränderten.

Dieses Ergebnis hat für neue Vorausschau-Experimente durchaus Konsequenzen, bedeutet es doch, dass eine getrennte Ausweisung von Ergebnissen sehr wichtig ist, weil ein Bias in den Antworten nicht ausgeschlossen werden kann, dass aber auch die Funktion von „Laien“ oder Personen mit mittlerer Expertise darin bestehen kann, ein Korrektiv darstellen zu können.

Für zwei Drittel der Themenfelder bestätigt sich sogar Hypothese 4: Experten mit besonders hohem Wissensstand scheinen sich dem Delphi-typischen Konvergenzprozess nicht in dem Ausmaß zu unterwerfen wie andere Personenkreise. Dies hat weit reichende Konsequenzen. Wollte man bisher für Delphi-Befragungen immer die Personen mit möglichst hoher Expertise gewinnen, so zeigt dieses Ergebnis, dass dies nicht immer notwendig ist, sondern in Fällen, in denen hochgradige Expertise notwendig ist (z. B. in bestimmten Technologiebereichen) für die Vorausschau besser auf Einfachbefragungen, Interviews oder andere Verfahren zurückgegriffen werden sollte. Wenn das Ziel also Konsensbildung heißt, sollten demnach nicht nur die kenntnisreichen Experten, sondern möglicherweise auch Laien („Betroffene“) oder Personen mit geringerem Kenntnisstand hinzugezogen werden, da diese eher konsensbereit scheinen. Der Inhalt des Themas muss jedoch Berücksichtigung finden, da zu vermuten ist, dass in einem kontroversen Gebiet wie der Energieversorgung auf allen Seiten eine geringe Konsensbereitschaft anzutreffen sein wird.

Lautet das Ziel einer Delphi-Studie jedoch festzustellen, ob es überhaupt einen Konsens im Thema gibt, ist eine der wichtigen Informationen, ob die Expertenwelt im Vergleich zur „Laienwelt“ oder den „Anwendern“ eine andere Meinung vertritt und diese auch in der zweiten Runde nicht ändert. Gleichzeitig bestätigt das Ergebnis das Vorurteil, dass (insbesondere technische) Experten nur schwer von ihrer einmal vorgefassten Meinung abzubringen sind. Dies kann insbesondere für die Forschungspolitik im Bereich der angewandten Forschung eine sehr wichtige Information sein und bestätigt indirekt den Ansatz des Bundesministeriums für Bildung und Forschung, im neuen Futur-Prozess nicht mehr nur Experten zu Rate zu ziehen, sondern in Diskursen auch Personen mit geringerem Vorwissen teilnehmen zu lassen. Auf diese Weise soll eine größere Bedarfsorientierung erreicht werden, damit nicht nur die technisch-naturwissenschaftlichen Angebote bei der Förderung von Einzelprojekten Berücksichtigung finden.

Korrespondenzadresse

Dr. Knut Blind, Dr. Kerstin Cuhls

Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI)

Breslauer Str. 48

76139 Karlsruhe

Tel: 0721-6809-212 bzw. -14

Fax: 0721-6809-260

kb@isi.fhg.de; cu@isi.fhg.de

Literatur

Bardecki, M. J., 1984: Participants' Response to the Delphi Method: An Attitudinal Perspective. *Technical Forecasting and Social Change* 24: 281-292.

Blind, K./Cuhls, K./Grupp, H., 2001: The Influence of Personal Attitudes on the Estimation of the Future Development of Science and Technology: A Factor Analysis Approach, erscheint in: *Technical Forecasting and Social Change* 2001.

Bundesministerium für Forschung und Technologie, BMFT (Hg.), 1993: *Deutscher Delphi-Bericht zur Entwicklung von Wissenschaft und Technik*, Bonn.

Cuhls, K., 2000: Opening up Foresight Processes. *Économies et Sociétés, Série Dynamique technologique et organisation* 5: 21-40.

Cuhls, K./Blind, K./Grupp, H., 2001: *Innovations for our Future*, Heidelberg: Physica Verlag.

Cuhls, K./Blind, K./Grupp, H. (Hg.), 1998: *Delphi '98 Umfrage. Zukunft nachgefragt. Studie zur globalen Entwicklung von Wissenschaft und Technik*, Karlsruhe.

Cuhls, K./Breiner, S./Grupp, H., 1995: *Delphi-Bericht 1995 zur Entwicklung von Wissenschaft und Technik - Mini-Delphi -*, Karlsruhe (auch herausgegeben als BMBF-Broschüre, Bonn 1996).

Dalkey, N.C., 1969a: Analyses From a Group Opinion Study, *Futures*, 2, Jg., 12: 541-551.

Dalkey, N.C., 1969b: *The Delphi Method: An Experimental Study of Group Opinion*, prepared for United States Air Force Project Rand, Santa Monica.

Dalkey, N.C./Brown, B./Cochran, S., 1969: *The Delphi Method, III: Use Of Self Ratings To Improve Group Estimates*, Santa Monica: Rand Corporation Paper RM-6115-PR.

Dalkey, N.C., 1968: *Predicting the Future*, Santa Monica.

Dalkey, N./Helmer, O. 1963: An Experimental Application Of The Delphi-Method To The Use Of Experts. *Journal of the Institute of Management Sciences, Management Science*, 9: 458-467.

Festinger, L., 1978: *Theorie der kognitiven Dissonanz*, Bern/Stuttgart/Wien: Huber.

Grupp, H. (Hg.), 1995: *Der Delphi-Report*, unter Mitarbeit von Breiner, S. und Cuhls, K. Stuttgart: dva-Verlag.

Häder, M., 2000: Subjektiv Sicher und Trotzdem Falsch? Methodische Ergebnisse einer Delphi-Studie zur Zukunft des Mobilfunks. *ZUMA-Nachrichten* 46: 89-116.

National Institute of Science and Technology Policy (NISTEP), 1997: Dai 6 kai Gijutsuyosoku Chôsa, NISTEP Report No. 52, Tokyo.

National Institute of Science and Technology Policy (NISTEP), 2001: Dai 7 kai Gijutsuyosoku Chôsa, NISTEP Report No. 71, Tokyo.

KONSEQUENZEN FÜR DIE METHODENAUSBILDUNG AUS DEM GUTACHTEN DER KOMMISSION ZUR VERBESSERUNG DER INFORMATIONELLEN INFRASTRUKTUR ZWISCHEN WISSENSCHAFT UND STATISTIK (KVI)¹

WALTER MÜLLER

Die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik geht auf ein Thesenpapier zurück, das im Sommer 1999 im Rahmen einer Konferenz im Statistischen Bundesamt verfasst wurde und in dem empfohlen wird, eine Kommission einzurichten, die „Lösungsvorschläge zu den inhaltlichen, organisatorischen und ressourcenbezogenen Fragen der Zusammenarbeit zwischen Wissenschaft und Statistik erarbeitet“ (Statistisches Bundesamt 1999: 237). Die Bundesministerin für Bildung und Forschung hat im Herbst 1999 die Kommission berufen. Diese hat im Frühjahr 2001 ihr Gutachten übergeben (Kommission 2001)². In diesem kurzen Beitrag will ich mit zwei Vorbemerkungen beginnen, nachher kurz auf die Arbeiten der Kommission und ihre Empfehlungen eingehen und dann mit einigen Gedanken zu möglichen Folgerungen für die Methodenausbildung schließen.

In autumn 1999 the German Federal Minister for Education and Research set up a “Commission to improve the informational infrastructure between official statistics and the scientific community”. The commission was asked to develop proposals to strengthen systematically the co-operation between these two bodies and to make recommendations on how to deal with the substantive and organisational issues involved

¹ Vortrag an der Tagung der Methodensektion der Deutschen Gesellschaft für Soziologie „Praxisrelevanz der Methodenausbildung“, Universität Bremen, 6./7. Juli 2001.

² Mitglieder der Kommission waren: Hans-Jürgen Krupp, Vorsitz in Kooperation mit Johann Hahlen (Präsident des Statistischen Bundesamtes); Eckhart Hohmann (Präsident Hessisches Statistisches Landesamt); Axel Reimann (Verband Deutscher Rentenversicherungsträger; Karl Ulrich Mayer und Walter Müller (für die Sozialwissenschaften); Gerhard Arminger (für die Statistikwissenschaft); Richard Hauser und Gert G. Wagner (für die Wirtschaftswissenschaften); Willem F.M. de Vries (UN Statistics Division, für den internationalen Bereich); das Sekretariat besorgte Gabriele Rolf-Engel. Kurzfassung des Gutachtens und der Empfehlungen unter: <http://www.bmbf.de/presse01/338.html>.

and the resources needed to achieve these goals. The Commission published its findings this spring in a report entitled "Ways to an improved informational infrastructure". The monograph is accompanied by a CD containing the full text of the Commission's recommendations, an English summary of the main points of the report, and expert appraisals commissioned in the course of the work. This article briefly describes the work, findings and recommendations of the Commission, then considers possible consequences for the teaching of empirical social research methods. A number of concrete suggestions are made on how to improve methods instruction in the empirical social sciences and for official statistics.

1. *Vorbemerkung: Die Nichtexistenz der amtlichen Statistik in der Methodenausbildung*

Ohne große Übertreibung kann man wohl feststellen: Eine mehr als über Marginalien hinausgehende Diskussion der Daten der amtlichen Statistik, ihrer Stärken und Schwächen, ihrer Einseitigkeiten und Lücken, aber auch ihrer Bedeutung für sozialwissenschaftliches Wissen und ihrer unverzichtbaren Leistungen für Politik, Staat, Wirtschaft, Gesellschaft und Öffentlichkeit findet gegenwärtig in der Methodenausbildung nicht statt. In den umfangreichen Stichwortverzeichnissen der meistgenutzten Methodenlehrbücher gibt es so gut wie keine Einträge zur amtlichen Statistik.

Bei Andreas Diekmann kann man immerhin lesen, dass sich die Scheinexaktheit der amtlichen Statistik schon im biblischen Text zur Volkszählung ankündigte und außerdem, dass die Volkszählung 1987 in Deutschland mehr als eine Milliarde DM „verschläng“ (Diekmann 1995: 327). Schnell/Hill/Esser geben in neueren Auflagen einige Hinweise auf den Stichprobenplan des Mikrozensus. Es ist ihnen aber auch eine längere Fußnote Wert, mit Noelle-Neumann mitzuteilen, dass die „Volkszählung durch David von Gott durch eine Pest, die 70 000 Tote forderte, bestraft wurde“ (S. 18).

In einem merkwürdigen Gegensatz dazu enthält Bernhard Schäfers Lehrbuch zur Sozialstruktur der Bundesrepublik – und in anderen entsprechenden Lehrbüchern verhält es sich ähnlich – praktisch nur Daten der amtlichen Statistik. Allerdings muss es den Studierenden nicht unbedingt auffallen, dass es sich um Daten der amtlichen Statistik handelt, denn sie werden oft über Sekundärliteratur referiert, z.B. als Informationen aus der Brockhaus-Enzyklopädie. Es ist wie mit dem Strom, der aus der Dose kommt, oder der Milch aus der Tüte; die Informationen sind einfach da; man muss nicht wissen, wo und wie sie entstehen.

Trotz der Soziale-Indikatoren-Bewegung wurde in den letzten Jahrzehnten in der Ausbildung die Vermittlung von Kenntnissen über die amtliche Statistik als Grundlage der Bevölkerungsstatistik und Demographie, allgemeiner gesellschaftlicher Kenngrößen und Maßzahlen des Wirtschaftsgeschehens wahrscheinlich reduziert. In meinem Studium war das noch Gegenstand von propädeutischen Veranstaltungen. Heute kann man auch in Mannheim ein Diplom in Sozialwissenschaften bekommen, ohne sich eingehender mit den Leistungen und Problemen der amtlichen Statistik befasst zu haben. Das gerade erschienene Lehrbuch von Rohwer und Pötter (2001) leitet vielleicht einen Wandel ein. Es charakterisiert wenigstens unterschiedliche Typen von Daten der amtlichen Statistik und verweist ausführlich auf Literatur, in denen diese näher diskutiert sind.

2. *Vorbemerkung: Bewegung in der amtlichen Statistik*

In mehrerer Hinsicht vollziehen sich gegenwärtig Veränderungen in der amtlichen Statistik und in den Bedingungen sozialwissenschaftlicher Datenproduktion, die zu einem neuen Verhältnis zwischen Statistik und Sozialwissenschaft führen können.

- Die amtlichen Datenbasen werden inhaltlich umfassender und methodisch anspruchsvoller. Daten werden in ihrem Inhalt reichhaltiger und erschließen zunehmend für die Sozialwissenschaften wichtige Bereiche (Beispiele: Zeitbudget, Umwelt, Gesundheit). Vermehrt werden Längsschnittdaten erhoben (Beispiele: European Community Household Panel; Labor Force Surveys mit Panelkomponenten; Sozialversicherungskonten; IAB-Betriebspanel). Vor allem für die an Bedeutung gewinnende gesellschaftsvergleichende Forschung ist die amtliche Statistik eine unverzichtbare Datenbasis, da bislang nur wenig andere sozialwissenschaftliche Daten kontinuierlich und systematisch vergleichbar für verschiedene Länder gesammelt werden (nicht einmal für die Länder der EU).
- Durch die Verknüpfung unterschiedlicher Daten (z.B. aus Registern und Umfragen) entstehen neue Datenbasen mit hohem Potential für die empirische Forschung.
- In einzelnen statistischen Ämtern ist hohe Methodenkompetenz mit Spitzenleistungen in der Forschung versammelt (z.B. in Frankreich, im Vereinten Königreich, in den USA). In Deutschland muss sich das Profil der Aufgaben und Arbeiten in den Statistischen Ämtern noch verstärkt in dieser Richtung ausbilden.
- Die Non-Response-Problematik und steigende Kosten der Datenproduktion erfordern zunehmend die optimale Nutzung aller vorhandenen Datenbasen. Wegen der bei verschiedenen Erhebungen bestehenden gesetzlichen Beteili-

gungspflicht bilden die amtlichen Daten einen zunehmend wichtigen Referenzrahmen für eigene Erhebungen der Sozialwissenschaften.

- Die Möglichkeiten der Nutzung von Mikrodaten der amtlichen Statistik für wissenschaftliche Sekundäranalysen haben sich in jüngster Zeit entschieden verbessert (Müller 1999a, 1999b). Zunehmend mehr amtliche Datensätze können in ähnlich flexibler Weise theoriegeleitet und mit adäquaten statistischen Verfahren analysiert werden wie die in den Sozialwissenschaften generierten Daten.

Diese und weitere Entwicklungen der letzten Jahre werden in der zukünftigen sozialwissenschaftlichen Forschung das Gewicht von Daten der amtlichen Statistik und anderer öffentlicher Datenproduzenten wahrscheinlich verstärken. In anderen Ländern sind entsprechende Entwicklungen teilweise schon wesentlich weiter vorangeschritten (vgl. Alba/Müller/Schimpl-Neimanns 1994).

I. Arbeitsschwerpunkte der Kommission

Für Deutschland kann ein Schub zu intensiverem Austausch und zu verbesserter Zusammenarbeit von den Arbeiten der KVI ausgehen. In dieser Kommission haben Vertreter der amtlichen Datenproduzenten und der Wissenschaft in einer Reihe von Sitzungen viele Probleme der Kooperation intensiv diskutiert und dann gemeinsam mehrere Empfehlungen formuliert. Das in Gang gesetzte Gespräch ist selbst ein wichtiges Ergebnis der Arbeiten der Kommission, die sich im übrigen auf folgende Arbeitsschwerpunkte konzentriert hat:

1. Verschiedene Datenproduzenten erstellen mehr oder weniger regelmäßig Datenbanken für ihre spezifischen Aufgaben, Bedürfnisse und Interessen. Die Kommission hat versucht, eine Übersicht darüber zu gewinnen und die bestehenden Lücken und Probleme in der empirischen Erfassung zentraler gesellschaftlicher und wirtschaftlicher Handlungsfelder zu identifizieren. Diese Analyse erfolgte im Wesentlichen durch Expertisen, die den Status quo in unterschiedlichen Informationsfeldern untersucht haben.³ Ein Fazit dieser Übersicht ist zweifelsohne: Insgesamt gibt es viel mehr Daten als man erwartet. Oft stehen aber die erhobenen Daten nur in geringer systematischer Beziehung zueinander. Teilweise werden vorhandene Daten kaum genutzt. Es gibt aber auch erhebliche Datenlücken in wichtigen Bereichen wie z.B. im Gesundheitswesen, im Bildungswesen oder der Erwerbstätigkeit. Die „Optimierung“ und bessere Koordination in der Dateninfrastruktur ist eine eigene große Zukunftsaufgabe, bei der Serviceeinrichtungen wichtige Beiträge leisten müssen und können.

³ Die Expertisen sind auf einer dem Gutachten beigelegten CD verfügbar.

2. Als Folge unterschiedlicher Bedingungskonstellationen haben sich in verschiedenen Ländern unterschiedliche Formen der Aufgabenteilung und Kooperation zwischen öffentlicher Statistik und sozial- und wirtschaftswissenschaftlicher Forschung herausgebildet. Die Kommission hat die im Ausland unter verschiedenen institutionellen Rahmenbedingungen der amtlichen Statistik und des Forschungssystems entstandenen innovativen Modelle der Kooperation analysiert. Die Ergebnisse ergaben nützliche Hinweise für die institutionelle Reformen zur Verbesserung der informationellen Infrastruktur in Deutschland.
3. Auf dieser Grundlage hat die Kommission eine Reihe von Empfehlungen zur Weiterentwicklung der informationellen Infrastruktur in Deutschland verabschiedet. Die Empfehlungen werden von den beteiligten Mitgliedern aus unterschiedlichen Bereichen der amtlichen Statistik und der Wissenschaft gemeinsam getragen und wurden auch vom auftraggebenden Ministerium für Bildung und Forschung positiv aufgenommen. Dies lässt hoffen, dass sie für die zukünftigen Rahmenbedingungen der Sozialforschung in Deutschland positive Auswirkungen haben können. Die Empfehlungen der Kommission können im Folgenden nur in Stichworten und unvollständig zusammengefasst werden.

II. Empfehlungen der Kommission⁴

Institutionelle Rahmenbedingungen der Koordination und Kooperation von Wissenschaft und Statistik bei der Datenproduktion

In Deutschland hat sich eine hochdifferenzierte Infrastruktur der Datenproduktion herausgebildet, zu der sowohl die amtliche Statistik als auch wissenschaftsgetragene und privatwirtschaftliche Einrichtungen beitragen. Im Hinblick darauf empfiehlt die Kommission, an der grundsätzlich bewährten Arbeitsteilung von amtlicher Statistik und wissenschaftsgetragener Datenproduktion festzuhalten, aber die *Koordination zwischen Wissenschaft und Statistik* und die *Mitwirkung der Wissenschaft bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen* durch *institutionelle Regelungen* zu verbessern. Diese institutionellen Regelungen sollen auf verschiedenen Wegen, kontinuierlich und auf Dauer zu einer Verbesserung der Kooperation und gegenseitigen Abstimmung beitragen. Besonders hervorzuheben sind:

⁴ Wegen des begrenzten Umfangs dieses Beitrages sind die Empfehlungen teilweise sehr verkürzt und anders gegliedert als im Kommissionsgutachten. Die Verkürzungen sind nicht mit der Kommission abgestimmt, sondern werden nur vom Verfasser verantwortet. Wörtliche Übernahmen aus den Empfehlungen sind nicht systematisch gekennzeichnet.

*Die Einrichtung eines Rates für Sozial- und Wirtschaftsdaten zur Weiterentwicklung und Bewertung der informationellen Infrastruktur.*⁵

Ein solcher Rat für Sozial- und Wirtschaftsdaten soll als bundesweites Steuerungs- und Initiativgremium eingerichtet werden, in dem die wichtigsten Datenproduzenten und Datennutzer sowie Bund und Länder als Mittelgeber vertreten sind. Er soll

- *die Dateninfrastruktur in Deutschland bewerten, weiterentwickeln sowie für Bund und Länder Empfehlungen zum Programm der wissenschaftstragenen Statistik und dessen Finanzierung erarbeiten;*
- *die Sozial- und Wirtschaftsberichterstattung sowie die gesellschaftliche Dauerbeobachtung fördern;*
- *die Einrichtung und Arbeit von Forschungsdaten- und Servicezentren beraten, empfehlen und evaluieren;*
- *Projektmittel vergeben oder zur Vergabe vorschlagen;*
- *in Zusammenarbeit mit den einschlägigen Fachverbänden und den übergreifenden Organisationen (DFG, MPG, WGL) eine systematische Mitwirkung bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen der amtlichen Statistik und bei den einschlägigen Anhörungen der Parlamente (einschließlich der Datenschutzgesetzgebung) organisieren.*

Dem Gesetzgeber wird darüber hinaus empfohlen:

- *die gesetzlichen Vorgaben in den einzelstatistischen Gesetzen auf den verfassungsrechtlich gebotenen Mindestumfang zu beschränken und Detailfestlegungen den statistischen Ämtern und ihren Beiräten zu überlassen;*
- *eine formelle Anhörungspflicht vor der Einführung, Abschaffung oder vor wesentlichen Änderungen amtlicher Statistiken vorzusehen. Die Zahl der Vertreter der Wissenschaft [im Statistischen Beirat] sollte erhöht und es sollte sichergestellt werden, dass Vertreter der empirischen Sozial- und Wirtschaftsforschung entsandt werden;*
- *für Erhebungen der Ressorts und von nicht-statistischen Behörden bzw. Körperschaften und Institutionen wie der Bundesbank, der Bundesanstalt für Arbeit oder aus der Sozialversicherung regelhafte wissenschaftliche Beratungsmöglichkeit zu schaffen.*

5 Inzwischen hat das Bundesministerium einen Gründungsausschuss für den Rat für Sozial- und Wirtschaftsdaten berufen. Seine Mitglieder sind: Richard Hauser, Johann Hahlen, Eckhart Hohmann, Wolfgang Jagodzinski, Gerhard Kleinhenz, Hans-Jürgen Krupp, Karl Ulrich Mayer (Vorsitz), Walter Müller und Gert G. Wagner.

Prioritäten bei der Datenproduktion

Von den Experten werden in den verschiedenen Bereichen zahlreiche Verbesserungen in den amtlichen und wissenschaftsgetragenen Datenerhebungsprogrammen vorgeschlagen. Diese können schon wegen der für die Zukunft gewünschten verbesserten Abstimmung nicht alle kurzfristig realisiert werden. Die Kommission hält aber die folgenden Vorhaben für dringlich:

Prioritäten bei der Fortführung und beim Ausbau wichtiger Statistiken:

- die Durchführung einer Volkszählung;
- die Weiterentwicklung des Mikrozensus (u.a. als unterjährige Erhebung, die eine Aufbereitung als rotierendes Panel erlaubt und als Access-Panel für die Ziehung freiwilliger Haushaltsstichproben nutzbar ist);
- Verbesserung der Einkommens- und Verbrauchsstichprobe (u.a. Verringerung der Zeitabstände, Einführung eines rotierenden Stichprobenkonzepts, Nachweis differenzierter Vermögensangaben);
- Schließung gravierender Lücken in der Statistik des Unternehmenssektors.

Prioritäten bei wissenschaftsgetragenen Datenerhebungen:

- dauerhafte Institutionalisierung des SOEP als breit angelegte, wissenschaftsgetragene und ausreichend große Längsschnitt-Stichprobe;
- Fortführung von ALLBUS, ISSP und Wohlfahrtssurveys;
- verstärkte Förderung von Kohortenstudien, wie z.B. eines „Entwicklungsbus“, der frühe Kindheit, Adoleszenz und das frühe Erwachsenenalter umfasst.

Förderung der Forschung zur Datenerhebung, -aufbereitung und -archivierung

Zur Verbesserung der Qualität der Daten und zur Erhöhung der Tragfähigkeit ihrer Analysen ist systematische methodische Forschung unverzichtbar. Im Vergleich mit dem Ausland besteht in Deutschland diesbezüglich ein erhebliches Defizit. Die in der Wissenschaft betriebene Forschung zu den Methoden der Datenerhebung und zur Weiterentwicklungen der Analyseverfahren werden zudem bislang in die deutsche amtliche Datenproduktion nicht systematisch einbezogen. Die Kommission empfiehlt deshalb,

- *den Stand der Ausbildung und der Forschung in den Methoden der empirischen Sozial- und Wirtschaftsforschung durch eine Kommission des Wissenschaftsrates prüfen zu lassen;*
- *die Einrichtung von Lehrstühlen oder Zentren mit besonderer Ausrichtung auf methodische Probleme der Umfrageforschung und der amtlichen Statistik;*

- *in der amtlichen Statistik die Wissenschaft stärker in die Weiterentwicklung des methodischen Instrumentariums einzubeziehen und die budgetären Voraussetzungen für eine nachhaltige Methodenforschung in der amtlichen Statistik zu schaffen;*
- *die Etablierung des Faches „Empirische Wirtschaftsforschung“ als Teildisziplin bei den Förderungs- und Begutachtungsverfahren der DFG (analog zur „Empirischen Sozialforschung“).*

Datenzugang

Im Vergleich zur Vergangenheit haben sich die Möglichkeiten der Nutzung von Daten der amtlichen Statistik für die Zwecke wissenschaftlicher Forschung in jüngster Zeit vor allem durch die Pilotprojekte zur Entwicklung und kostengünstigen Bereitstellung von Mikrodaten als Scientific Use Files verbessert. Die positiven Erfahrungen sollten auf möglichst viele weitere Datensätze übertragen und Lösungen für die Finanzierung der dafür erforderlichen Arbeiten gefunden werden. Je nach Art der Daten bieten sich dabei unterschiedliche Modalitäten der Bereitstellung, des Zugangs und der Finanzierung an, um zugleich den Interessen des Datenschutzes und einer möglichst guten Datennutzung in der Forschung gerecht zu werden. Die Kommission empfiehlt:

für aggregierte Daten der amtlichen Statistik

- *ein gemeinsames Datenbanksystem der amtlichen Statistik, auf das via Internet zugegriffen werden kann und das die Daten aller Bundesstatistiken in regionalisierter Form enthält und die Daten – zumindest für die Wissenschaft – weitgehend kostenfrei oder gegen geringe Schutzgebühr zur Verfügung stellt.*

für Mikrodaten der amtlichen Statistik

- *die Entwicklung von Scientific Use Microdata Files als wichtigstes Instrument des Mikrodatenzugangs weiter voranzutreiben; auch geeignete ältere Daten sowie Regional- oder Panelfiles (z.B. des Mikrozensus) in dieser Weise zugänglich zu machen; den Datenproduzenten sollen die Kosten für die Erstellung von Scientific Use Files pauschal über Mittel der Forschungsförderung erstattet werden; für die Forscher soll nur ein Entgelt in Höhe der marginalen Kosten der Abwicklung der Datenlieferung anfallen;*
- *die Entwicklung von Public Use Microdata Files, u.a. für die Lehre;*
- *die Einrichtung von Gastwissenschaftlermodellen in den Ämtern sowie von Forschungsdatenzentren mit der Möglichkeit der kontrollierten Ferndatenverarbeitung;*
- *Verbesserung der Nutzungsmöglichkeiten internationaler Mikrodaten durch entsprechende Serviceeinrichtungen.*

Die wichtigste Neuerung ist die Einrichtung von Forschungsdatenzentren für Daten, bei denen schon die faktische Anonymisierung die Aussagekraft zu sehr einschränkt, oder für Untersuchungen, die das Zusammenführen verschiedener Datensätze erfordern. In Forschungsdatenzentren werden solche Daten unter besonderen Schutzvorkehrungen und Kontrolle durch die Datenbesitzer für Zwecke der wissenschaftlichen Analyse bereitgestellt. Damit eine Aufsplitterung nach Datenbesitzern verhindert wird, sollen solche Zentren im Grundsatz Daten unterschiedlicher Datenbesitzer und -produzenten zugänglich machen, auch besonders zu schützende und nicht hinreichend anonymisierbare Daten aus der wissenschaftlichen Forschung. Forschungsdatenzentren sollen sowohl bei Datenbesitzern als auch bei einzelnen wissenschaftlichen Einrichtungen aufgebaut werden.

Datenschutz

Der Ausbau dieser verbesserten Infrastruktur erfordert auch die Weiterentwicklung der gesetzlichen Grundlagen des Datenschutzes und ihrer Anwendung. U.a. empfiehlt die Kommission in diesem Zusammenhang

- *Verfahren zur Klärung der Nutzungsberechtigung zu entwickeln, beispielsweise die Formulierung eines Verhaltenscodex (Code of Conduct) oder die Zertifizierung von Einrichtungen, wenn die Wissenschaftszugehörigkeit nicht eindeutig institutionell geklärt ist;*
- *die Möglichkeit exakter Datenverknüpfungen – ohne ausdrückliche Zustimmung aller einzelnen Befragten – für rein statistische Zwecke gesetzlich zu regeln;*
- *ein Forschungsdatengeheimnis einzuführen, das sowohl das Zeugnisverweigerungsrecht des Wissenschaftlers im Hinblick auf Forschungsdaten wie ein entsprechendes Beschlagnahmeverbot gesetzlich verankert.*

Serviceeinrichtungen

Zur Steigerung der Effizienz in der Nutzung von Mikrodaten in der Forschung sollten auch in Zukunft Grundlagenarbeiten für die gesamte Wissenschaftsgemeinschaft in zentralen Serviceeinrichtungen geleistet werden. Entsprechende Serviceleistungen sollten ausgebaut und neuen Bedarfen vor allem im Bereich der Mikrodatennutzung angepasst werden.

Diese und weitere Empfehlungen können – auch wenn nur ein Teil davon wirklich umgesetzt wird – die Arbeitsgrundlagen für die empirisch orientierten Sozialwissenschaften entscheidend verbessern. Auch in Deutschland wird sich die Nutzung amtlicher Mikrodaten in der Forschung in Zukunft deutlich verstärken. Welche Folgerungen ergeben sich daraus für die Methodenausbildung?

III. Folgerungen für die Methodenausbildung

Die gegenwärtige Ausbildung in den Methoden der empirischen Sozialforschung ist noch stark durch ein Verständnis geprägt, das die Studierenden sozialwissenschaftliche Methoden lehrt, die einen Kontrast zur amtlichen Statistik bilden. Die amtliche Statistik wird als weitgehend fixiert auf die Zählung rechtlicher und administrativer Tatbestände und ihre Darstellung in Aggregattabellen perzipiert. Dagegen hat die empirische Sozialforschung ein durch systematische Methodenforschung fundiertes Methodeninstrumentarium entwickelt, das gesellschaftliche Realitäten nicht nur in vielen von der amtlichen Statistik nicht berücksichtigten Bereichen abzubilden erlaubt, sondern dieses auch professioneller tut. Mit unterschiedlichen Techniken und Verfahren werden Daten erhoben und analysiert, die jeweils möglichst passgenau auf ein spezifisches Forschungsproblem bezogen und auf die Prüfung spezieller Hypothesen ausgerichtet sind. Durch die Sammlung und Analyse eigener Daten kann zielgerichteter auf inhaltliche Fragestellungen der Soziologie geantwortet werden. Diese Vorstellungen haben mit wenigen Ausnahmen in vielen Teilen der Forschung zu einer Interesselosigkeit an der amtlichen Statistik und zu ihrer weitgehenden Vernachlässigung in der sozialwissenschaftlichen Methodenlehre geführt. Wenn sich auch nach wie vor Spuren dieses Dualismus finden, so tun beide Seiten gut daran, ihn zu überwinden.

Auch in der wirtschaftswissenschaftlichen Ausbildung hat die amtliche Statistik in den letzten Jahrzehnten Boden verloren. Zwar hat sich an einzelnen Universitäten die Tradition von Lehrstühlen zur Sozial- und Wirtschaftsstatistik erhalten, weil im Rahmen der Statistikausbildung für Wirtschaftswissenschaftler traditionsgemäß ausgewählte Teilgebiete und methodische Grundlagen der amtlichen Statistik gelehrt wurden. Allerdings wurden diese Teile mehr und mehr durch den verstärkten Ausbau der mathematischen Statistik und der Ökonometrie abgelöst.⁶ Dennoch hat die frühere Tradition der Vermittlung wirtschafts- und sozialstatistischer Grundlagen in den wirtschaftswissenschaftlichen Studiengängen bemerkenswerte Konsequenzen. Sie kommen in einem interessanten Aufsatz im Allgemeinen Statistischen Archiv durch einen Mitarbeiter des Statistischen Bundesamtes mit aller Klarheit zur Sprache. Dieser Aufsatz von Klaus Schüler (1998) hat den Titel „Was wünscht sich die amtliche Statistik von einer Statistik-Ausbildung an wirtschaftswissenschaftlichen Fakultäten?“ Als Ausbildungsbedarf besonderer Dringlichkeit nennt Schüler eine

⁶ Zur Diskussion dieser Entwicklung in den wirtschaftswissenschaftlichen Studiengängen vgl. u.a. Krug (1996, 1998), Rendtel (1998), von der Lippe (1998, 1999); lesenswert und amüsant zur Statistik-Lehre Krämer (1995).

Reihe von Gebieten, die das ureigene Ausbildungsangebot der empirischen Sozialforschung betreffen, wie

- Stichprobentheorie und Praxis;
- Methodenkenntnisse zu Datenerhebung, -verarbeitung und -darstellung;
- Gestaltung von Erhebungsunterlagen, d.h. Fragebögen;
- Technologien bei der Datenerhebung und Aufbereitung;
- Fehlerquellen und Methoden zur Vermeidung/Behebung solcher Fehler;
- Probleme des Erhebungsmanagements;
- Aufbau und Nutzung von Registern.

Besonders bemerkenswert an diesem Aufsatz ist, dass diese Ausbildungswünsche an die Wirtschaftswissenschaften gerichtet werden. Es gäbe, so kann man lesen, im Höheren Dienst des Statistischen Bundesamtes auch Nicht-Wirtschaftswissenschaftler wie Juristen, Mediziner oder Geographen. Soziologen wurden vom Autor im Statistischen Bundesamt jedoch offensichtlich noch nicht gesichtet. Die genannten Ausbildungsgegenstände werden bei den Wirtschaftswissenschaftlern als zu wenig angeboten vermisst, aber im Angebot der empirischen Sozialforschung nicht wahrgenommen. Bezeichnenderweise habe ich keinen Aufsatz mit einem ähnlichen Wunschkatalog für die Methodenausbildung in den sozialwissenschaftlichen Fakultäten gefunden.

In Wirklichkeit beschäftigen die Statistischen Ämter natürlich Soziologen, aber nur wenige⁷. Offensichtlich ist es der Disziplin bislang nicht ausreichend gelungen, die Überzeugung zu vermitteln, dass das Ausbildungsprofil in den Methoden der empirischen Sozialforschung gute Voraussetzungen für wichtige Aufgabenbereiche in der amtlichen Statistik schafft. Dies ist zum Nachteil beider Seiten. Der amtlichen Statistik entgeht die Einbeziehung des kontinuierlichen methodischen Fortschrittes, der durch die weltweite Methodenforschung gemacht wird und (hoffentlich) auch in der Ausbildung ihren Niederschlag findet. Den Absolventen sozialwissenschaftlicher Studiengänge entgehen Arbeitsplätze, in denen die im Studium erworbene Kompetenz eingesetzt und weiterentwickelt werden kann. An den Aufgaben der amtlichen Statistik gemessen sind dort die im Studium erworbenen Fachkenntnisse insgesamt umfassender nutzbar als in vielen Instituten der Markt-, Meinungs- und Werbeforschung. Dies gilt nicht nur für das Methodenwissen, sondern auch für viele sub-

⁷ Ein sehr löbliches Beispiel bildet das Institut für Forschung und Entwicklung in der Bundesstatistik im Statistischen Bundesamt, in dem seit einigen Jahren mehrere Soziologen als wissenschaftliche Mitarbeiter tätig sind.

stanzwissenschaftliche Bereiche des Soziologiestudiums. Aber schon im Hinblick auf die engere Frage der Praxisrelevanz der Methodenausbildung lohnt es sich aus wenigstens zwei Gründen, Studierende so auszubilden, dass die Institutionen der amtlichen Datenproduktion in der Rekrutierung von Soziologen einen Gewinn sehen.

Erstens: Wenn in diesen Einrichtungen mehr methodisch gut ausgebildete Soziologen arbeiten, ist das langfristig vorteilhaft für uns alle. Wir bekommen Daten aus der amtlichen Statistik, die methodisch überzeugender und inhaltlich für die Sozialwissenschaften interessanter werden.

Zweitens: Insgesamt kann für Soziologen ein ganz erheblicher Arbeitsmarkt entstehen. Beschäftigungsfelder für Sozialwissenschaftler bestehen nicht nur in der Bundes- und Länderstatistik im engeren Sinn, sondern auch in der Gemeinde- und Städtestatistik, in der Arbeitsverwaltung, bei den Sozialversicherungsträgern, in den Ministerien mit ihrer Ressortforschung oder in den vielen nachgeordneten Behörden, die im Grenzgebiet zwischen Statistik und Sozialforschung arbeiten, wie das IAB, das BIBB, das Bundesgesundheitsamt, das Bundeskriminalamt, Umweltbehörden und andere mehr. Schließlich sind auch viele internationale Organisationen ein großer Arbeitsmarkt, in dem viel Personal mit Kenntnissen in der angewandten Sozial- und Wirtschaftsstatistik gesucht wird.⁸

Will man die Praxisrelevanz der Methodenausbildung verbessern, so gibt es mit einer stärkeren Ausrichtung am Aufgabenspektrum der amtlichen Statistik zweifelsohne eine Feld, in dem sich Investitionen lohnen.

Die KVI hat sich selbst nicht intensiv mit der Lehre und mit Ausbildungsfragen befasst. Sie hat jedoch einige Empfehlungen formuliert, die allgemein zur Verringerung offensichtlicher Defizite beitragen können. Dabei ging es vor allem den ökonomischen Fachkollegen darum, die Ausbildung in empirischen Methoden generell stärker als Ausbildungsgegenstand für ihre Disziplin zu verankern und zu forcieren. In Verbindung mit den stärkeren Ausbildungskomponenten in mathematischen Grundlagen und in der formalen Statistik kann sich daraus die Konkurrenz für Soziologen und andere Sozialwissenschaftler verstärken, und dieses nicht nur in der

⁸ Nach Schätzungen von Krug (1998: 366) werden europaweit mehr als 100.000 Personen direkt oder indirekt im Bereich der amtlichen Statistik beschäftigt. Grünewald (1999: 78) berichtet aus der Einstellungspraxis europäischer Ämter: „Bedingt durch weitgehend fehlende Kenntnisse insbesondere im Bereich der Wirtschafts-, Sozial- und Bevölkerungsstatistik haben Absolventen deutscher Studiengänge, die auch Statistikelemente enthalten, deutlich eingeschränkte Einstellungschancen auf europäischer bzw. supranationaler Ebene.“

amtlichen Statistik. Die von der Kommission formulierten Empfehlungen sind in der folgenden Übersicht zusammengestellt.

Übersicht zu den Ausbildungsempfehlungen der KVI

- *Im Grundstudium: Verstärkte Ausbildung in Statistik, Ökonometrie und angewandte Informatik unter Verwendung realer Datensätze;*
- *im Hauptstudium: Vertiefung der Ausbildungsmöglichkeiten in Statistik und Ökonometrie und in den Bereichen Datenerhebung, Aufbereitung der Daten, Datenschutz und Auswertung;*
- *Etablierung einer regulären Graduiertenausbildung nach amerikanischem Muster, wobei das Erlernen neuer empirischer Methoden und die Vertiefung in Statistik und Ökonometrie obligatorisch sein sollten;*
- *Präsenz der amtlichen Statistik im Lehrangebot verstärken, insb. durch vermehrte Vergabe von Lehraufträgen zur Verbesserung der Kenntnis in spezifischen Methoden und Datenbanken der amtlichen Statistik;*
- *Einrichtung von Kontaktseminaren und Weiterbildungskursen in Zusammenarbeit von Universitäten und statistischen Behörden bis hin zu einem kontinuierlichen Aus- und Weiterbildungsprogramm oder Ergänzungsstudium für junge Wissenschaftlerinnen und Wissenschaftler, aber auch für Nachwuchskräfte der amtlichen Statistik.*

Was ergibt sich daraus konkret für die Methodenausbildung in den soziologischen und sozialwissenschaftlichen Ausbildungsgängen?

Bei dem ohnehin großen Ausbildungspensum und dem bestehenden Trend zur Verschlankung der Ausbildungssyllabi können die Lehrpläne gewiss nicht fundamental im Hinblick auf eine starke Umgewichtung zu Gunsten von mehr Berücksichtigung amtlicher Datenproduktion umgeschrieben werden. Dennoch gibt es entschiedenen Ausbaubedarf, der aber nicht nur die Methodenausbildung betrifft und dem in Grund- und Hauptstudium unterschiedliches Gewicht zukommt.

Im *Grundstudium* geht es primär um die Vermittlung von Grundkenntnissen über die sozialwissenschaftliche Dateninfrastruktur und die Methoden, durch die die Daten entstehen. Amtliche Datenquellen sind und bleiben ein zentraler Teil der Informationsbasis, die Soziologie als Gesellschaftsanalyse nutzt. Deshalb sollten systematisch minimale Grundlagen auch über diesen Teil des Beobachtungsinstrumentariums vermittelt werden, allerdings eher nicht als spezieller Kursus, sondern integriert in die generelle Methodenlehre und die Sozialstrukturveranstaltungen.

Wichtige Bereiche sind:

- Arbeitsteilung und Abhängigkeiten zwischen sozialwissenschaftlicher und amtlicher Datenproduktion; methodische Gemeinsamkeiten sowie Stärken und Schwächen beider Bereiche;
- wichtigste Typen amtlicher Daten (Erhebungen, Buchführung, Verwaltungsregister), Veröffentlichungsquellen und Zugangs- und Nutzungsmöglichkeiten der Daten;
- Überblick über die wichtigsten für die Sozialwissenschaften relevanten Datenbasen, einschließlich der methodischen Besonderheiten ihrer Generierung;
- Verständnis der Definitionen, Operationalisierungen und Datengrundlagen zentraler Indikatoren (Beispiele: Erwerbstätigkeit, Arbeitslosenquote, Nettoreproduktionsrate).

Im *Hauptstudium* können Daten der amtlichen Statistik ein Vertiefungsgebiet in einem Methodenschwerpunkt darstellen, ähnlich wie beispielsweise qualitative Methoden, Evaluationsforschung oder ein Vertiefungsgebiet zu den Grundlagen des Befragtenverhaltens und speziellen Befragungstechniken. Es gibt einen ganzen Kanon von Themen, in denen eine sinnvolle Spezialisierung des methodischen Wissens denkbar und eine Professionalisierung der Soziologieausbildung wünschenswert ist. Alle im Grundstudium nur an der Oberfläche angesprochenen Themen können vertieft werden. Dabei muss es nicht unbedingt darum gehen, Spezialisten für Statistische Ämter auszubilden, obwohl im einen oder anderen Studiengang eine solche Spezialisierung durchaus sinnvoll sein könnte.⁹ Ein anderes – wahrscheinlich sinnvollerer, weil allgemeiner verwertbares – Modell würde darin bestehen, am Beispiel von Problemen, die sich u.a. besonders in der amtlichen Statistik zeigen, allgemeine Methodenprobleme zu behandeln. In einem solchen Modell würde zugleich die allgemeine Methodenkompetenz gestärkt und ein wichtiges Praxisfeld für Berufstätigkeiten in Einrichtungen der öffentlichen Datenproduktion erschlossen. Elemente davon können sinnvoll auch in Studiengänge eingebaut werden, die nicht speziell auf die amtliche Statistik ausgerichtet sind.

⁹ In Frankreich beispielsweise bestehen mit der ENSAE (Ecole Nationale de la Statistique et de l'Administration Economique) und mit der ENSAI (Ecole Nationale de la Statistique et de l'Analyse de l'Information) zwei Ausbildungseinrichtungen auf dem Niveau von Grandes Ecoles, die sich zentral auf die Ausbildung von Personal für die Statistischen Ämter konzentrieren. Außerdem verfügt das Französische Statistikamt INSEE mit dem LECEPE (Centre de Formation) über eine eigene große und professionelle Einrichtung zur Fortbildung seines

Beispiele für solche Veranstaltungen sind:

- Quellenkunde und Quellenkritik der zentralen Datenbasen der amtlichen Statistik und empirischen Sozialforschung;
- „Zur Metaphysik der Sozialstatistik“;
- Komplexe Datendesigns: der European Labour Force Survey als Beispiel;
- Sekundäranalyse amtlicher Daten und Untersuchungen;
- Organisations- und Qualitätsmanagement von Großerhebungen;
- Datenverknüpfung und Datenschutz;
- Theorie, Methoden, Daten und Modelle der gesellschaftlichen Dauerbeobachtung;
- Wie kann man aus Verwaltungsvorgängen interessante sozialwissenschaftliche Daten gewinnen und welche Tücken haben sie?

Alle diese und viele weitere Themen wären auch sinnvolle und mit großer Wahrscheinlichkeit nachgefragte Gegenstände für Angebote in *Aufbaustudiengängen, Weiterbildungskursen oder einer Sozialforschungsakademie*. Welches konkrete Bedarfe sind, sieht man am besten daran, was bereits am Ausbildungsmarkt angeboten und dort nachgefragt wird. Zur Illustration habe ich eine Liste mit ausgewählten Kursen zusammengestellt, die von dem in Luxemburg in enger Anbindung an EUROSTAT existierenden TES-Institut (Training for European Statisticians) für die Weiterbildung von Statistikern angeboten werden. Schon in dieser kleinen Auswahl zeigt sich, wie vielfältig die Zahl der Themen ist, zu denen die empirische Sozialforschung in der Ausbildung unmittelbar praxisbezogen beitragen kann.

Übersicht: Ausgewählte Kurse aus dem Angebot von TES (Training for European Statisticians)

Sampling Techniques and Practice
 Dealing with Non-Response
 Variance estimation and Discrete Data Analysis for Complex Surveys

Measurement in Surveys
 Nomenclatures, Classifications and their Harmonisation
 Concepts and Measurement of Inequality and Poverty
 Measurement of the Quality of Statistics

Theory and Application of Household Panel Surveys

Personals. Die herausragenden wissenschaftlichen Leistungen im INSEE sind damit ohne Zweifel verbunden.

Theory and Application of Enterprise Panel Surveys
Systems of Social Statistics
Living Conditions, Social Indicators & Social Reporting
Systems of Health Statistics
Surveys for the Service Sector
The European Labour Force Survey
Regional Accounts and Indicators
Demographic Data and Their Analysis
The European System for the Collection of Economic Information on the Environment

Regression Modelling for Complex Survey Data
Techniques of Electronic Data Dissemination
Statistical Presentation with Graphic Software
Geographical Information System

The Role of Statistics in a Democracy
Confidentiality and Protection of Privacy

Eine weitere wichtige praxisrelevante Komponente kann durch *Praktikumsplätze* in Statistischen Ämtern oder anderen Einrichtungen der amtlichen Datenproduktion erschlossen werden. In den entsprechenden Einrichtungen sind vielfältige Möglichkeiten dieser Art vorhanden. Studierende müssen nur darauf hingewiesen werden. Dies kann dann das Interesse bei den Studierenden stärken, Diplom- und andere *Abschlussarbeiten* aus dem Bereich der amtlichen Statistik zu schreiben. Solche Arbeiten könnten natürlich auch unabhängig davon vermehrt vergeben werden. Das Statistische Bundesamt schreibt kontinuierlich eine Liste mit Themen fort, zu denen es gerne Arbeiten sehen würde (die angebotenen Themen überfordern allerdings oft die Möglichkeiten einer Abschlussarbeit und sind oft eher für Betriebs- und Volkswirte geeignet als für Sozialwissenschaftler). Das Interesse des Statistischen Bundesamtes an entsprechenden Arbeiten wird auch durch einen Preis signalisiert, den es jährlich für herausragende Abschlussarbeiten und Dissertationen vergibt, die in einem sehr breiten Sinne Bezug zum Arbeitsgebiet der amtlichen Statistik haben.¹⁰

Insgesamt bin ich der Auffassung, dass auch hier die Wechselwirkung von Forschung und Lehre entscheidend ist. Das Lehrangebot wird sich vor allem dann stärker in Richtung amtlicher Daten entwickeln, wenn in der Wissenschaft mehr mit

¹⁰ http://www.statistik-bund.de/allg/d/veroe/fue_txt.htm

diesen Daten geforscht wird. Und es wird umso mehr mit solchen Daten geforscht werden, je interessanter sie für sozialwissenschaftliche Fragestellungen werden und je besser zugänglich sie sind. Damit diese positiven Feedbackprozesse in Gang kommen, erscheint es mir sehr wichtig, dass die Sozialwissenschaften positiv auf die Öffnungssignale aus der Statistik reagieren.

Korrespondenzadresse

*Prof. Dr. Walter Müller
Universität Mannheim
Lehrstuhl für Methoden der Empirischen Sozialforschung
und angewandte Soziologie
Seminargebäude A 5, 68131 Mannheim
Tel.: 0621-181-2047
wmueller@sovi.uni-mannheim.de*

Literaturhinweise

Alba, R./Müller, W./Schimpl-Neimanns, B., 1994: Secondary Analysis of Official Microdata. S. 57-78 in: Borg, I./Mohler, P. Ph. (eds.): Trends and Perspectives in Empirical Social Research. New York: de Gruyter.

Diekmann, A., 1995: Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt.

Grünewald, W., 1999: Über Kenntnisse in der angewandten Statistik im europäischen Kontext. S. 75-78 in: Statistisches Bundesamt (Hrsg.), Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven. Band 34 der Schriftenreihe Forum der Bundesstatistik. Stuttgart: Metzler-Poeschel.

Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik, 2001: Wege zu einer besseren informationellen Infrastruktur. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Freiburg: Nomos.

Krämer, W., 1995: Was ist faul an der Statistik-Grundausbildung an deutschsprachigen Wirtschaftsfakultäten? Allgemeines Statistisches Archiv, 79: 196-211.

Krug, W., 1998: Some proposals for enhancing cooperation between academic and official statisticians. Allgemeines Statistisches Archiv, 82: 242-250.

Müller, W., 1999: Amtliche Statistik und empirische Forschung: Wege in eine kooperative Zukunft. S. 17-35 in: Statistisches Bundesamt (Hrsg.), Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven. Band 34 der Schriftenreihe Forum der Bundesstatistik. Stuttgart: Metzler-Poeschel.

Müller, W., 1999: Der Mikro-Zensus als Datenquelle sozialwissenschaftlicher Forschung. S. 7-27 in: Lüttinger, P., (Hrsg.), Sozialstrukturanalysen mit dem Mikrozensus. ZUMA-Nachrichten Spezial, Band 6.

Rendtel, U., 1998: Praxisnähe im Grundstudium. Allgemeines Statistisches Archiv, Bd. 82: 396-406.

Rohwer, G./Pötter, U., 2001: Grundzüge der sozialwissenschaftlichen Statistik. Weinheim: Juventa.

Rohwer, G., 2000: Zur Metaphysik der Sozialstatistik. MS, <http://www.stat.ruhr-uni-bochum.de/scrip.html>

Schäfers, B., 1995: Gesellschaftlicher Wandel in Deutschland. Ein Studienbuch zur Sozialstruktur und Sozialgeschichte. 6. Auflage, Stuttgart: Enke.

Schnell, R./Hill, P. B./Esser E., 1999: Methoden der empirischen Sozialforschung. 6. Auflage, München u. Wien: Oldenbourg.

Schüler, K., 1998: Was wünscht sich die amtliche Statistik von der Statistik-Ausbildung an wirtschaftswissenschaftlichen Fakultäten? Allgemeines Statistisches Archiv, 82: 407-414.

Statistisches Bundesamt (Hrsg.) 1999: Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven. Stuttgart: Metzler-Poeschel.

von der Lippe, P., 1998: Mit Mikro-Daten einen Makro-Wirbel machen. Allgemeines Statistisches Archiv, 82: 380-386.

von der Lippe, P., 1999: Plädoyer für (wieder) mehr Wirtschaftsstatistik in der Statistikausbildung. S. 79-91 in: Statistisches Bundesamt (Hrsg.): Kooperation zwischen Wissenschaft und amtlicher Statistik – Praxis und Perspektiven. Band 34 der Schriftenreihe Forum der Bundesstatistik. Stuttgart: Metzler-Poeschel.

THE SWISS HOUSEHOLD PANEL 1999-2003: DATA FOR RESEARCH ON MICRO-SOCIAL CHANGE

*MONICA BUDOWSKI, ROBIN TILLMANN, ERWIN ZIMMERMANN, BORIS
WERNLI, ANNETTE SCHERPENZEEL & ALEXIS GABADINHO*

Die Umfrage Leben in der Schweiz des Schweizer Haushalt Panels (SHP) bietet eine einzigartige longitudinale Datenbasis an. 1999 wurden 7.799 Personen von 5.074 Haushalten über ihre Lebensbedingungen befragt. Sämtliche 14-jährigen und älteren Personen, die in diesen Haushalten wohnen, sollen fortan während zehn bis fünfzehn Jahren in jährlichem Abstand befragt werden. Die Erhebung wird mittels computerunterstützten Telefoninterviews (CATI: Computer Assisted Telephone Interviewing) durchgeführt. Mittlerweile konnten die ersten zwei Befragungswellen erfolgreich realisiert werden. Anders als bei den vorwiegend auf sozioökonomischen Bedingungen ausgerichteten Panels - wie dem SOEP in Deutschland und dem BHPS in England - deckt das SHP ein breites Spektrum von Themen und sozialwissenschaftlichen Forschungsansätzen ab. Das Faktenmaterial wird ergänzt durch „subjektive“ Beurteilungen. Die Trägerschaft des SHPs besteht aus dem Schwerpunktprogramm SPP „Zukunft Schweiz“, dem Bundesamt für Statistik und der Universität Neuchâtel.

The Living in Switzerland survey of the Swiss Household Panel (SHP) provides a unique longitudinal database in Switzerland. In 1999, 7,799 members of 5,074 households – from a stratified random sample of the permanent resident population of Switzerland – were interviewed about their living conditions. All household members aged 14 years and older are to be interviewed annually for 10 to 15 years. The SHP survey is conducted using computer-assisted telephone interviewing (CATI). To date, the first two waves have been carried out successfully. In comparison with panels such as the SOEP in Germany and the BHPS in Britain concentrating on socio-economic conditions, the SHP covers a broader range of topics and approaches in the social sciences. Subjective assessments complement the factual information. The SHP

is a joint project run by the Swiss Priority Programme (SPP) „Switzerland Towards the Future“, the Swiss Federal Statistical Office and the University of Neuchâtel.

1. Introduction

The purpose of this article is to publicise a unique Swiss database that was set up recently based on the Swiss Household Panel survey. Its main characteristics are as follows: 1) it is a comprehensive survey covering a broad range of social fields and a variety of topics; 2) all members of the households in the panel aged 14 years and older are interviewed; 3) the database offers opportunities for analysing gross social change and 4) it enables research on social trajectories of the individuals and groups comprising the resident population of Switzerland.

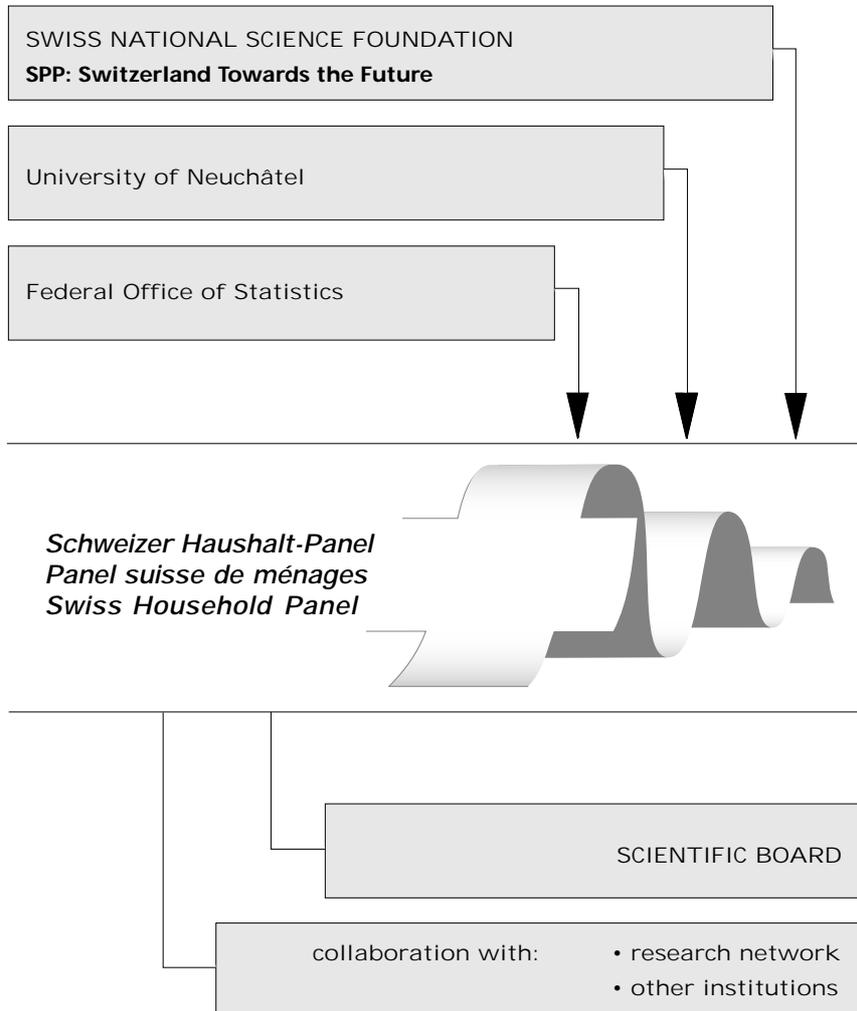
The SHP represents one element of the structural measures undertaken by the Swiss Priority Programme (SPP) „Switzerland Towards the Future“ for collecting basic data. Its purpose is to contribute to improving the situation for the social sciences in terms of quantitative data, to act as a reliable barometer of social change and provide an in-depth analysis of social dynamics in Switzerland. To date, it represents the largest financial investment ever made in a single social sciences project in Switzerland.

The Swiss Household Panel survey is a joint project run by the SPP „Switzerland Towards the Future“, the University of Neuchâtel and the Swiss Federal Statistical Office. These three partners are represented on the SHP steering committee and make strategic decisions as regards the content and running of the Living in Switzerland panel survey. The SHP team, based at the University of Neuchâtel, has operational responsibility for designing and managing the survey, and for appropriately distributing the data to interested researchers.

The data is collected by means of telephone interviews by M.I.S. Trend in Lausanne. Since spring 2001 the Scientific Council (comprising experts from Switzerland and abroad) have the dual role of guaranteeing the scientific quality of the collection and the analysis of the data, of providing financial support to promote projects which arise from the SHP's data analysis and of ensuring scientific publications. The SHP is also linked to a research network (Living in Switzerland), comprising all researchers using the survey in any way. The SHP team periodically organises seminars and working sessions that are to a large extent based on the contributions of the researchers themselves.



ORGANIZATION:



Before it was launched, the SHP was the subject of various preparatory steps which focused in particular on: 1) clarifying the objectives of the systematic observation of social change in Switzerland; 2) assessing European experiences, such as the Socio-economic Panel (SOEP) in Germany, the British Household Panel Study (BHPS) in Great Britain and the European Community Household Panel (ECHP) run by EUROSTAT; and 3) defining topics and indicators, and the data collection method (Farago 1996; Joye/Scherpenzeel 1997). The SHP structure reflects its main objective, which is to measure the multiple dimensions of social change as they occur in Switzerland over the next 10 to 15 years. Designed as a „data generator”, the SHP enables the changes to be analysed according to various disciplines and theoretical approaches. Household data is collected about the following domains: accommodation, standard of living, financial situation, the household, and the family; individual level data from the household and the family, health and victimisation, social origins, education, employment, income, participation, integration and networks, politics and values, and finally leisure and the media. The SHP has in addition been designed and operationalised so that international comparative analyses are possible. For this purpose, the SHP participates in the CHER programme (Consortium of Household Panels for European Socio-Economic Research).

2. Structure and Content of the SHP

2.1 Characteristics of a household panel

Household panels are considered to be tools for fine-tuning our conceptions and analyses of social dynamics (Berthoud/Gershuny 2000; Rose 1995). This general proposition can be made more specific. The dynamics (or social changes) at the macrosocial level do not directly belong to the field of observation covered by a panel survey. What panel surveys are intended to investigate, however, are the effects of changes at the macrosocial level on the living conditions of individuals and households, the manner in which these changes affect the individuals and households, and how they produce social change on a microsocial level. The main purpose of household panels is therefore to understand the processes, causes and effects of the social changes currently occurring. Of course, panel surveys are not the only tools used to measure social change. The repetition of a cross-sectional survey makes it possible to calculate, for example, *net transitions* between two dates (e.g. a drop in the proportion of the population receiving social benefits, or a rise in unemployment), but not *gross transitions* (e.g. the number of unemployed still without a job one year later). The data collected from household panels supplies unique information, allowing not only to estimate gross transitions but also providing an „understanding” of the transitions observed i.e. the circumstances (family events, a change in the activity status,

events related to the state of health, etc.) causing movements in and out of a given state (e.g. the fact that an individual or a household is living below the poverty line). In other words, by observing the same individuals over the course of time it is not only possible to study the change in numbers but also the flow of movement between the various states of being and to establish links of causality between different factors and events (temporal sequences).

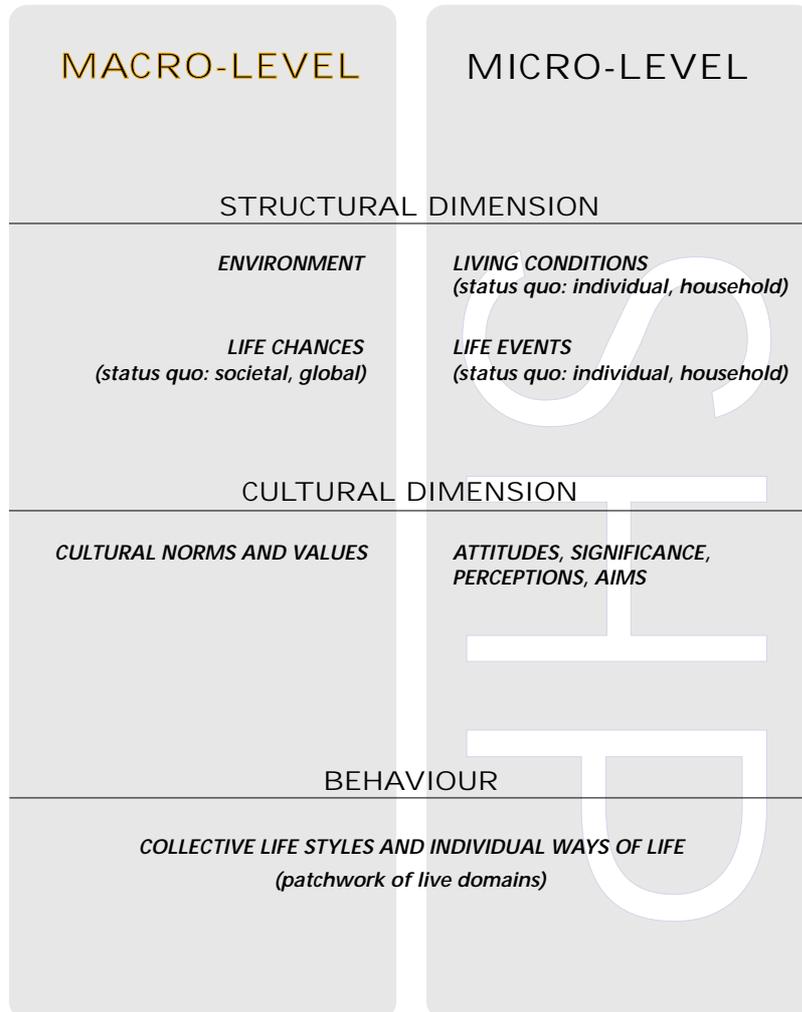
However, household panels, such as the SHP survey, are obviously not the only type of longitudinal surveys (Buck/Ermisch/Jenkins 1996). Longitudinal surveys include, notably, retrospective (biographical) surveys and other types of panels (cohort panels, rotating panels). In all cases, the main characteristic of the longitudinal survey is to provide repeated observations over a period of time for a set of variables for the group of individuals and households interviewed. Consequently these allow for richer dynamic analyses than is possible by repeating cross-sectional surveys.

2.2 Structure of the SHP

The construction of the SHP survey is drawn from international knowledge of the social sciences and the experiences of various panel surveys in Europe and North America (Budowski et al. 1998a; Joye/Scherpenzeel 1997). The structure of the Living in Switzerland survey must satisfy its primary objective, which is to measure the multiple dimensions of social change as they occur over the next 10 to 15 years in Switzerland. Given that the survey is designed to enable analyses in different disciplines using different conceptual approaches, the design of the SHP survey is based on the one hand on theoretical work related to the structure and development of contemporary societies (Beck 1986; Bourdieu 1983; Eisenstadt 1990; Haferkamp 1990; Konietzka 1995; Leisering/Walker 1998; Mayer 1991; Müller/Schmid 1995) and on the other hand on recent analyses of Swiss society and the way it functions (Höpflinger/Charles/Debrunner 1991; Leu/Burri/ Priester 1997; Levy/Joye/Guye/Kaufmann 1997). Contemporary societies are currently undergoing profound structural changes, which are the consequence of globalisation of the economy, uneven regional developments around the world, and world-wide demographic changes. Despite the new opportunities this evolution offers for households and their members, these changes are affecting current living conditions and lifestyles as well as aspirations regarding the future.



PANEL-ARCHITECTURE:



The way individuals and households adapt to the new realities is only partially the result of their own preferences and abilities; it is also to a major extent dependent on the social structures, standards and values present in the immediate social environment and society at large. Individual behaviour patterns therefore result from a dynamic relationship between the structural and the cultural dimensions. These influence the preferences and behavioural tendencies of individuals. Collective lifestyles and perceptions mediate the impact of these two dimensions. Collective behaviour patterns, in turn, may lead to transformations in the structural and cultural dimensions. In one way the SHP survey is designed to distinguish between structural, cultural and behavioural dimensions, in another, since it essentially collects data on individuals, it distinguishes between the macrosocial and the microsocial level. The contents of the survey, i.e. the actual questions used when collecting the data, reflect as a result, the three dimensions mentioned and the microsocial level. Macrosocial information (from other sources) can be added to the panel's database.

2.3 Content

The Living in Switzerland survey is a comprehensive survey. The questionnaires (household and individual) cover a broad range of social fields and topics. They are also designed to collect both „objective” data (resources, social position, participation, etc.) and „subjective” data (satisfaction, values, evaluation, etc.). The whole constitutes an operationalisation of the different elements of the microsocial level of the panel structure: living conditions, life events, attitudes and perceptions, and lifestyles/ways of life (Budowski et al. 1998b):

2.3.1 Areas covered by the survey

By definition, a household panel collects data at two levels: the household and the individual (for all individuals in the household aged 14 and older). In the case of the SHP-survey, the interview at *the household level* covers the following areas:

1. *composition of the household*: containing basic information about all the members of the household, such as the age, sex, relations between the members of the household, nationality, level of education and occupational status;
2. *accommodation*: containing „objective” elements, such as the type and size of the accommodation, home ownership or tenancy, the cost of and/or the subsidies received for housing, as well as „subjective” elements, such as satisfaction with the accommodation, evaluation of the state of the accommodation and assessment of perceived nuisances;

3. *standard of living*: referring to a list of goods owned by the household or activities that its members can carry out, together with the reason (financial or otherwise) why goods are not owned or activities not carried out;
4. *the household's financial situation*: containing „objective” information such as the existence of financial difficulties (and the household's reactions to different situations), indebtedness and the reasons for it, the total household income, the amount of tax paid, and the social and private transfers, as well as „subjective” elements, such as satisfaction, an estimate of the minimum income the household considers necessary or an evaluation of how the household's financial situation has evolved;
5. *the household and the family*: collecting information on any external help available to the household for housework or child-care, the sharing of tasks, and decision-making within the household.

Individual interviews cover the following topics:

1. *the household and the family*: comprising „objective” elements, such as the existence of children living outside the household, the sharing of housework and child-care, as well as „subjective” elements, such as satisfaction with private life and with the sharing of the housework;
2. *health and „victimisation”*: covering „objective” elements, such as general illness and health problems, visits to the doctor and hospitalisation, long-term handicaps, threats or attacks endured, together with „subjective” elements such as the self-perceived state of health, the estimated evolution of the state of health, or satisfaction with one's own health;
3. *social origins*: referring to information related to profession, professional position, educational level, and the nationality of both parents together with possible financial difficulties in the family of origin;
4. *education*: covering the various levels of achieved education, education currently being pursued, fluency in foreign languages, and participation in on-the-job training;
5. *employment*: considering four different aspects: firstly, the collection of information necessary to determine the status of the interviewee in the labour market, secondly, information covering the current main employment, thirdly, information on second jobs, and finally details about the last main job held; these modules also comprise „objective” elements, such as profession, status of the profession, the number of hours worked, work schedule, atypical work, as well as „subjective” elements such as satisfaction with various aspects of the job, the evaluation of promotion prospects or of personal qualifications;

6. *income*: including „objective” elements such as total personal income, total professional income, received social transfers, received private transfers, and other income, and „subjective” elements, such as satisfaction with the financial situation and an evaluation of changes concerning the personal financial situation;
7. *participation, integration, networks*: taking into account „objective” elements, such as frequency of social contacts, non-remunerated work outside home, participation in associations, membership of and participation in religious groups, and „subjective” elements such as the assessment of social capital by means of evaluation of potential practical help and emotional support (from various social networks);
8. *politics and values*: referring to „objective” elements such as political participation, membership, party identification, political positioning; and „subjective” elements such as satisfaction with the political system, the evaluation of issues or even political values; and finally
9. *leisure and media*: comprising „objective” elements, such as leisure activities and the use of the media as well as „subjective” elements, such as satisfaction with leisure and free time.

From the second wave on, the questionnaire also includes a „life events” module and an „occupational calendar” module (covering the 12 months prior to the interview).

2.3.2 Nomenclatures, standard variables, and comparability with other surveys

From the outset, the SHP was designed and built up in such a way as to make international comparative analyses possible. Furthermore, it was always situated in relation to the work of the Swiss Federal Statistical Office (SFSO). To achieve this, the survey uses a certain number of nomenclatures and classifications and also offers a series of standard variables making comparisons possible.

On the subject of the nomenclatures and classifications, for example, the professions of the interviewees (and their parents) are classified using a uniform nomenclature drawn up by the SFSO. This means it is possible to apply the International Standard Classification of Occupations (ISCO), which enables comparisons with the results of various surveys of the Swiss Federal Statistical Office and with various other surveys carried out in Europe. Furthermore, the interviewees give the name of the company in which they work (or used to work). The attribution of the company to one of the economic sectors of the EUROSTAT general nomenclature of economic activities is carried out by means of the Register of Enterprises drawn up by the SFSO.

The data from the Living in Switzerland survey also allows for the construction of most of the *recommended variables* on working life (status in the labour market, status of activity, level of occupation, atypical working conditions, duration of the activity, etc.), education, state of health as well as income. By means of these standardised objective variables the comparison with various surveys and analyses carried out either in Switzerland or abroad is made possible. In particular, the results of the Living in Switzerland survey can be compared or combined with results from different data sets held by the SFSO, such as the *General Census*, the *Swiss Labour Force survey*, the *Swiss Health survey* or the *Household Budget survey*, by means of variables comparable to the *standard variables* presented in Table 1. Finally, the data set contains or will contain in near future a certain number of indicators aimed at making their use easier: various household typologies (Budowski/Wanner 2001) various indices of social position (such as ISCO, CAMSIS, Wright), the yearly equivalised household income, the typology of communes, the economic sector of companies, etc.

Table 1: Standard variables in Swiss Federal Statistical Office household surveys and equivalents in the Living in Switzerland survey

SFSO standard variables	Level of data collection for the SHP survey		
	Household grid	Household questionnaire	Individual questionnaire
Household size	X		
Commune of residence	X	X	
Household income		X	
Telephone lines	X		
Home ownership, tenant or rent-free accommodation		X	
Number of rooms in accommodation		X	
Year of birth			X
Date of birth			X
Sex	X		X
Position in the household	X		
Marital status	X		X
Nationality	X		X
Type of residential permit	X		
Highest level of education completed	X		X
Current activity	X		X
Level of labour force participation	X		X
Professional situation			X
Unpaid housework		X	X

In addition to the nomenclatures, definitions and questions taken from the household surveys carried out by the official statistics, the Living in Switzerland survey has also picked up on a series of questions from surveys recently carried out in Switzerland, in particular the *Quality of Life and Poverty in Switzerland* and *The Swiss and their Society*,

(Leu/Burri/Priester 1997; Levy/Joye/Guye/ Kaufmann 1997) and abroad (notably the questionnaires of the European Community Household Panel).

3. Methodology of the Living in Switzerland survey

3.1 The reference population: the permanent resident population of Switzerland

The reference population for the Living in Switzerland survey is the population permanently resident in Switzerland. The sample must therefore be representative of the whole of Switzerland without regional imbalances. It includes households of various nationalities provided that their members live on Swiss territory throughout the year. Seasonal workers, cross-border workers, and foreign tourists are not part of the permanent resident population and are therefore not taken into account in the sample. Switzerland is divided into seven large statistical regions. The methodology section of the Swiss Federal Statistical Office drew a random sample in each of these on the basis of the SWISSCOM's electronic telephone directory, which covers over 95% of all private households. The households selected in this way are a representative sample of the various social groups in all regions of Switzerland. However, as the interviews are carried out in the three official national languages (German, French, and Italian) only, there might be a certain bias concerning how population groups who have recently migrated to Switzerland are represented.

3.2 The survey unit: private households

The concept of household refers not only to households comprising individuals or groups of individuals but also collective households (e.g. homes or prisons) and non-profit organisations (NPO) such as charitable organisations, political parties, trade unions, religious communities. Only private households are included in the Living in Switzerland survey, collective households and NPOs are excluded. The definition of a „private household” may vary. In panel surveys, such as the panels in the CHER project (Schmaus/Riebschläger 1997) or the ECHP (EUROSTAT 1996a) private households must fulfil a first criterion: the sharing of a dwelling unit. A second criterion consists in the existence of certain common living arrangements. However, the definition of „common living” is not always the same, which implies that certain choices need to be made.

Table 2: Resident population in Switzerland and the SHP sample characteristics

Code	Regions of the national territory (cantons)	Population in 1999 (in thousands, rounded off)	in %	Sample of Living in Switzerland 1999 (1st wave, absolute figures, eligible persons)	in %	Sample of Living in Switzerland 1999 (1st wave, absolute figures, interviewees)	in %
1	Lake Geneva region (VD, VS, GE)	1,295	18%	1,797	17%	1,366	17%
2	Mittelland (BE, FR, SO, NE, JU)	1,656	23%	2,622	25%	2,001	26%
3	North-west Switzerland (BS, BL, AG)	988	14%	1,491	14%	1,146	15%
4	Zurich	1,199	17%	1,680	16%	1,254	16%
5	Eastern Switzerland (GL, SH, AR, AI, SG, GR, TG)	1,042	15%	1,406	14%	1,012	13%
6	Central Switzerland (LU, UR, SZ, OW, NW, ZG)	677	9%	920	9%	693	9%
7	Ticino	308	4%	453	4%	327	4%
		7,164	100%	10,369	100%	7,799	100%

The Living in Switzerland survey has adopted a fairly broad definition of „household” in order not to underestimate new types of cohabitation (in particular collective non-institutionalised ways of living) as compared with the classic concept of the „household/family”. Therefore the definition of a household is based on five fundamental criteria, which must all be fulfilled: 1) sharing at least one common dwelling room; 2) sharing certain expenses; 3) taking at least one meal together per week; 4) stability (the household is considered a long-term arrangement); and 5) the individuals consider it to be their main dwelling (rather than a second home, work-related accommodation, etc.). If blood ties, marriage (in the broad sense, including the spouse’s family), adoption or stable partnerships are included in the traditional definition of household/family, the SHP definition also allows non-institutionalised forms of living to be included, provided they fulfil the above-mentioned criteria. In each household a „reference person” needs to be defined. In the Living in Switzerland survey the reference person is designated freely by the household itself. However, it is specified that this person should know the household very well. The reference person should also, as far as possible, be 1) an adult (age 18 or older), 2) a longitudinal respondent of the panel – an original sample member (OSM), and 3) the same person from one wave to the next. In the first wave, it is the reference

person who answers the questions of the grid and the household questionnaire (in addition to the individual questionnaire) as well as the proxy questionnaire.

3.3 Data collection tools

In a household panel, the information is collected at various levels (household/individual), so several questionnaires are used. There may also be many different questionnaires for any one level. For example, a complete questionnaire for adults, and a shorter questionnaire for children or adults unable to respond for themselves (for reasons of illness, disability, or absence). Most household panels use three types of questionnaires, i.e. the household grid, the household questionnaire and the individual questionnaire. This is also the case for the SHP survey.

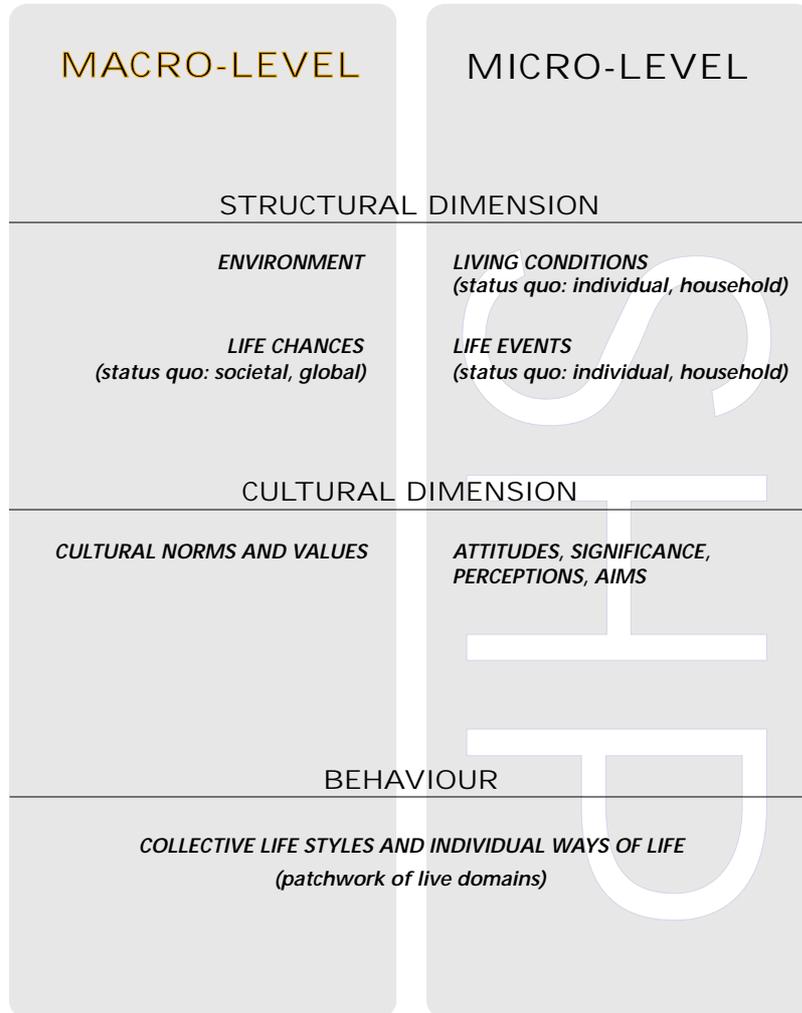
3.3.1 The household grid

The data in the household grid is not usually intended for analysis. However, as it is an operational control tool, it is absolutely crucial to keep it up to date to ensure the smooth operation of the data collection process. Every year the new grid, which is based on the previous one is updated taking into account address changes in between waves. This makes it possible to contact households and check whether their composition has changed, how many people they comprise, who must be interviewed and with which questionnaire (by personal interview or by proxy, etc.).¹ The grid's main functions are: 1) to check and trace changes within the sample (for both households and individuals) over time; 2) to collect information on the non-respondents; 3) to provide information necessary to link households and individuals over time; 4) and to collect basic data on the households and individuals.

¹ In the first wave, the grid was completed by the household's reference person. From the second wave onwards, this grid is amended in accordance with information from the first contact with any adult in the household.



PANEL-ARCHITECTURE:



In terms of the data collected, the household grid includes a list of all the household members and a series of basic details about them, such as: surname, Christian name, age, sex, date of birth, marital status, nationality, exact relationship among all the household members (not just in relation to the reference person), the highest level of education achieved, and finally the occupational status. The grid therefore provides information on the composition of the households and any changes concerning the structure of the households' „resources" (education, work, nationality, etc.). Given that much of this information is subject to data protection laws, researchers receive access to it in the form of aggregate variables (such as the relationship of the interviewee with the reference person or household typologies). It is furthermore important information for the calculation of the weights.

3.3.2 The household questionnaire

In the first wave, the household questionnaire was answered by the household's reference person. As far as possible, this questionnaire follows the household grid, to reduce the number of contacts and increase the response rate. From the second wave on, it can, if necessary, be completed by any adult member of the household. The household questionnaire contains questions about accommodation, living standards, the household's financial situation, the household's organisation, and the family. The Living in Switzerland survey has opted for a single version of the household questionnaire but adapts the wording and the questions depending on the type of household (i.e. households comprising a single adult or households with more than one adult).

3.3.3 The individual questionnaire

To collect data from individuals various questionnaire forms are used. Alongside the complete questionnaire, the CHER and ECHP panels have a variety of other instruments such as a specific questionnaire for children (to be filled in by an adult or the adolescent or pre-adolescent concerned). Sometimes a self-administered questionnaire is used when dealing with more delicate issues that are susceptible to the influence of another person in the usual face-to-face interview setting. Such questionnaires are also used for people who refuse to answer questions by telephone, or who are difficult to reach. Finally, there is a general proxy questionnaire for those who are absent for a long period or who are handicapped or too ill to respond.

The Living in Switzerland survey utilises only two questionnaires for the individual. The „standard" questionnaire contains all the questions in various sections: family, health, social origins, education, employment, income, networks, religion, leisure, media, politics and values. The „proxy" questionnaire comprises a limited number of

questions and is intended for ineligible candidates (children under 14 years), those unable to respond (handicapped, too old, etc.) or temporarily absent. Unlike the „standard“ questionnaire, it only contains factual questions about health, education, professional activity and income. The household's reference person answers the proxy questionnaire in the first wave. In the following waves the same adult in the household who responds to the household questionnaire can answer the proxy questionnaires. For ethical reasons, the proxy questionnaire is not carried out with those who are eligible but refuse to be interviewed individually. From the second wave on, the various questionnaires use a filter system to take account of the status of the interviewees: those who are already members of the panel and the newcomers (the cohabitants and children born to first wave members).

3.4 Decision to use CATI as the data collection mode

In the original project proposal, data was to be collected by the CAPI (Computer Assisted Personal Interviews) mode similar to the mode (face-to-face interviewing) used by most other European household panel studies (BHPS, SOEP, EURO-STAT). However, during the practical set-up of the project in spring and summer 1998, doubts first arose in terms of the sample size and the costs, followed by concerns about the quality of the interview situation and the monitoring of the interviewers if the CATI-mode were to be used. Therefore, the SHP decided to use CATI (Computer Assisted Telephone Interviewing). As no previous major household panel had started off with the CATI data collection mode, the SHP-team decided to carry out a method study jointly with the Swiss Federal Statistical Office (Scherpenzeel/Eichenberger 1998; Scherpenzeel/ Eichenberger 2000) in order to compare the „real life“ data collection procedures, that is: to study the *overall* effects of the two data collection modes CAPI and CATI². A methodological experiment was then designed in which the two data collection strategies were compared. The experimental design was a combination of the split-ballot design and the Multitrait-Multimethod approach. The Multitrait-Multimethod design is somewhat related to the test-retest approach. It also consists of repeating the same questions to the same people, but in addition, one aspect of the repeated questions is systematically varied. In this way, random variance can be distinguished from systematic method variance and estimates of both reliability as the complement of random error variance and validity the complement of systematic method variance are obtained. The criteria for comparison of the data collection strategies provided by the split-ballot design were the classical ones: costs, speed,

² The significance of a mere experimental study of the marginal effect of the medium of communication was considered to be very limited for survey practitioners.

response rates, answer distributions and summary statistics. The criteria for comparison provided by the Multitrait-Multimethod design were the reliability and validity of the data obtained (Andrews 1984; Saris/Andrews 1991; Scherpenzeel/Saris 1995; Scherpenzeel/Saris 1997). The experiment was carried out in the Swiss (German speaking) agglomeration of Bern between January 25 and April 13, 1999. The result of the experiment was: CATI saves time and costs in comparison with CAPI and should not be regarded a second choice among data collection techniques. The response rates and data obtained by CATI are at least as good as those by CAPI. The response rates on the household level are in fact somewhat better with CATI, although this difference is compensated for by a slightly lower individual response rate. Very similar in both modes of interviewing are most answer distributions, as are the overall validity and reliability of the data.

The only reason to select CAPI over CATI is when many sensitive questions or questions that require a good memory are asked. Indeed, interaction effects of data collection mode and these types of questions were found. However, strategies exist that help to overcome this shortcoming, such as stimulating confidentiality between interviewer and respondent, avoiding to formulate questions in a way that provokes socially desirable answers, introducing memory aids and stimulating respondents to take their time to think³.

3.5 Periodicity of the survey

Like other household panels, the SHP chose to interview the households and individuals annually. Due to memory lapses, answers to questions covering a longer period may be significantly biased. In a study focusing on such problems, Bailer (1989) showed that the longer a reference period gets, the greater the likelihood that the interviewees' memories are random. It further showed that certain phenomena studied are minimised, leading to significant underreporting. In addition, a prolonged reference period has been shown to pose major problems of telescoping and omissions (Cantor 1989). Owing to a long time span under review, interviewees tend to lose the dimension of time, with the risk of that they might either quote the same event twice in the same period or, forget it altogether. The periodicity of the panel surveys also has repercussions on the response rate, because of the greater probability of respondents moving and the difficulties encountered in finding them again for the following wave. Finally, problems can arise with regard to the motivation of respondents to continue in the panel when the periodicity is extended to 24 months between waves. The in-

3 For more details see: Scherpenzeel (2000c).

interviewees need to feel that they are part of an „experiment” which is of use to society and that the panel „lives” through them. For this feeling to be maintained they must stay in regular contact with the survey institute and the panel team. This is less feasible if the periodicity increases.

3.6 Follow-up rules

Follow-up rules make it possible to determine which people must be interviewed in the first wave and which people must be interviewed again in subsequent waves. The following rules were chosen for the Living in Switzerland survey. At the *household level*, in the first wave, all the households from the sample drawn are to be interviewed. A household becomes an *original sample member* (OSM) if the household questionnaire is completed together with at least one individual questionnaire. From the second wave on, the households are screened according to whether they satisfy the conditions for becoming original sample members. Among the original sample members the households which do not meet the following criteria are dropped: 1) those which do not respond to two subsequent waves; 2) those which have given a „refusal” considered as final; 3) those which move out of national territory; 4) those whose members are totally and „definitively” institutionalised. At the *individual level*, a distinction is made between individuals that are *original sample members* and those considered *cohabitants*. Those persons stemming from the first wave households are defined as original sample members or longitudinal members. Children born to original sample members after the first wave of the study also become such. The original sample members (OSM) defined in this way are followed constantly.⁴ Individuals joining or living in a household of at least one ordinary sample member in the second or any consecutive wave of the survey are defined as cohabitants. Cohabitants are, however, only followed while they continue to live with an ordinary sample member.

4. Development and phases of the project

4.1 Launching the SHP

In January 1998 the Swiss National Science Foundation took the final decision to fund the SHP. Thanks to a close collaboration with the SIDOS⁵, all researchers in Switzerland were contacted to participate and roughly 80 researchers stated their

⁴ In principal therefore until they die or are definitively institutionalised.

⁵ Swiss Information and Data Archive Service for the Social Sciences, Neuchâtel.

interest in certain topics and methodologies. These became the core of the Living in Switzerland research network. In tandem with establishing the scientific basis, defining the relevant indicators and drawing up the questionnaires, an important decision was made: that telephone interviewing would be used. The decision to carry out the successive surveys using CATI (computer-assisted telephone interviewing) rather than face-to-face using CAPI (computer-assisted personal interviewing) was based on dual considerations of quality and cost (Scherpenzeel 2000c; Scherpenzeel/Wernli/Eichenberger 1999). A call for tenders was launched in early September 1998 among institutes that are members of the nation-wide survey research association called SWISS-Interview. It contained the specific requirements for this type of survey such as sufficient CATI capacity to carry out 8,000 interviews each year in a pre-determined time frame, high quality of interviewers, capacity for management and updating of addresses of panel households, ability to maintain the loyalty of participants by applying an appropriate communication strategy, and a long-term commitment by the institute. After a detailed examination of the tender from interested parties, and discussions and negotiations with the various institutes, M.I.S. Trend in Lausanne was chosen in mid-December 1998.

In January 1999, the household and individual questionnaires were submitted to the Living in Switzerland research network for consultation then passed to the Steering Committee for approval. Over 50 researchers provided valuable comments and suggestions, which were systematically evaluated according to various criteria: compatibility with the panel structure, necessity of a longitudinal perspective, and implications for the questionnaire in the long term. In the light of these criteria, however, it was not possible to take all the remarks into account. A series of qualitative and quantitative pre-tests were conducted from February to the end of June 1999 in order to fine-tune the questionnaires and the survey procedure (the computerised management of appointments, the training of interviewers, the CATI interview procedure, and the data extraction) as well as to improve communication with the households and the interviewees (covering letter and explanatory brochure).

4.2 The first two waves of the survey

The first wave was carried out from September 1999 to February 2000. Out of the 14,174 addresses receiving an invitation to participate in the survey, telephone contact was possible with 12,084 households (85%). At the household level, the net response rate was 61%. To obtain reliable extrapolations for the whole resident population of Switzerland and a panel of about 4,000 households and 6,600 individuals for the second wave, it was necessary to request the participation of at least 5,000 households

in the first wave. This goal was achieved, as the first wave yielded valid data for 5,074 households and 7,799 individuals.⁶

Table 3: Breakdown of households and individuals into major regions, first wave

Regions of the national territory	Number of households	Number of people in the household	Number of eligible persons (14 years +)	Number of interviewees	Interviewees as % of eligible persons
Lake Geneva region	907	2,260	1,797	1,366	76
Mittelland	1,281	3,286	2,622	2,001	76
North-West Switzerland	726	1,826	1,491	1,146	77
Zurich	865	2,069	1,680	1,254	75
Western Switzerland	716	1,787	1,406	1,012	72
Central Switzerland	357	1,149	920	693	75
Ticino	222	554	453	327	72
Total	5,074	12,931	10,369	7,799	75

The households interviewed are single-person households (27%), couples without children (29%), couples with children (36%), single parents with children (6%) and 2% of other households. The household interviews lasted an average of 12 minutes and individual interviews an average of 55 minutes. In total, slightly fewer than 8,500 hours of telephone interviews were carried out. The first set of provisional data was provided in August 2000. The finalised set, with transversal weighting and constructed variables, was available in spring 2001.⁷

The data from the first wave of the SHP is very similar to that obtained in other household surveys carried out periodically by the SFSSO. For example, in Table 4 the distribution of the population according to status within the labour market is compared.

⁶ The number of panel households in the first wave is similar to those in the main comparable panels: 5,000 households in the PSID (Panel Study of Income Dynamics), 5,500 households in the BHPS (British Household Panel Survey) and 5,900 households in the GSOEP (German Socio-Economic Panel).

⁷ The Living in Switzerland survey was weighted by scientific specialists at the SFSSO.

Table 4: Status in the labour market according to sex (percentages, weighted results)

Status	SLFS 1999			ERC98 Living conditions			Living in Switzerland 1999		
	men	women	total	men	women	total	men	women	total
Working employed	74.4	55.5	64.7	76.4	56.7	66.2	74.4	55.1	64.4
Unemployed	3.7	2.6	3.1	1.9	2.3	2.1	1.4	2.3	1.9
Non-working	21.9	41.8	32.2	21.7	40.9	31.6	24.2	42.6	33.7

The second wave started in September 2000 and ended in February 2001. Valid data was collected for approximately 4,300 households and 7,000 individuals. The longitudinal sample (individuals interviewed in the first and second waves) comprises approximately 6,000 individuals. The data is likely to be made available to researchers in the autumn 2001.

Future plans envision a survey with a self-administered, written „biographical“ questionnaire to be completed by the panel interviewees after the third wave, to re-constitute their life trajectory since childhood, fundamentally of the living arrangements/family and professional life.⁸

5. Diffusion and exploitation of the data

5.1 Diffusion of data

By and large, there are no restrictions on the use of SHP data. However, interested parties must sign a contract with the SHP undertaking to 1) use the data only for predefined research purposes; 2) inform the SHP of their experiences and their publications; 3) refrain from transmitting the data to third parties; and 4) play an active role in the Living in Switzerland research network. The topics being analysed and the identity of members of the network are listed on the SHP website <http://www.unine.ch/psm>. On signature of the contract, the data is provided in a labelled

⁸ During May and June 2001 a pretest was conducted in a random subsample of 600 households in order to evaluate (i) response rates to the self-administered questionnaire with and without incentives and (ii) the repercussions on the participation rate in the annual panel interview (the third wave). Depending on the results, the survey will be carried out after the third wave or not. This decision will be taken by the Steering Committee in March 2002.

SPSS format and a SAS format together with the codebook⁹ and other information on a CD-ROM for PC or Mac.

5.2 The analysts

The considerable effort and financial resources invested in the SHP can only be justified if the data is used by a large number of analysts. It is mainly researchers in the field of the social sciences interested in the study of social change and relations of causality who need panel data. Nevertheless, this type of scientific perspective is new in Switzerland and there are few researchers actually trained in the methods and techniques of longitudinal analysis. To enhance such skills, the SPP „Switzerland Towards the Future“ launched a course on the analysis of longitudinal data at its summer school in 1999. A second group that can benefit from panel data is the staff of the Swiss Federal Statistical Office since analysis of SHP data complements the official statistics and the major surveys conducted by the SFSSO, mainly by supplying an estimation of the magnitude of gross transitions.

5.3 Types of publications

The diffusion of results through appropriate channels is essential to ensure their scientific and political impact, which is the true measure of the „return on investment“. Although the authors of the analyses may well be in the best position to choose an appropriate publication depending on the type of results and their own personal interests, the SHP considers that socially and politically pertinent results must not remain confined to scientific publications (journals, books, etc.) or specialised reports (publications by the Swiss Federal Statistical Office, Swiss Social Report, etc.) that are intended for a limited audience. Instead they should be communicated, via print or the broadcasting media, to target audiences defined by their profession or their socio-political role as well as to the general public. In other words, they should be used to stimulate public debate and serve as a basis for political decision-making.

6. Conclusion

The Swiss Household Panel was launched in 1998 and the first two waves have been carried out successfully. The first wave (1999) yielded valid data for 5,074 households and 7,799 individuals, with a net participation rate among the households of 61%. In

⁹ The documentation system and the code-book are a product of the close collaboration between the SHP and SIDOS, the Swiss Information and Data Archive Service for the Social Sciences.

the second wave (2000), approximately 4,300 households and 7,000 individuals were interviewed. The longitudinal sample comprises approximately 6,000 individuals.

Interdisciplinary collaboration is now emerging on topics such as the composition of households and the new forms of shared lifestyles, the phenomena of social insecurity and poverty, connections between living conditions and health, the influences of the context on political values and behaviour patterns, the multiple approaches and dimensions of social position, intergenerational aspects of the acquisition of Swiss nationality, etc.

Since the SHP data is representative of the resident population in Switzerland, it is a necessary complement to official statistics and will be used in the various publications of the Swiss Federal Statistical Office.

Since the beginning of 2001, the SHP's Scientific Board guarantees the quality of the work for collection and analysis of the data. The financial resources recently available to the board make it possible to support the scientific analyses and promote publications in which these appear.

The SHP has a stable institutional and financial basis, which guarantees the completion of five yearly waves of data collection. The management of SPP „Switzerland Towards the Future“ is currently exploring the possible ways of guaranteeing the future of the SHP beyond 2003.

Contact

Dr. Monica Budowski

Schweizer Haushalt-Panel, Panel suisse de ménages, Swiss Household Panel

Case Postale 1820

2002 Neuchâtel - Schweiz

E-mail: swiss.panel@unine.ch

monica.budowski@unine.ch

<http://www.unine.ch/psm>

References

- Andrews, F.M., 1984: Construct validity and error components of survey measures: a structural modelling approach. *Public opinion quarterly* 48:409-422.
- Bailar, B., 1989: Information needs, surveys, and measurement errors. Pp. 1-24 in: Kasprzyk, D./G. Duncan/G. Kalton/M.P. Singh (eds.), *Panel Surveys*. New York: John Wiley & Sons.
- Beck, U., 1986: *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp.
- Berthoud, R./J. Gershuny, 2000: *Seven Years in the Lives of British Families. Evidence on the dynamics of social change from the British Household Panel Survey*. Bristol: The Policy Press.
- Bourdieu, P., 1983: Oekonomisches Kapital, kulturelles Kapital, soziales Kapital. Pp. 183-198: in Kreckel, R. (ed.), *Soziale Ungleichheiten. Sonderheft 2 der Zeitschrift Soziale Welt*. Göttingen: Otto Schwarz.
- Buck, N./J. Ermisch/S.P. Jenkins, 1996: Choosing a longitudinal survey design: the issues. *Occasional Papers of the ESRC Research Centre on Micro-social Change. Occasional Paper 96-1*. Colchester: University of Essex.
- Budowski, M./M. Niklowitz/A. Scherpenzeel/R. Tillmann/B. Wernli/E. Zimmermann, 1998a: *Aims and Architecture of the Swiss Household Panel. Working Paper 1/98*. Neuchâtel: Swiss Household Panel (SHP).
- Budowski, M./M. Niklowitz/A. Scherpenzeel/R. Tillmann/B. Wernli/E. Zimmermann, 1998b: *Description of Life Domains and Indicators of the Swiss Household Panel. Working Paper 2/98*. Neuchâtel: Swiss Household Panel (SHP).
- Budowski, M./P. Wanner, 2001: *The Swiss Household Panel (SHP): A Rich Data Set for Research on the Household and the Family*. Manuscript submitted to the *Swiss Sociological Review*.
- Cantor, D., 1989: Substantive implications of longitudinal design features: The national crime survey as a case study. Pp. 25-51 in: Kasprzyk, D./G. Duncan/G. Kalton/M.P. Singh (eds.), *Panel surveys*. New-York.
- Eisenstadt, S., 1990: Kultur und Sozialstruktur in der neueren soziologischen Analyse. Pp. 7-20 in: Haferkamp, H. (ed.), *Sozialstruktur und Kultur*. Frankfurt am Main.

- EUROSTAT, 1996a: The European Community Household Panel: Survey methodology and implementation. Volume 1. Luxembourg: Eurostat.
- Farago, P., 1996: Gesellschaftliche Dauerbeobachtung im SPP "Zukunft Schweiz/Demain la Suisse". Working Paper 1/96. Bern: Swiss Priority Programme "Switzerland Towards the Future", Swiss National Science Foundation.
- Haferkamp, H. (ed.), 1990: Sozialstruktur und Kultur. Frankfurt am Main: Suhrkamp.
- Höpflinger, F./M. Charles/A. Debrunner, 1991: Familienleben und Berufsarbeit. Zum Wechselverhältnis zweier Lebensbereiche. Zürich: Seismo.
- Joye, D./A. Scherpenzeel, 1997: Observation à long terme : Projet de panel. Working Paper 2/97. Berne: Programme Prioritaire "Demain la Suisse". Fonds national suisse de la recherche scientifique.
- Konietzka, D., 1995: Lebensstile im sozialstrukturellen Kontext: ein theoretischer und empirischer Beitrag zu Analyse soziokultureller Ungleichheiten. Opladen: Westdeutscher Verlag.
- Leisering, L./R. Walker, 1998: The Dynamics of Modern Society. Bristol: The Policy Press.
- Leu, R.E./S. Burri/T. Priester, 1997: Lebensqualität und Armut in der Schweiz: Haupt.
- Levy, R./D. Joye/O. Guye/V. Kaufmann, 1997: Tous égaux? De la stratification aux représentations. Zürich: Seismo.
- Mayer, K.U., 1991: Soziale Ungleichheit und die Differenzierung von Lebensverläufen. Pp. 667-687: in Zapf, W. (ed.), Die Modernisierung moderner Gesellschaften. Verhandlung des 25. Deutschen Soziologentages 1990. Frankfurt am Main.
- Müller, H.-P./M. Schmid (eds.), 1995: Sozialer Wandel. Modellbildung und theoretische Ansätze. Frankfurt am Main: Suhrkamp.
- Rose, D., 1995: Household Panel Studies: an Overview. *Innovation* 8 (1):7-24.
- Saris, W.E./F.M. Andrews, 1991: Evaluation of measurement instruments using a structural modelling approach. Pp. 575-597 in: Biemer, P.P./R.M. Groves/L.E. Lyberg/N.A. Mathiowetz/S. Sudman (eds.), *Measurement errors in surveys*. New York: Wiley and Sons.
- Scherpenzeel, A., 2000c: Mode effects in panel surveys: A split-ballot comparison of CAPI and CATI. Submitted to *Sociological Methods and Research*.
- Scherpenzeel, A./P. Eichenberger, 1998: An Experimental Study of two Data Collection Strategies: Face-to-face Interviews (CAPI) vs. Telephone Interviews (CATI). Demande no. 5004-056052 de subside Div.IV, PNR/PP. Demande de subside Div.IV, PNR/PP 5004-056052. Neuchâtel: Swiss Household Panel (SHP).
- Scherpenzeel, A./P. Eichenberger, 2000: An Experimental Study of Two Data Collection Strategies: Face-to-face Interviews (CAPI) versus Telephone Interviews (CATI). Division IV, Swiss Priority Programme, Final Report of the partial project, Funding Period 1996-1999. Neuchâtel: Swiss Household Panel (SHP).
- Scherpenzeel, A./W.E. Saris, 1995: Effects of data collection technique on the quality of survey data: an evaluation of interviewer- and self-administered computer assisted data collection

techniques. Pp. 91-116 in: Scherpenzeel, A. (ed.), *A Question of Quality. Evaluating survey questions by multitrait-multimethod studies*. Leidschendam: Royal PTT Nederland NV, KPN Research.

Scherpenzeel, A./B. Wernli/P. Eichenberger, 1999: *An experimental study of data collection strategies: report of the first stage*. Working Paper 7/99. Neuchâtel: Swiss Household Panel.

Scherpenzeel, A.C./W.E. Saris, 1997: *The validity and reliability of survey questions: a meta-analysis of Multitrait-Multimethod studies*. *Sociological Methods and Research* 25: 341-383.

Schmaus, G./M. Riebschläger, 1997: *PACO User Guide*. Luxembourg: CEPS/ INSTEAD.

DIE GESIS IM URTEIL DER PROFESSION: ERGEBNISSE EINER BEFRAGUNG VON SOZIOLOGIEPROFESSORINNEN UND –PROFESSOREN IN DEUTSCHLAND, ÖSTERREICH UND DER SCHWEIZ

GISBERT BINDER, MARKUS KLEIN, ROLF PORST, MATTHIAS STAHL¹

Die Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS) ist eine Serviceeinrichtung für die Sozialwissenschaften. Eine solche Einrichtung muss sich daran messen lassen, inwieweit ihre Produkte und Dienstleistungen den Bedürfnissen ihrer Zielgruppe gerecht werden. Im Rahmen der Qualitätssicherung der GESIS wurde daher eine Befragung deutschsprachiger Soziologieprofessorinnen und -professoren in Deutschland, Österreich und der Schweiz durchgeführt, die sogenannte *GESIS-Potentialanalyse* (vgl. ausführlich Binder/Klein/Porst/Stahl 2001). In ihrem Rahmen wurde die Nutzung sowie die Bewertung der Produkte und Dienstleistungen der GESIS erfasst. Die Gruppe der Soziologieprofessorinnen und -professoren wurde ausgewählt, da sie zur Kerngruppe der (potentiellen) Nutzer der GESIS gehören und außerdem über eine besondere Expertise zur Beurteilung der Produkte und Dienstleistungen der GESIS verfügen sollten.

Die Durchführung der Befragung

Die GESIS-Potentialstudie wurde als Gemeinschaftsprojekt der drei GESIS-Institute Informationszentrum Sozialwissenschaften (IZ), Zentralarchiv für Empirische Sozialforschung (ZA) sowie Zentrum für Umfragen, Methoden und Analysen (ZUMA) durchgeführt. Die Feldarbeit lag bei ZUMA. Angestrebt war eine Befragung

¹ Gisbert Binder (IZ), Markus Klein (ZA), Rolf Porst (ZUMA), Matthias Stahl (IZ). Der Beitrag erscheint zugleich in der Ausgabe 49 der ZA-Information.

aller Soziologieprofessorinnen und -professoren in Deutschland, Österreich und der Schweiz. Insgesamt konnten vom IZ in den Vorlesungsverzeichnissen (WS 1999/2000) der wissenschaftlichen und sonstigen Hochschulen der drei genannten Ländern 636 Soziologieprofessorinnen und -professoren identifiziert werden. Von diesen wurden in der Zeit vom 9. November bis 14. Dezember 2000 insgesamt 431 telefonisch befragt. Die Ausschöpfungsquote lag damit bei beachtlichen 77,2 Prozent (vgl. zur Feldarbeit ausführlich Schneid/ Stiegler 2000, 2001).

Der Bekanntheitsgrad der GESIS und ihrer drei Mitgliedsinstitute

Das erste wichtige Kriterium, an dem sich die GESIS messen lassen muss, ist, ob sie von der Profession überhaupt wahrgenommen wird. Im Rahmen der Befragung wurde daher der Bekanntheitsgrad der GESIS insgesamt sowie seiner drei Mitgliedsinstitute ermittelt. Es zeigt sich, dass die GESIS 86 Prozent der befragten Soziologieprofessorinnen und -professoren bekannt ist. Der Bekanntheitsgrad von IZ und ZA liegt bei jeweils 87 Prozent, während das ZUMA 93 Prozent der Befragten kennen.

Die Nutzung und Bewertung des IZ

235 der 431 befragten Soziologieprofessorinnen und -professoren haben das IZ bereits für ihre wissenschaftlichen Arbeiten persönlich genutzt oder haben ihren Mitarbeitern den Auftrag gegeben, die Angebote des IZ für sie zu nutzen. Damit liegt der Anteil der Nutzer bei 55 Prozent aller Befragten. Betrachtet man ausschließlich die Befragten, die das IZ kennen, liegt der Anteil der Nutzer bei 63 Prozent.

Von den 181 persönlichen Nutzern des IZ kennen praktisch alle die Literaturnachweisdatenbank SOLIS. Drei Viertel haben bereits damit gearbeitet. Die Forschungsprojektdatenbank FORIS ist ebenfalls der überwiegenden Mehrheit der persönlichen Nutzer bekannt. SOLIS wird von knapp der Hälfte der Nutzer für ihre wissenschaftliche Arbeit als sehr wichtig oder wichtig eingeschätzt. FORIS wird von ca. 40 Prozent der Nutzer als sehr wichtig oder wichtig angesehen. Nur 5 der 42 Nutzer, die SOLIS für sich persönlich als eher unwichtig oder ganz unwichtig einschätzen, halten diese Datenbank auch für die Sozialwissenschaft insgesamt für unwichtig. Die entsprechenden Werte für FORIS sind sehr ähnlich.

Bei den einzelnen Produkten und Dienstleistungen ist die Nutzung des gedruckten Sozialwissenschaftlichen Fachinformationsdienstes soFid am weitesten verbreitet: Die persönlichen Nutzer des IZ haben zu ca. 40 Prozent schon mit dem soFid gearbeitet.

Die Themendokumentationen werden von 34 Prozent der Nutzer in Anspruch genommen.

Jeder dritte Nutzer hat bereits beim IZ eine oder mehrere Recherchen in Auftrag gegeben. Die Möglichkeiten, in den Datenbanken SOLIS und FORIS direkt online bzw. über die CD-ROM WISO III persönlich zu recherchieren, wurden von den Nutzern bislang seltener in Anspruch genommen: 33 Nutzer haben bereits selbst mit der WISO III gearbeitet. 43 Befragte geben an, dass ihre Mitarbeiter die CD-ROM WISO III in ihrem Auftrag benutzen. Diejenigen, die die CD-ROM WISO III persönlich kennen, beurteilen sie überwiegend mit „sehr gut“ oder „gut“.

Die Nutzer kommen nicht nur über Produkte und Dienstleistungen mit dem IZ in Kontakt, sondern in bedeutsamer Weise auch über die jährlich durchgeführte Forschungserhebung bei allen sozialwissenschaftlichen Forschungsinstituten im deutschsprachigen Bereich. Die Forschungserhebung ist die mit Abstand am häufigsten genannte „Brücke“ zwischen dem IZ und seinen persönlichen Nutzern. Das Internetangebot der GESIS und die Beschreibung des IZ-Angebots im Prospekt „Produkte und Dienstleistungen“ werden von jeweils 42 Prozent der Nutzer wahrgenommen.

Diejenigen, die das IZ persönlich nutzen, beurteilen die Gesamtleistung des IZ überwiegend positiv: 156 der 169 persönlichen Nutzer, die ein Votum abgaben, wählten die Kategorien „ausgezeichnet“, „sehr gut“ oder „gut. Als „annehmbar“ wurden die Leistungen von lediglich 13 Befragten eingestuft. Keine einzige Person wählte die Kategorie „schlecht“. Die Bereitschaft, die Produkte und Dienstleistungen des IZ an Kollegen/innen weiterzuempfehlen, ist weit verbreitet: 138 Befragte (79 Prozent der Nutzer) würden dies „bestimmt“ oder „wahrscheinlich“ tun. Ein großer Teil der Nutzer (ca. 82 Prozent) wird auch in Zukunft wieder Produkte und Dienstleistungen des IZ in Anspruch nehmen. Etwas mehr als die Hälfte der Nutzer sehen einen „sehr großen“ oder „großen“ Vorteil in der Möglichkeit, das IZ nutzen zu können. Diese Befunde entsprechen weitgehend den Hauptergebnissen der 1997 gemeinsam mit INFRATEST BURKE durchgeführten Befragung zur Zufriedenheit der Mehrfachkunden (TRI:M-Studie).

Die Analyse des Nutzer- und Nutzungspotentials bei den deutschsprachigen Soziologieprofessoren gibt Hinweise auf die relative Bedeutsamkeit von Datenbanken im allgemeinen und von SOLIS und FORIS im besondern unter Berücksichtigung des gesamten Informationsbeschaffungsverhaltens der Befragten. Der Ausbau des Internet-Angebots der GESIS mit der Möglichkeit, vom eigenen PC aus kostenlos in SOLIS und FORIS recherchieren zu können, dürfte zu einer Steigerung der Nut-

zungsintensität führen. Voraussetzung ist allerdings, dass die potentiellen Nutzer umfassend und aktuell über die Angebote des IZ informiert werden.

Die Nutzung und Bewertung des ZA

Ungefähr zwei Drittel der Soziologieprofessorinnen und -professoren, die das ZA kennen, haben seine Leistungen bereits in Anspruch genommen. Von denjenigen Befragten, die das ZA nicht *selbst* nutzen, geben 53 Prozent als Grund der Nicht-Inanspruchnahme an, dass sie nicht empirisch arbeiten. Alle nachfolgenden Zahlen beziehen sich nur noch auf die Personen, die das ZA kennen *und* bereits genutzt haben.

Die bekannteste Dienstleistung des ZA ist der Datenservice (92 Prozent). Es folgen die ZA-Information (90 Prozent), der Datenbestandskatalog (80 Prozent), Workshops und Konferenzen (75 Prozent), das Frühjahrsseminar (72 Prozent), die Methodenberatung (68 Prozent), das Herbstseminar (65 Prozent) sowie die Spezialbibliothek „Empirische Sozialforschung“ (55 Prozent). Am häufigsten genutzt wird die ZA-Information (72 Prozent), der Datenservice (64 Prozent) sowie der Datenbestandskatalog (55 Prozent). 79 Prozent der Befragten fühlen sich ausreichend über das Dienstleistungsangebot des ZA informiert.

52 Prozent der Befragten haben schon einmal einen Datensatz zur Archivierung an das ZA weitergegeben. 79 Prozent haben außerdem schon einmal in den Datenbeständen des ZA nach einem Datensatz gesucht. Die Suche erfolgte dabei telefonisch (26 Prozent), über das Internet-Angebot des ZA (24 Prozent), über den gedruckten Datenbestandskatalog (22 Prozent), schriftlich (17 Prozent) oder persönlich (16 Prozent). 26 Prozent führten die Recherche über einen Mitarbeiter durch. Unabhängig von der konkret gewählten Recherchemöglichkeit waren die Befragten mit deren Qualität im Durchschnitt sehr zufrieden. 67 Prozent konnten einen für ihr Forschungsthema geeigneten Datensatz finden. 46 Prozent waren mit der Aktualität der vom ZA angebotenen Datensätze zufrieden. Nur 23 Prozent der Befragten geben an, dass das ZA den Nutzern die Recherche in seinen Datenbeständen weiter erleichtern könnte. Genannt werden dabei vor allem erweiterte Zugangs- und Recherchemöglichkeiten im Internet, insbesondere aber die Möglichkeit des Downloads von Datensätzen.

Insgesamt 79 Prozent der Befragten haben schon einmal einen Datensatz über das ZA bezogen, 61 Prozent von ihnen innerhalb der letzten drei Jahre. 21 Prozent der Befragten geben an, bislang keinen Datensatz über das ZA bezogen zu haben. Als Gründe hierfür werden vor allem der fehlende Bedarf an empirischen Daten im allgemeinen und an Sekundäranalysen im besonderen genannt. Der am häufigsten

nachgefragte Datensatz ist der ALLBUS. 73 Prozent der Befragten, die bereits einmal einen Datensatz über das ZA bezogen haben, geben an, es habe sich bei einem dieser Datensätze um den ALLBUS gehandelt. Es folgen die Eurobarometer (40 Prozent), der ISSP (38 Prozent), die Jugendstudien (37 Prozent), der Wohlfahrtssurvey (29 Prozent), das Politbarometer (27 Prozent), der World Values Survey (23 Prozent), die Wahlstudien (21 Prozent) und der European Values Survey (13 Prozent).

84 Prozent der Befragten sind mit der Geschwindigkeit, mit der ihnen die bestellten Daten geliefert wurden zufrieden. 82 Prozent geben an, dass die Auswertung der über das ZA bezogenen Daten problemlos möglich gewesen sei. Genutzt werden die über das ZA bezogenen Daten vor allem für Sekundäranalysen mit eigenen Fragestellungen (85 Prozent), den Einsatz in der Lehre (59 Prozent), die Vorbereitung eigener Forschungsprojekte (55 Prozent), zur Ergänzung eigener Daten in zeitlicher Hinsicht (54 Prozent) sowie zur Ergänzung eigener Daten um einen internationalen Vergleich (36 Prozent).

All diejenigen Befragten, die das ZA kennen und bereits in Anspruch genommen haben, wurden gebeten, ihre Erfahrungen mit dem ZA auf einer fünfstufigen Zufriedenheitsskala summarisch zu bilanzieren. Immerhin 37 Prozent der Befragten wählen dabei mit der „5“ den höchsten Zufriedenheitswert. Weitere 43 Prozent wählen mit der „4“ die zweitbeste Bewertung. Eine nur durchschnittliche Bewertung vergaben mit der „3“ noch 10 Prozent, während die beiden Werte „1“ und „2“, die Unzufriedenheit indizieren, jeweils nur von einem Prozent der Befragten gewählt wurden.

29 Prozent der Befragten gaben an, es gebe Dinge, die das ZA besser machen könnte. Auf Nachfrage nannten diese Personen konkret die folgenden Punkte: Mehr Informationen über die Produkte, Dienstleistungen und Datensätze des ZA, ein besseres Internetangebot sowie die Möglichkeit des Downloads von Datensätzen sowie bessere Recherchemöglichkeiten und Beratung.

78 Prozent der Befragten können keine Leistung nennen, die ihrer Ansicht nach vom ZA noch zusätzlich angeboten werden sollte. Von denjenigen Befragten die dies konnten wurde in erster Linie der Download von Daten über das Internet sowie die verstärkte Berücksichtigung qualitativer Daten genannt. 62 Prozent der Befragten kennen keine anderen Einrichtungen, die vergleichbare Produkte und Dienstleistungen wie das ZA anbieten. Diejenigen Befragten, die solche Einrichtungen kennen, nennen im wesentlichen andere Datenarchive außerhalb von Deutschland.

Die Nutzung und Bewertung des ZUMA

ZUMA ist in der deutschsprachigen Profession ein Begriff: 93 Prozent der befragten Professorinnen und Professoren kennen ZUMA. Mehr als die Hälfte (52 Prozent) dieser Personen hat bereits selbst Produkte oder Dienstleistungen von ZUMA in Anspruch genommen, fast die Hälfte (46 Prozent) hat Mitarbeiter beauftragt, ZUMA in ihrem Namen zu nutzen. Auch weisen mehr als die Hälfte der Soziologieprofessorinnen und -professoren, die ZUMA kennen, ihre Studenten auf ZUMA hin.

55 Prozent der Personen, die ZUMA schon in Anspruch genommen haben, haben dies in Form konkreter Beratungsleistungen getan, die meisten davon mehr als einmal. Die Zufriedenheit mit den Leistungen ZUMA's in diesem Bereich ist sehr hoch (Mittelwert 4,4 auf einer Skala von 1 = „gar nicht zufrieden“ bis 5 = „sehr zufrieden“).

35 Prozent der Personen, die ZUMA schon in Anspruch genommen haben, haben dies in Form von Projektbetreuung getan, die meisten bei mehr als einem ihrer Projekte. Die Zufriedenheit mit der Projektbetreuung bei ZUMA ist ebenfalls sehr hoch (Mittelwert 4,4 auf einer Skala von 1 = „gar nicht zufrieden“ bis 5 = „sehr zufrieden“).

78 Prozent der Personen, die ZUMA schon in Anspruch genommen haben, haben Produkte und Dienstleistungen von ZUMA genutzt. Der weitaus größte Teil davon nennt dabei die ZUMA-Nachrichten, gefolgt von ALLBUS und Datenreport. Am wenigsten bekannt sind die ZUMA-Nachrichten Spezial, sonstige ZUMA-Publikationen und das ZUMA-Informationssystem ZIS.

Der Großteil der Befragten, die bereits mit ZUMA zu tun hatten, werden sich „ganz bestimmt“ (45 Prozent) oder „wahrscheinlich“ (28 Prozent) auch in Zukunft wieder an ZUMA wenden oder ZUMA an Kolleginnen oder Kollegen weiterempfehlen (61 Prozent „ganz bestimmt“, 27 Prozent „wahrscheinlich“).

Die Zufriedenheit mit ZUMA insgesamt ist mit 4,1 auf einer Skala von 1 = „gar nicht zufrieden“ bis 5 = „sehr zufrieden“ recht hoch; dennoch gibt es eine Reihe von Verbesserungsvorschlägen, die auf systematischere Öffentlichkeitsarbeit und mehr Information über das Leistungsspektrum ZUMA's abzielen. Man würde sich wünschen, dass ZUMA auch mehr im Bereich von Dienstleistungen allgemeiner Art für die Profession anbieten würde, z. B. im Bereich von Workshops oder durch Vermittlung der Erfahrungen ZUMA's in die Profession hinein. Ein gewisses Interesse an mehr Leistungen im Bereich der qualitativen Forschung wird sichtbar, allerdings in einer nicht allzu starken Häufigkeit.

Alternativen zu ZUMA werden von 22 Prozent der Professorinnen und –professoren genannt, die mit ZUMA bereits zu tun hatten, wobei aber etwa die Hälfte davon auf Institute verweisen, die inhaltliche Arbeit verrichten.

Befragte, die ZUMA zwar kennen, aber selbst bisher noch nicht unmittelbar mit ZUMA zu tun hatten, beziehen ihre Kenntnis vor allem aus dem Kollegenkreis. Wer ZUMA nicht in Anspruch nimmt, hat aufgrund seiner Forschungsausrichtung oder ohne dies näher zu spezifizieren keinen Bedarf an ZUMA-Leistungen. Aber auch (allerdings wenige) vorrangig qualitativ arbeitende Wissenschaftler sehen in ZUMA nicht vorrangig eine Anlaufstelle, vermissen dort eine einschlägige Kompetenz.

Literatur

Binder, G./Klein, M./Porst, R./Stahl, M., 2001: Die GESIS-Potentialanalyse. IZ, ZA und ZUMA im Urteil von Soziologieprofessorinnen und –professoren. GESIS-Arbeitsbericht Nr. 2.

Schneid, M./Stiegler, A., 2000: Pretestbericht Projekt „Potential-Umfrage GESIS“. ZUMA-Projektbericht 03/2000.

Schneid, M./Stiegler, A., 2001: Feldbericht Projekt „Potential-Umfrage GESIS“. ZUMA-Projektbericht 08/2001. Methoden der Empirischen Sozialforschung.

PREIS DER FRITZ THYSSSEN STIFTUNG AN BERNHARD SCHIMPL-NEIMANNS (ZUMA)

Mit einem zweiten Platz wurde Bernhard Schimpl-Neimanns (ZUMA) bei der diesjährigen Preisverleihung des Fritz-Thyssen Preises ausgezeichnet. Er erhielt diese Auszeichnung für seinen Aufsatz „Soziale Herkunft und Bildungsbeteiligung. Empirische Analysen zu herkunftsspezifischen Bildungsungleichheiten zwischen 1950 und 1989“, der in der Kölner Zeitschrift für Soziologie und Sozialpsychologie erschienen ist (Jg. 52, Heft 4: 636-669). Der Preis ist mit 1.000 DM notiert.

Der Fritz-Thyssen Preis ist der einzige Zeitschriftenpreis in den Sozialwissenschaften außerhalb des englischsprachigen Bereichs. Er wurde in diesem Jahr zum zwanzigsten Mal durch eine Jury vergeben.¹ Die Auswahl der Arbeiten erfolgt in zwei Stufen. Die Herausgeber und Redakteure von dreizehn deutschsprachigen Zeitschriften schlagen jeweils bis zu drei Aufsätze vor, die anschließend von der Jury begutachtet werden. Die am Verfahren beteiligten Zeitschriften sind: Angewandte Sozialforschung, Berliner Journal für Soziologie, Geschichte und Gesellschaft, Kölner Zeitschrift für Soziologie und Sozialpsychologie, Leviathan, Österreichische Zeitschrift für Soziologie, Politische Vierteljahresschrift, Schweizerische Zeitschrift für Soziologie, Sociologia Internationalis, Soziale Welt, Zeitschrift für Politik, Zeitschrift für Sozialpsychologie, Zeitschrift für Soziologie.

Für das Jahr 2000 wurden von den Zeitschriftenredaktionen 24 Arbeiten zur Prämierung vorgeschlagen. Die Jury entschloss sich aufgrund der relativen Gleichwertigkeit der in die Endausscheidung gekommenen vier Arbeiten, für diesen Jahrgang keine Arbeit mit einem ersten Preis auszuzeichnen und statt dessen den zweiten und dritten Preis jeweils doppelt zu vergeben.

Neben Bernhard Schimpl-Neimanns erhielt einen weiteren zweiten Preis NEIL FLIGSTEIN (Berkeley) für seine Arbeit: „Verursacht Globalisierung die Krise des Wohlfahrtsstaates?“ (Berliner Journal für Soziologie, Jg. 10, Heft 3, S. 349-378). Einen dritten Preis (dotiert mit jeweils DM 500,-) haben erhalten: JENS ALBER (Konstanz): „Sozialstaat und Arbeitsmarkt. Produzieren kontinentaleuropäische Wohlfahrtsstaaten

¹ Die Mitglieder dieser Jury sind zur Zeit: Prof. Dr. Geißler (Universität-GHS Siegen), PD. Dr. Nunner-Winkler (MPI für Psychologische Forschung, München), Prof. Dr. Scheuch (Universität zu Köln, Vorsitzender), Prof. Dr. Soeffner (Universität Konstanz), Prof. Dr. Tenfelde (Ruhr Universität Bochum), Prof. Dr. Weiß (Universität-GHS Kassel), Prof. Dr. Windolf (Universität Trier).

typische Beschäftigungsmuster? - Gleichzeitig eine Abhandlung über einige Probleme komparativer statistischer Analyse“ (Leviathan, Jg. 28, Heft 4, S. 535-569) und VOLKER BORN SCHIER (Zürich): „Befähigung zur Sozialkapitalbildung und wirtschaftlicher Erfolg im entwickelten Kapitalismus - Neue Evidenzen aus Ländervergleichen 1980 - 1997“ (Schweizer Zeitschrift für Soziologie, Jg. 26, Heft 2, S. 373 - 400).

Die Jury begründet die Vergabe des zweiten Preises an Bernhard Schimpl-Neimanns wie folgt: „Die andere, ebenfalls mit einem zweiten Preis ausgezeichnete Arbeit von SCHIMPL-NEIMANNS ist bezüglich ihrer Fragestellung, ihrer theoretischen Fundierung und der Art der sich darauf beziehenden Datenanalyse der Prototyp eines gelungenen soziologischen Zeitschriftenaufsatzes. Die Arbeit besticht nach Meinung der Jury durch ihre differenzierte Sichtweise und eine klare und ausgewogene Argumentation. Die Methodik der Datenanalyse ist auf dem Stand der Forschung, ohne den Leser mit Statistik zu überfordern. Die einzelnen Schritte im Ablauf der Argumentation sind begründet und jederzeit nachvollziehbar. Dies trifft auch für den kurzen Exkurs zur Verwendung multinomialer Logit-Modelle statt der in der anglo-amerikanischen Forschung üblichen sequentiellen Modelle zu.

Die vom Autor aufgeworfene Frage, ob die Bildungsexpansion in Westdeutschland seit den 60er Jahren den von ihren Promotoren erwarteten Abbau herkunftsspezifischer Ungleichheiten in Gang setzen konnte wird in der Literatur bislang uneinheitlich beantwortet. Keine Zweifel bestehen dabei jedoch an den seit dieser Zeit sich entwickelnden strukturellen Veränderungen der Bildungsbeteiligung, die auch hier in den deskriptiven Befunden kurz herausgestellt werden. Aufgrund der nachfolgenden, systematisch durchgeführten multivariaten Analysen mit sog. multinomialen Logit-Modellen wird die These einer unverändert fortbestehenden sozialen Ungleichheit in der Bildungsbeteiligung letztlich verworfen. Allerdings ist sie zeitlich und je nach Herkunftsindikator unterschiedlich zu beurteilen: Einmal ist der größte Teil des Abbaus von Ungleichheit bereits bis Ende der 70er Jahre eingetreten, danach ist eher eine Stagnation zu beobachten. Zum anderen hängt die Bildungsbeteiligung und damit der Ungleichheitsabbau im Zeitverlauf offensichtlich stärker mit der beruflichen Stellung des Vaters als mit dem Bildungsniveau des Elternhauses zusammen.“

BUCHBESPRECHUNGEN

Das Buch beginnt mit den Worten „Statistik ist ein wichtiger Bestandteil der Sozialwissenschaften...“, ein Bestandteil, der von vielen Studenten wegen des mathematischen Formelapparates nicht besonders geliebt wird. Die Verfasser künden im Vorwort an, Formeln nicht nackt zu präsentieren, sondern zu erklären und zu erläutern. Ohne statistische Kenntnisse gerät man - nach Auffassung der Autoren - in die „Gefahr fehlerhafter Interpretationen und unzutreffender Schlussfolgerungen“.

STEFFEN-M. KÜHNEL & DAGMAR KREBS

Statistik für die Sozialwissenschaften.

Grundlagen Methoden Anwendungen

Rowohlt Taschenbuch Verlag 2001

ISBN 3-499-55639-1, 654 Seiten, DM 34,90

Der Aufbau des Buches ist in die Teile A bis F gegliedert, die ihrerseits jeweils mehrere Kapitel umfassen. Nach den Grundlagen der statistischen Datenanalyse behandelt das Buch in Teil B Beziehungen zwischen Grundgesamtheit und Stichprobe, dem sich in Teil C das Schätzen von Kennzahlen der Grundgesamtheit und das Testen von Hypothesen anschließt. Die Analyse des Zusammenhangs zwischen zwei Merkmalen einschließlich Regression ist Gegenstand von Teil D. In Teil E werden statistische Modelle der Drittvariablenkontrolle vorgestellt und auch Fragen in Bezug auf Korrelation und Kausalität untersucht. Kurze Ausblicke auf multivariate Analyseverfahren bilden den Abschluss in Teil F, dem noch ein kurzer Anhang mit Rechenhilfen und Tabellen folgt.

Beim Lesen des Buches erkennt man schnell, dass die Autoren über viel praktische Erfahrung verfügen und auch den Zwiespalt zwischen Theorie und Praxis aufzeigen. Ein Beispiel dafür ist etwa der Satz: „In der Praxis der sozialwissenschaftlichen Datenanalyse erfolgt häufig aus rein pragmatischen Gründen eine Missachtung des Skalenniveaus“. Eine für den Leser sehr schöne Facette des Buches ist, dass am Ende

jeden Kapitels eine Zusammenfassung der Formeln und eine Erläuterung der verwendeten Begriffe in einem Glossar erfolgt.

Das vorliegende Buch liefert ein breites Spektrum an statistischen Methoden zur Analyse von Daten aus dem Bereich der Sozialwissenschaften, wobei bei Austausch der Beispiele auf wirtschaftswissenschaftliche Daten das Buch genauso für Studierende der Wirtschaftswissenschaften geeignet ist. Gerade die Wahl der Beispiele aus dem ALLBUS ist als positiv anzuerkennen. Nicht empfehlenswert wäre das Buch eigentlich nur für Studenten, die sich für wirtschaftsstatistische Fragestellungen interessieren und sich über Indizes oder Zeitreihen informieren wollen, da diese nicht behandelt werden. Für solche ist das Buch auch nicht geschrieben. Erfreulich ist, dass die Symbolik recht einheitlich durchgehalten wird. Bei der Trennung zwischen Zufallsvariablen und deren Ausprägungen verschwimmen manchmal die Grenzen (z.B. Formel 13.19). Dies wird allerdings in Fußnote 11 auf Seite 199 auch erwähnt. Es kommt auch vor, dass man im Inhaltsverzeichnis bei der Suche nach gängigen Begriffen nicht fündig wird, etwa bei der Suche nach „Prognose“, obwohl im Buch etliches darüber steht, allerdings unter dem Begriff Vorhersagewert. Bei der multiplen Regression wird man auf die Seiten 507 und 538 verwiesen, aber nicht auf Seite 519, wo in (16.9) die eigentliche Definition steht.

Ich bin davon überzeugt, dass das Buch seinen Platz in der Reihe der Lehrbücher in Statistik für Sozialwissenschaftler haben wird und schon bald nach einer zweiten Auflage verlangt wird. Dies liegt sicherlich auch an dem günstigen Preis-Leistungsverhältnis, das dieses Buch als lohnenswert auch zur privaten Anschaffung macht.

SIEGFRIED GABLER

Das Buch von GÖTZ ROHWER & ULRICH PÖTTER versteht sich als Einführungstext in die sozialwissenschaftliche Statistik, grenzt sich aber deutlich von der Vielzahl anderer Statistik-Lehrbücher ab. Programm der Autoren ist nicht in erster Linie die Darstellung von Techniken aus dem konventionellen Kanon der sozialwissenschaftlichen Statistik. Sie verfolgen vielmehr ein kritisches Anliegen bzgl. Inhalten und Formen sozialwissenschaftlicher Statistik. Im Vorwort heißt es dazu gleich zu Beginn: „Ein wesentliches Anliegen des Textes ist es, die Grundbegriffe der Statistik zu begründen und in ihrem begrifflichen Zusammenhang deutlich zu machen. Es wird deshalb vergleichsweise wenig über einzelne Rechentechniken gesprochen. Stattdessen konzentrieren wir uns auf Begriffsbildungen und versuchen, ihren theoretischen Sinn zu verdeutlichen.“

GÖTZ ROHWER & ULRICH PÖTTER
Grundzüge der sozialwissenschaftlichen Statistik.
Grundlagentexte Soziologie
Juventa Verlag, Weinheim und München, 2001
ISBN 3-7799-1470-0, 384 Seiten, DM 49,80

Die vorgestellte statistische Methodenlehre wird von den wissenschaftstheoretischen Grundpositionen der Autoren bestimmt, ohne deren Kenntnis die einzelnen Teile ihres Buches unverständlich bleiben. Ein Aspekt wird näher im Abschnitt „Reflexion von Fragestellungen“ (Seite 40 ff.) erläutert: „... als Methodenlehre verfolgt sie (die Statistik, M.W.) keine inhaltlichen Fragestellungen. Nach einem verbreiteten Selbstverständnis ist sie auch keine Methodenlehre für irgendein inhaltlich abgrenzbares Wissensgebiet, so dass man fragen kann, worauf sich *ihre* Fragestellungen richten. In einer kurzen Formulierung könnte man antworten, dass sich die Fragestellungen der statistischen Methodenlehre darauf richten, *was man mit Daten machen kann*, und zwar mehr oder weniger unabhängig davon, ob die aus einer Beschäftigung mit dieser Frage resultierenden Methoden in irgendeinem inhaltlich bestimmten Kontext zu sinnvollen Überlegungen führen.“ Daraus erklärt sich, wieso der Text in weiten Strecken einen ziemlich hermetischen Eindruck macht. Die kurz vorher unter Berufung auf R.A.Fisher erklärte Absicht (Seite 40), „daß es nicht nur darum geht, Daten übersichtlich darzustellen, sondern daß das Ziel darin besteht, ihren Informationsgehalt für das Nachdenken zugänglich zu machen.“ kommt m.E. in weiten Teilen des Buches zu kurz. Die Frage, wie Verfahren zum und beim Nachdenken eingesetzt werden, das

sprichwörtliche „proof of the pudding“ statistischer Verfahren, wird im Text zugunsten einer, unter der gegebenen Prämisse formal konsistenten Darstellung vernachlässigt.

Der zweite Aspekt ist der konsequente Ausschluss des wahrscheinlichkeitstheoretischen Kalküls. Diese Position wird mit einer Kritik am main stream der statistischen Methodenlehre begründet. Die einzige Konzession an den main stream findet – teilweise – in Teil IV über Stichproben statt, da die verbreitete Praxis von Zufallsstichproben und deren Begründung schlechterdings nicht in Zweifel gezogen werden können. Im übrigen ist die Abstinenz von w-theoretischen Elementen nicht so klar begründet wie die o.g. Position der Unabhängigkeit von inhaltlichen Fragestellungen. Es scheint aber, als ob w-theoretische Begriffe verdächtigt werden, unzulässige Spekulationen ins statistische Denken einzuschmuggeln. Insbesondere dann, wenn die inhaltliche Theorieentwicklung mit Methoden der Hypothesentestung arbeitet. Auf Seite 42 heißt es dazu: „In der Literatur wird manchmal die Auffassung vertreten, daß man einen sehr unmittelbaren Zusammenhang zwischen inhaltlichen Fragestellungen der empirischen Sozialforschung und methodischen Fragestellungen der Statistik ... herstellen könne. Manchmal wird sogar die Ansicht vertreten, daß statistische Methoden gewissermaßen eine ‚Prüfinstanz‘ zur Beurteilung theoretischer Überlegungen ... eingesetzt werden können; einige Autoren sprechen davon, daß mithilfe statistischer Methoden ‚Hypothesen getestet‘ werden können.“ Das geht den Autoren „beim gegenwärtigen Stand der empirischen Sozialforschung, in dem es noch kaum explizit formulierte theoretische Modelle für die begriffliche Repräsentation des gesellschaftlichen Lebens und seiner Veränderungen gibt.“ entschieden zu weit. Ich gebe zu, dass mir ihr Argument, soweit ich es verstehe, nicht einleuchtet, auch wenn man ohne weiteres zugeben kann, dass die konkreten Anwendungen des w-theoretischen Kalküls in den Sozialwissenschaften mitunter kritikwürdig sind. Für die Methodenlehre von Rohwer und Pötter besteht dieses Risiko nicht mehr, da sie w-theoretische Begriffe radikal verbannen.

Das Buch gliedert sich in fünf Hauptteile. Unter der Überschrift „Variablen und Verteilungen“ werden endliche Grundgesamtheiten diskutiert. Der zweite Teil befasst sich mit „Regressionsrechnung“ als einer „der wichtigsten statistischen Methoden in der empirischen Sozialforschung“. Im dritten Teil „Darstellung von Zustandsverläufen“ werden dynamische Prozesse, im wesentlichen Darstellungsformen und Modelle für Übergangsraten zwischen einer endlichen Anzahl von Zuständen, behandelt. Es folgt ein Abschnitt „Stichproben zur Datengewinnung“ und schließlich „Ungenau und unvollständige Daten“. Der erste Teil „Variablen und Verteilungen“ umfasst - auch wenn die Autoren diesen Begriff verwerfen - den Bereich der deskriptiven Statistik. Allerdings

wirkt der Text wie eine Demonstration, dass man über endliche Grundgesamtheiten in den Begriffen der Wahrscheinlichkeitstheorie reden kann, wenn man „Wahrscheinlichkeit“ durch „relative Häufigkeit“ ersetzt. Der Subtext der W -Theorie drängt sich natürlich auf, wenn beispielsweise die Tschebyscheffsche Ungleichung (Seite 77 f.) oder den Begriff der bedingten Verteilung (Seite 101 ff.) in den Kontext finiter Grundgesamtheiten transponiert werden. Dies dient aber auch der Vorbereitung der nachfolgend dargestellten Regressionsrechnung und der Zustandsverläufe.

Die Vermeidung probabilistischer Formulierungen kann für die Operationen der Regressionsrechnung einigermaßen weitgehend durchgehalten werden, die Abgrenzung vollzieht sich hier eher in der Interpretation von Analyseergebnissen. Die Entwicklung der Regressionsrechnung orientiert sich wie im ersten Teil an den Begriffen der W -Theorie. Hier ist es der bedingte Erwartungswert bzw. die bedingten Wahrscheinlichkeit, die in deterministischer Lesart als bedingte relative Häufigkeit verwendet wird. Auch in dieser Version ist dieser Begriff der Regression wesentlich umfassender als die übliche lineare Regression zur Modellierung von Mittelwerten. Trotz der damit verbundenen formal anspruchsvolleren Darstellung wird durch den allgemeinen Begriff und die Betonung des Unterschieds zwischen der Regression als bedingter Verteilung und speziellen parametrischen Funktionen, mit denen die Abhängigkeit bestimmter Parameter der bedingten Verteilung von den Regressorvariablen modelliert wird, auch der allgemeine Anwendungsrahmen aufgezeigt. Positiv hervorzuheben ist hier auch Kapitel 10 „Zum Verständnis von Regressionsmodellen“, in dem der Unterschied zwischen der Residualvariablen des bedingten Verteilung und dem „Störterm“ in parametrischen Modellfunktionen diskutiert wird.

Die Ablehnung einer probabilistischen Fassung der Regression führt nun allerdings nicht nur dazu, dass die Autoren auch nicht von Prognosen sprechen können und wollen, auch die Frage nach der Signifikanz von Effekten ist in ihrem Begriffsrahmen nicht möglich. Dies ist m.E. nicht nur ein echtes Manko, es steht in gewisser Weise auch im Gegensatz zu Gesamtkonzept des Buchs, das eben auch Zufallsstichproben umfasst, deren Daten i.d.R. die Basis für Schätzungen der Regressionsparameter bilden.

Die Darstellung von Zustandsverläufen bildet Teil III des Buchs und ist dem gleichen wissenschaftstheoretischen Dogma unterworfen wie die Regressionsrechnung. Im übrigen bietet dieser Text aber einen durchaus lesbaren Überblick über die Statistik von Zeitverlaufsdaten. Auch im Teil IV über Gesamtheiten und Stichproben ist zwar von zufälligen Stichproben die Rede, der Boykott der W -Theorie erlaubt aber nicht, die Qualität von Stichproben aus Ziehungsverfahren abzuleiten. Was bleibt, ist die Hoff-

nung auf Repräsentativität, d.h. dass relative Häufigkeiten zwischen Stichprobe und Grundgesamtheit übereinstimmen. Stichprobenfehler – sonst eine wichtige Größe für die Bestimmung von Ziehungsdesigns – haben in der Theorie von Rohwer und Pötter keinen Platz. Es wird gleichwohl der Versuch gemacht, Schätzfunktionen als Zufallsvariable zu behandeln und w-theoretische Begriffe wie Erwartungstreue und Konfidenzintervalle einzuführen. Leider teilt sich dem Leser deutlich das Unbehagen mit, das die Autoren bei der Abfassung dieses Teils ihres Buches überkommen haben dürfte und der nur eine pflichtschuldigen Tribut an allgemeine Gepflogenheiten darzustellen scheint.

Der letzte Teil V über „Ungenau und unvollständige Daten“ befasst sich mit Fragen der Genauigkeit von gemessenen Variablen, die unterschiedlichen Fehlerquellen bei der Messung, aber auch grundsätzlichen „Unschärfen“ geschuldet sein könnten. Da natürlich auch hier w-theoretische Begriffe nichts zu suchen haben, weichen die Autoren auf mengenwertige z.B. intervallwertige Variable aus. Über der abstrakten Begriffsbildung wird hier aber die Perspektive der Anwendung bei realen Datenanalyseproblemen nicht recht deutlich. Wie beispielsweise werden für eine bestimmte Beobachtung die Genauigkeitsgrenzen erhoben? Der „Genauigkeitsgrad“ (Seite 343) einer empirischen Variablen, der für jede Beobachtung als Durchmesser des Intervalls der möglichen Werte definiert ist, erweist sich jedenfalls in den beispielhaften Anwendungen nur dadurch handhabbar, dass er einfach einen angenommenen Zahlenwert erhält. Insofern ähneln die Überlegungen dieses Teils denen von Sensitivitätsrechnungen, bei denen ebenfalls mit hypothetischen Werten für Variable gearbeitet wird, um den Effekt von Abweichung bei Schätzungen zu studieren.

Die eingangs geschilderte wissenschaftstheoretische Position der Autoren wird mit wechselnder Stringenz durchgehalten. Dem Buch kommt entgegen, dass sich probabilistische Formulierung und ihre deterministische Lesart für finite Gesamtheiten nicht wesentlich unterscheiden. Inferenzstatistische Verfahren sind beim Ansatz von Rohwer und Pötter bewusst ausgeschlossen. Leider rangiert generell der Anwendungs- und Problemlösungsaspekt bei der Entwicklung und Darstellung in der zweiten Reihe.

MICHAEL WIEDENBECK

Das Buch, das inzwischen in der 2. Auflage vorliegt, hat sich als Ausgangspunkt den im Titel erwähnten „Sozialarbeiter“ und seinen professionellen Statistikbedarf gewählt. Dabei darf der „Sozialarbeiter“ nicht wörtlich verstanden werden, er kann ohne weiteres durch jede Profession ersetzt werden, bei der es auf den kompetenten Umgang mit Daten ankommt, d.h. sowohl auf kritische Lektüre anderweitig erstellter Statistiken als auch auf eigene Analysen mit Hilfe deskriptiver und einfacher inferenzstatistischer Methoden. Es wendet sich an alle, die lernen wollen, mit Hilfe empirischen Materials zwischen Handlungsalternativen zu entscheiden, wozu nach Auffassung der Verfasser auch gehört, die Grenzen der eigenen Kompetenz zu erkennen und gegebenenfalls statistische Experten zu Rate zu ziehen.

RÜDIGER OSTERMANN & KARIN WOLF-OSTERMANN
Statistik für Studierende der Sozialarbeit und Sozialpädagogik.
Oldenbourg: München, 2. Auflage, 1999
ISBN 3-486-25203-8, 404 Seiten, DM 48,50 (€ 24,80)

Das Buch umfasst drei große Bereiche: Deskriptive Statistik, Einführung in die Wahrscheinlichkeitsrechnung und inferenzstatistische Verfahren. Es schließen sich die Tabellen einiger bekannter Verteilungen und ein ausführliches Kapitel mit den Lösungen der Übungsaufgaben sowie mit Probeklausuren und den zugehörigen Lösungen an.

Die Darstellung der deskriptiven Statistik beginnt mit der Definition grundlegender Begriffe wie Untersuchungseinheiten und Merkmale, Variablen unterschiedlicher Skalenniveaus und deren Häufigkeitsverteilungen in finiten Grundgesamtheiten. Es schließen sich die Beschreibung der Variablen und ihrer Häufigkeitsverteilungen durch Konzentrationsmaße, Lage- und Skalenparameter sowie die verschiedenen graphischen Darstellungsmöglichkeiten der Verteilungen an. Die Darstellung beschränkt sich dabei im wesentlichen auf univariate Statistik, ergänzt aber den deskriptiven Teil um ein letztes Kapitel über zweidimensionale Häufigkeitsverteilungen und verbindet damit die Diskussion verschiedener Assoziationsmaße zweier Variabler.

Der Text ist - und das lässt sich generell für das gesamte Buch feststellen - sehr klar und verständlich verfasst. Auch die formalen Teile sind exakt ohne überflüssigen Ballast, die auch für einen Anfänger in der Statistik keine besondere Schwierigkeit bedeuten dürften. Anhand vieler Fallbeispiele aus sozialpädagogischer Berufspraxis

wird die Anwendung statistischer Argumente demonstriert. In allen Beispielen wird deutlich, wie durch die vorgestellten Verfahren Daten auch kleinen Umfangs transparenter werden als durch die bloße Inspektion des Datenrohmaterials. Die Beispiele zeigen aber nicht nur den Zugewinn an Übersichtlichkeit etwa durch graphische Darstellungen, es werden auch Fragen zum Vergleich von Einzeldaten mit Stichproben oder zum Vergleich zwischen verschiedenen Stichproben mit Mitteln der deskriptiven Statistik behandelt. Der Vergleich einer quantitativen Größe, etwa einer Mengenangabe, erlaubt z.B. erst im Vergleich mit der Ordnungsstatistik von Angaben einer Stichprobe vergleichbarer Untersuchungseinheiten, ob die einzelne Mengenangabe als klein oder groß angesehen werden kann. In einem anderen Beispiel wird die Einheitlichkeit zweier Personengruppen bzgl. bestimmter Merkmale ihrer Mitglieder durch den Vergleich von Konzentrationsmaßen der beiden Verteilungen der Merkmale verglichen. In diesen Beispielen wird in sehr anschaulicher Weise deutlich, dass die Beantwortung von Fragestellungen mit Hilfe von Daten immer von der Definition bestimmter Kenngrößen bzw. von den angewendeten Verfahren abhängt. Die genaue Kenntnis davon, welche Charakteristiken der Daten durch verschiedenen Kenngrößen abgebildet werden und wovon andererseits abstrahiert wird, welche Kenngrößen für die gleiche Fragestellung herangezogen werden können und wodurch sich die Alternativen unterscheiden, ist Grundlage für eine angemessene Beurteilung statistischer Befunde.

Der mittlere, relativ kurze Teil zur Wahrscheinlichkeitsrechnung dient im wesentlichen der Vorbereitung des anschließenden inferenzstatistischen Teils und ist entsprechend anwendungsorientiert abgefasst. Der Begriff der Wahrscheinlichkeit ist bekanntlich schwierig, nur im Rahmen einer umfangreichen Theorie einigermaßen angemessen zu entfalten und obendrein im Bereich der Statistik und ihrer Schlussweisen auch noch strittig. Von einer elementaren Einführung kann man daher nicht mehr erwarten als sich der vorliegende Text vornimmt: erste Formalisierung des Begriffs der Wahrscheinlichkeit und plausible Interpretation der Kolmogorov'schen Axiome sowie die daraus folgenden Rechenregeln. Bedingte Wahrscheinlichkeit und Unabhängigkeit werden eingeführt und anhand einfacher Würfelexperimente diskutiert. Der Stellenwert der bedingten Wahrscheinlichkeit für Schlüsse auf die Wahrscheinlichkeit von Ereignissen gegeben bestimmte Informationen/Daten wird mit dem Satz von der totalen Wahrscheinlichkeit und dem Theorem von Bayes unterstrichen. Die Diskussion von diskreten und stetigen Zufallsvariablen, ihren Verteilungen und die Einführung der Normal-, t -, F - und Chi-quadrat-Verteilungen runden den Abschnitt über Wahrscheinlichkeitsrechnung ab.

Der dritte, inferenzstatistische Teil des Buches handelt von statistischen Testverfahren und Regressionsrechnung. Die Darstellung der Testverfahren beginnt wieder mit ausführlichen Fallbeispielen, in denen konkrete Fragestellungen und Untersuchungen vorgestellt werden. Der Leser lernt auch hier die Situation für die Anwendung von Tests und die Probleme einer Entscheidung über konkrete Datenbeispiele kennen. Erst dann wird der formale Rahmen von Tests als Entscheidungsverfahren über Hypothesen auf der Grundlage von Daten (Hypothese und Gegenhypothese, Fehler 1. und 2. Art etc.) erläutert. Die Darstellung ist auch hier sehr bemüht, den in formalen Darstellungen ungeübten Leser durch illustrierende Beispiele „mitzunehmen“.

Nach dieser Einführung in Sinn und Zweck statistischer Tests werden systematisch spezielle Tests zu ausgewählten Verfahren abgehandelt: Mittelwertvergleiche bei unverbundenen und verbundenen Stichproben, Tests auf Gleichheit von Varianzen in unterschiedlichen Stichproben, Testung von Korrelationskoeffizienten und schließlich verschiedene Tests auf Unabhängigkeit in Kontingenztafeln, wobei die zugrundeliegenden Verteilungsmodelle Normalverteilungsmodelle mit stark vereinfachenden Annahmen sind, die dann schrittweise bei zunehmend realistischeren Annahmen zu komplexeren Modellen (wie der t-Verteilung) abgeändert werden. Alle Verfahren werden wieder extensiv durch Anwendungsbeispiele erläutert.

Die abschließende Einführung in die Regressionsrechnung ist mit der gleichen didaktischen Sorgfalt verfasst. Behandelt werden einfache und multiple Regression, die OLS-Schätzung, das Bestimmtheitsmaß und die Testung der Effekte unter der Annahme normalverteilter Residuen. Besonders hervorzuheben ist, dass der graphischen Residuenanalyse ein eigenes Kapitel gewidmet ist, in dem die Notwendigkeit und Möglichkeit der Überprüfung der Modellannahmen demonstriert wird.

Das Buch von Ostermann ist m.E. eine didaktisch sehr gelungene Einführung in die Anfangsgründe der Statistik, die man nicht nur den angesprochenen Sozialarbeitern und -pädagogen zur Lektüre empfehlen möchte, sondern auch Studierenden der empirischen Sozialwissenschaften. Der Umfang der dargestellten Verfahren ist hier ein Großteil dessen, was in der täglichen Arbeit benötigt wird.

Der Problemkreis der Datenerhebung ist jedoch kein Schwerpunkt des Buchs, wenngleich im Zusammenhang mit der Einführung der grundlegenden Begriffe auf Probleme der Stichprobenziehung Bezug genommen wird (Seite 13 ff.) bzw. einige Anmerkungen zu fehlenden Werten im Kapitel über Lageparameter gemacht werden (Seite 37 ff.) Die Erziehung zu aufgeklärten Statistik-Konsumenten ist allerdings ohne eine Vorstellung davon, wie und in welchem Ausmaß der Prozess der Datenerhebung

die Qualität der Daten beeinflussen kann, unvollständig. In künftigen Auflagen könnte dieser Punkt mit großem Nutzen ergänzt werden.

MICHAEL WIEDENBECK

PUBLIKATIONEN

ZUMA-ARBEITSBERICHTE

Die ZUMA-Arbeitsberichte informieren über das gesamte Spektrum der Arbeit bei ZUMA und enthalten überwiegend Beiträge, die zur Veröffentlichung in Fachzeitschriften vorgesehen sind. Nachfolgend werden die ZUMA-Arbeitsberichte, die seit der letzten Ausgabe der ZUMA-Nachrichten publiziert worden sind, in Form von Abstracts kurz vorgestellt. ZUMA-Arbeitsberichte sind auch im Internet als PDF-Datei abrufbar (www.gesis.org/publikationen/berichte/zuma_arbeitsberichte/). Einzelhefte können unter folgender Adresse bestellt werden:

Zentrum für Umfragen, Methoden und Analysen
ZUMA-Publikationen
Postfach 12 21 55
68072 Mannheim

* * * * *

ZUMA-Arbeitsbericht 2001/03: Georgios Papastefanou: Paying attention to eco-labels in purchase decisions: socio-economic and demographic determinants.

Eco-labeling of consumer goods is an important measure to support sustainability in society but there is little information on how eco-label awareness depends on social structure. This paper examines social structural differentiation of consumer's eco-label information awareness in purchase decisions. Survey data from 1993 are analyzed with logistic regression analysis. Eco-label attention is modeled as a non-linear function of age, employment status, size of residence, education and subjective social class.

ZUMA-METHODENBERICHTE

In der ZUMA-Reihe „Methodenberichte“ werden Beiträge publiziert, die wichtige Aspekte der Arbeit von ZUMA dokumentieren, aber in der Regel nicht für eine Veröffentlichung in einer Fachzeitschrift geeignet sind, wie z.B. Dokumentationen, Bibliographien, Literaturberichte und Datenbeschreibungen. Die Berichte sind im Internet als PDF-Datei abrufbar (www.gesis.org/publikationen/berichte/zuma_methodenberichte/). Einzelhefte können auch bei folgender Adresse bestellt werden:

Zentrum für Umfragen, Methoden und Analysen
ZUMA-Publikationen
Postfach 12 21 55
68072 Mannheim

* * * * *

ZUMA-Methodenbericht 2001/04

Bernhard Schimpl-Neimanns & Ulrich Rendtel
SAS-, SPSS- und STATA-Programme zur Berechnung der Varianz von
Populationsschätzern im Mikrozensus ab 1996

ZUMA-Methodenbericht 2001/05

Achim Koch, Martina Wasmer, Janet Harkness & Evi Scholz
Konzeption und Durchführung der Allgemeinen Bevölkerungsumfrage der
Sozialwissenschaften (ALLBUS) 2000

ZUMA-Methodenbericht 2001/06

Michael Blohm & Ursula Lamas Perez
ALLBUS-Bibliographie (17. Fassung; Stand: Juni 2001)

ZUMA-Methodenbericht 2001/07

Birgit Neugebauer & Rolf Porst
Patientenzufriedenheit. Ein Literaturbericht.

ZUMA-HOW-TO-REIHE

In der ZUMA-Reihe „How-to“ werden ausgewählte Themen aus dem Bereich der sozialwissenschaftlichen Methodik anwendungsorientiert dargestellt. Der Schwerpunkt liegt auf der Vermittlung von Know-how, das für den Anwender relevant, wichtig und nützlich ist. Die Berichte sind im Internet als PDF-Datei abrufbar (www.gesis.org/publikationen/berichte/zuma_how_to/). Einzelhefte können auch bei den jeweiligen Autoren bestellt werden.

* * * * *

How-to Nr. 8: *Cornelia Züll & Peter Ph. Mohler: Computerunterstützte Inhaltsanalyse: Codierung und Analyse von Antworten auf offene Fragen.*

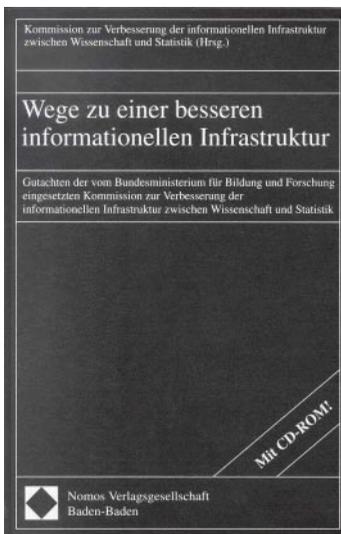
Das Vorgehen bei der Verwendung der computerunterstützten Inhaltsanalyse zur Codierung von Antworten auf offene Fragen eines Fragebogens wird dargestellt. Der Schwerpunkt liegt dabei auf der Beschreibung der Entwicklung und der Validierung eines Kategoriensystems/Diktionärs zur automatischen Codierung.

How-to Nr. 9: *Rolf Porst: Wie man die Rücklaufquote bei postalischen Befragungen erhöht.* Der Beitrag beschäftigt sich mit Strategien und Maßnahmen zur Erhöhung der Rücklaufquote bei postalischen Befragungen. Dabei werden sowohl Designaspekte wie auch Aspekte der Befragungsmaterialien behandelt.

How-to Nr. 10: *Michael Wiedenbeck & Cornelia Züll: Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren.*

Nach einer Einführung in die Ziele der Clusteranalyse werden die Grundprinzipien der Algorithmen hierarchisch-agglomerativer und K-means-Verfahren dargestellt. Ein Schwerpunkt liegt auf der graphischen Darstellung der Ergebnisse. Außerdem werden einige Verfahren zur Validierung von Clusterlösungen, wie der Vergleich von Lösungen hierarchisch-agglomerativer Verfahren mit K-means-Lösungen sowie Monte-Carlo-Verfahren zur Exploration des Einflusses von Startbedingungen bei K-means-Verfahren, vorgestellt.

WEGE ZU EINER BESSEREN INFORMATIONELLEN INFRASTRUKTUR



U nter anderem aus Anlass des schwierigen Zugangs zu Mikrodaten der amtlichen Statistik in Deutschland wurde von der Bundesministerin für Bildung und Forschung eine „Kommission zur Verbesserung der informationellen Infrastruktur“ eingesetzt. Im Frühjahr 2001 hat die Kommission unter Vorsitz von Prof. Dr. Hans-Jürgen Krupp und dem Präsidenten des Statistischen Bundesamtes, Johann Hahlen, ihre Empfehlungen übergeben, die in diesem Band veröffentlicht sind. Themenschwerpunkte des Gutachtens sind u.a. der Bedarf an statistischen Daten in den verschiedenen Informationsfeldern, eine Bestandsaufnahme der gegenwärtigen informationellen Infrastruktur in Deutschland, der Zugang der Wissenschaft zu Mikrodaten und Empfehlungen für eine bessere informationelle Infrastruktur in Deutschland. Das Buch enthält auch eine CD-ROM mit mehr als 40 Expertisen von Wissenschaftlern und Datenanbietern.

Wege zu einer besseren informationellen Infrastruktur.
Nomos Verlagsgesellschaft 2001, Baden-Baden
297 Seiten, 98 DM, ISBN 3-7890-7388-1

ZUMA-WORKSHOPS

Erstes Halbjahr 2002

Thema	Datum
Stichproben in Theorie und Praxis	26. Februar
Questionnaire Translation	11. - 12. März
Eine Einführung in die Analyse linearer Strukturgleichungsmodelle mit LISREL 8.50	6. - 7. Juni
Umfragen in den Massenmedien: Ein Grundkurs für Journalisten	14. Juni
Das Internet als Setting für experimentelle Forschung in den Sozialwissenschaften	20. - 21. Juni
Eine Einführung in Datenmanagement und Datenrecodierung	26. - 27. Juni
Fragen verstehen und beantworten: Kognitive und kommunikative Grundlagen von Befragungen	16. - 17. Juli
Simulation for Social Science	23. - 27. September

Hinweise zu weiteren Veranstaltungen der GESIS finden Sie auch unter www.gesis.org/veranstaltungen/

* * *

Stichproben in Theorie und Praxis

26. FEBRUAR 2002

Um Informationen über Personen in einer Population zu gewinnen, werden häufig Stichproben gezogen. Die Art und Weise der Auswahl wird dabei durch die verschiedenen Stichprobenverfahren geregelt. In den Sozialwissenschaften sind dies vor allem zufällige Auswahlverfahren, aber auch Quotenverfahren. Beide unterscheiden sich durch die zugrundeliegenden statistischen Theorien, die wesentliche Implikationen für

die Aussagefähigkeit der gewonnenen Daten haben. Bei der Wahl der Verfahren spielen jedoch auch Kosten- und Zeiterwägungen eine Rolle. In Form von Vorträgen werden mathematisch-statistische Grundlagen und praktische Anwendungen der Stichprobentheorie behandelt und Gelegenheit zu Diskussionen gegeben. Folgende Fragen sollen insbesondere erörtert werden: Welchen Einfluss hat die Grundgesamtheit auf die Verwendung des Stichprobendesigns? Welche Vor- und Nachteile bieten die verschiedenen Auswahlprozeduren? Wo werden die Verfahren in der Praxis eingesetzt? Wie groß sollte der Stichprobenumfang gewählt werden? Welche Rolle spielen Nonresponse und Gewichtung? Was ist der Designeffekt? Der Workshop richtet sich an Personen, die eine erste Ausbildung in Methoden der empirischen Sozialforschung absolviert haben und ihre Kenntnisse in Stichproben vertiefen wollen.

Die Teilnehmerzahl ist auf 20 Personen begrenzt. Interessenten werden gebeten, sich bis zum 15. Januar 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621-1246-221). Für die Teilnahme wird ein Beitrag von 35 Euro erhoben. Der Workshop wird von SIEGFRIED GABLER, SABINE HÄDER und MICHAEL WIEDENBECK geleitet.

* * *

Questionnaire Translation

MARCH 11-12, 2002

The two-day workshop is intended for people involved in questionnaire translation – those responsible for organising and overseeing translations and those involved in producing and assessing translations. It will be held in English. Contents of workshop: Organising translations, writing translation specifications; materials for training or accessing translators; overview of basic translation procedures for questionnaires; practical demonstration of common problems; overview of translation assessment procedures; documenting questionnaires. The workshop will be taught by an international team of expert survey translators and co-ordinators including ALICIA SCHOVA GLUSBERG, Director of Survey Operations, PHDCN, Harvard School of Public Health, BETH ELLEN PENNELL, Associate Director of Survey Operations, Survey Research Centre, University of Michigan, AND FONS VAN DE VIJVER (cultural psychology) Tilburg University, Netherlands.

Maximum number of participants: 18. Registration deadline: 25.1.2002. Fee: (includes materials) Euro 60. Please register with our workshop secretariat and indicate what languages you understand (workshop@zuma-mannheim.de, Tel. (INT) + (0)621-1246-221). Workshop co-ordinator: JANET HARKNESS.

* * *

Eine Einführung in die Analyse linearer Strukturgleichungsmodelle mit LISREL 8.50

6. UND 7. JUNI 2002

Das von Karl Jöreskog und Dag Sörbom entwickelte LISREL-Programm ist das bekannteste Statistiksystem zur Analyse von linearen Strukturgleichungsmodellen. In seiner aktuellen Version LISREL 8.50 können konfirmatorische Faktorenanalysen und Pfadmodelle mit beobachteten und latenten Variablen auf der Basis metrischer oder ordinaler Messungen, simultanen Gruppenvergleiche und Analysen von Mittelwertstrukturen berechnet werden, wobei unterschiedliche Schätzmethoden zur Verfügung stehen. Erweitert wurde die Behandlung fehlender Fälle durch die Implementation verschiedener Imputationsverfahren und die ML-Schätzung mit unvollständigen Daten. Neben den klassischen Strukturgleichungsmodelle enthält das Programmsystem auch Module zur explorativen Datenanalyse und zu Mehrebenenmodellen. In dem zweitägigen Workshop wird neben einer Einführung in das Programm auch die Logik der statistischen Modellierung vorgestellt. Es werden keine Vorerfahrungen in der Anwendung linearer Strukturgleichungsmodelle vorausgesetzt.

Die Teilnehmerzahl ist auf 14 Personen begrenzt. Interessenten werden gebeten, sich bis zum 25. April 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621-1246-221). Für die Teilnahme wird ein Beitrag von 150 Euro erhoben. Der Workshop wird von STEFFEN KÜHNEL (Universität Göttingen) geleitet.

Umfragen in den Massenmedien: Grundkurs für Journalisten

14. JUNI 2001

In Presse und Fernsehen wird immer häufiger auf die Ergebnisse von Umfragen Bezug genommen. Es ist abzusehen, dass die Demoskopie im Vorfeld der Bundestagswahl 2002 einen wichtigen, möglicherweise auch umstrittenen Gegenstand der innenpolitischen Berichterstattung bilden wird. Publiizierte Umfrageergebnisse sind oft informativ, manchmal aber auch desinformativ. In dieser Veranstaltung, die sich primär an Journalisten richtet, wird ein Einblick in die Praxis der politischen Umfrageforschung gegeben. Es wird diskutiert, wie Umfragen durchgeführt werden, anhand welcher Kriterien man gute von schlechten Umfragen unterscheiden kann und wie man die Ergebnisse von Umfragen sinnvoll interpretiert. Ziel ist die Vermittlung von Basiskompetenzen, um die Aussagekraft von Umfragen bewerten und ihre Befunde adäquat vermitteln zu können.

Die Teilnehmerzahl ist auf 20 Personen begrenzt. Interessenten werden gebeten, sich bis zum 3. Mai 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621-1246-221). Für die Teilnahme wird ein Beitrag von 80 Euro erhoben. Der Workshop wird von RÜDIGER SCHMITT-BECK geleitet.

* * *

Das Internet als Setting für experimentelle Forschung in den Sozialwissenschaften

20. UND 21. JUNI 2002

Wie können die Vorteile der Internet-basierten Forschung mit denen der experimentellen Methode verbunden werden? Web-Experimente gibt es seit 1995. Sie werden seither in immer breiterem Umfang und von einem immer größer werdenden Forscherkreis eingesetzt. Beispiele finden sich im Web-Labor für Experimentelle Psychologie (<http://www.genpsy.unizh.ch/Ulf/Lab/WebExpPsyLabD.html>). In den letzten Jahren wurde in der Anwendung von Web-Experimenten eine Reihe von Techniken entwickelt, die dabei helfen sollen, die

Vorteile der Methode zu nutzen und die Datenqualität bei der Forschung über das Internet zu sichern. Zu diesen Techniken gehören zum Beispiel Ernsthaftigkeitschecks, die „Hohe Hürde“-Technik, die Warm-up-Technik, Sub-sampling, die „Multiple site entry“-Technik, Passwort-Verfahren, nicht-offensichtliche Dateibenennung, Motivationstechniken, abbruchreduzierendes Design, die Kontrolle multipler Teilnahmen und die Kontrolle motivationaler Konfundierung. Der Workshop wird in diese Techniken und die Benutzung von Werkzeugen zur Erstellung von Web-Experimenten einführen (z.B. WEXTOR <http://www.genpsylab.unizh.ch/wextor/index.html>).

Dieser Workshop stellt eine Kooperation zwischen der Deutschen Gesellschaft für Psychologie DGPs (verantwortlich FRIEDRICH WILKENING) und ZUMA (verantwortlich JÜRGEN H. P. HOFFMEYER-ZLOTNIK) dar. Der Workshop wird von ULF REIPS (Universität Zürich) durchgeführt. Die Teilnehmerzahl ist auf 20 Personen begrenzt. Interessenten werden gebeten, sich bis zum 3. Mai 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621–1246–221). Für die Teilnahme wird ein Beitrag von 60 Euro erhoben.

* * *

Eine Einführung in Datenmanagement und Datenrecodierung

26. UND 27. JUNI 2002

Anhand eines konkreten Datensatzes soll demonstriert werden, wie die Daten vom Fragebogen bis zum analysefähigen Datensatz zu bearbeiten sind über die Schritte: Codeplanerstellung, (Vercodung,) Datenerfassung, Datensatzerstellung, Recodierung von Daten zu analysefähigen Variablen. Der Schwerpunkt liegt auf der Recodierung von Daten, die am konkreten Datensatz nicht nur demonstriert werden soll, sondern auch geübt werden kann. Die Handhabung geschieht mit den Programmen WORD und SPSS.

Die Teilnehmerzahl ist auf 20 Personen begrenzt. Interessenten werden gebeten, sich bis zum 3. Mai 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621–1246–221). Für die Teilnahme wird ein Beitrag von 60 Euro

erhoben. Der Workshop wird von JÜRGEN H.P. HOFFMEYER-ZLOTNIK (ZUMA) und FRANZ BAUSKE (ZA) geleitet.

* * *

Fragen verstehen und beantworten: Kognitive und kommunikative Grundlagen von Befragungen

16. UND 17. JULI 2002

Umfragedaten sind nur so nützlich wie die Antworten, die Befragte zu Protokoll geben. Die kognitiven und kommunikativen Prozesse, die der Datenerhebung zugrunde liegen, sind daher von essentieller Bedeutung. Der Workshop gibt einen Überblick über relevante theoretische Konzeptualisierungen, experimentelle Befunde und ihre praktischen Implikationen. Zentrale Themen sind: (1) Probleme des Fragenverständnisses; (2) die Validität retrospektiver Berichte; (3) die Entstehung von Kontexteffekten in der Einstellungsmessung; (4) die Nutzung kognitiver Methoden in der Fragebogenentwicklung.

Die Teilnehmerzahl ist auf 20 Personen begrenzt. Interessenten werden gebeten, sich bis zum 31. Mai 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 0621-1246-221). Für die Teilnahme wird ein Beitrag von 80 Euro erhoben. Die Leitung hat NORBERT SCHWARZ (University of Michigan, USA).

* * *

Simulation for the Social Sciences

23. BIS 27. SEPTEMBER 2002

This will be a workshop on the techniques of developing simulations to help with the exploration and understanding of social and economic issues. It will provide a rationale for using simulation in the social sciences and outline a number of approaches to social simulation at a level of detail that would enable participants to understand the literature and, for some selected approaches, to develop their own simulations. The workshop covers the basics of modelling and simulation in the social sciences from

different points of view (mathematics, computer science, philosophy of science) and of seven different approaches to computer simulation in the social sciences. This workshop is co-organized with ZUMA Mannheim and will be held at the **University of Koblenz-Landau: in Koblenz, Metternich Campus**. An overview of the module can be found at <http://www.uni-koblenz.de/~kgt/SICSS/ModSimNew.html>.

Der Workshop wird von KLAUS G. TROITZSCH (Universität Koblenz-Landau) und NIGEL GILBERT (Universität Surrey) geleitet. Die Teilnahme ist auf 15 Personen begrenzt. Interessenten werden gebeten, sich bis zum 10. Juli 2002 beim ZUMA-Tagungssekretariat anzumelden (workshop@zuma-mannheim.de, Tel. 06 21- 1246-221). Für die Teilnahme wird ein Beitrag von 100 Euro erhoben.