

**Herausgeber:**

Zentralarchiv für Empirische Sozialforschung  
Universität zu Köln

Das Zentralarchiv ist Mitglied der GESIS

Direktor: Prof. Dr. Wolfgang Jagodzinski  
Geschäftsführer: Dr. h.c. Ekkehard Mochmann

**Postanschrift:**

Postfach 410 960  
50869 Köln

**Hausanschrift:**

Bachemer Straße 40  
50931 Köln

**Telefon:**

Zentrale 0221 / 4 76 94 - 0  
Telefax - 44  
Redaktion - 50

Redaktion:

Franz Bauske

E-mail: [bauske@za.uni-koeln.de](mailto:bauske@za.uni-koeln.de)

Internet: <http://www.gesis.org/za>

ISSN: 0723-5607

© Zentralarchiv

Die ZA-Information erscheint jeweils im Mai und November eines Jahres.

Sie wird kostenlos an Interessenten und Benutzer des Zentralarchivs abgegeben.

## Inhaltsverzeichnis

Mitteilungen der Redaktion .....	5
Begrüßung zur Akademischen Gedenkfeier der Universität zu Köln am 23.4.2004 für <i>Erwin K. Scheuch</i> (*9.6.1928 †12.10.2003) von <i>Heiner Meulemann</i> .....	6
<i>Erwin K. Scheuch</i> und seine Bedeutung für die empirische Soziologie in Deutschland von <i>Franz Urban Pappi</i> .....	8
<i>Erwin K. Scheuch</i> als akademischer Lehrer, Vorgesetzter und Kollege. Ansprache zur Akademischen Gedenkfeier für <i>Erwin K. Scheuch</i> an der Universität zu Köln am 23.4.2004 von <i>Heinrich Best</i> .....	17

---

## Forschungsnotizen

Online-Umfragen – schnell, billig, aber auch valide? Ein Vergleich zweier Internetbefragungen mit persönlichen Interviews zur Bundestagswahl 2002 von <i>Harald Schoen</i> .....	27
Die angemessene Erfassung der Stärke und Richtung von Anreizen durch soziale Erwünschtheit von <i>Volker Stocké</i> und <i>Christian Hunkler</i> .....	53
Unterschiede zwischen Wertorientierungen nach Geschlecht, Alter, Berufsgruppe, Parteienpräferenz und Bundesland: Ergebnisse der Standardisierung des Österreichischen Wertfragebogens von <i>Walter Renner</i> und <i>Ingrid Salem</i> .....	89
Richtungseffekte bei Bildblattvorlagen im demoskopischen Face-to-face-Interview von <i>Erich Lamp</i> .....	113
Labour Migration, Market Competition and Ethnocentrism: Guest workers in Israel and Germany – Ergebnisse eines Forschungsprojektes und Beschreibung des Datensatzes (ZA-Studien-Nr. 3974) by <i>Julia Iser</i> and <i>Silvia Keil</i> .....	147

---

## Berichte aus dem Archiv

Archivierung und Sekundärnutzung qualitativer Interviewdaten – eine Machbarkeitsstudie von <i>Reiner Mauer</i> .....	152
---	-----

Dimensional Analysis 34th Spring Seminar at the Zentralarchiv 14 Feb. - 4 March, 2005 .....	154
ZHSF-Methodenseminar 2004 Forschungsmethoden, Datenbankmanagement und Statistik in der Historischen Sozialforschung .....	159
Daten und Datenbanken der Historischen Statistik am Zentralarchiv für Empirische Sozialforschung, Abteilung ZHSF – Bericht zum Workshop am 9. u. 10. Oktober 2003 in Köln von <i>Volker Müller-Benedict</i> .....	161
Neue Studien im ZHSF von <i>Gabriele Franzmann</i> .....	164
The Central and Eastern Eurobarometer 1990 - 1997: TRENDS CEEB 1 - 8 The cumulative Data set is now available on CD-ROM by <i>Brigitte Hausstein, Wolfgang Zenk-Möltgen</i> and <i>Meinhard Moschner</i> .....	171
Election Studies from Eastern Europe 1989 - 2002 New Release - Data and Documentation on CD-ROM .....	180
ZA-Direktor wird Mitglied des Center of Excellence in Nishinomiya .....	181

---

### **Berichte aus anderen Instituten**

Der European Social Survey (ESS) 2002/2003 von <i>Katja Neller</i> .....	182
The Establishment of the Social Science Data Archive in Bulgaria (SSDA) by <i>Yantsislav Yanakiev</i> and <i>Brigitte Hausstein</i> .....	189

---

### **Buchhinweise**

<i>Buttler, Günter; Fickel, Norman</i> : Statistik mit Stichproben .....	191
<i>Zöfel, Peter</i> : SPSS-Syntax: Die ideale Ergänzung für effiziente Datenanalyse .....	192

---

Bei Beiträgen, die nicht von Mitarbeitern des Zentralarchivs verfasst wurden, ist die Anschrift der Autoren beim jeweiligen Artikel angegeben. Die Inhalte der Beiträge entsprechen der Meinung der Autoren und geben nicht unbedingt die Ansicht der Redaktion wieder.

Alle inhaltlichen Beiträge sind Gegenstand einer Beurteilung durch externe Gutachter.

## Mitteilungen der Redaktion

Der langjährige Direktor des ZA, *Erwin K. Scheuch*, wurde in einer Akademischen Gedenkfeier am 23.4.04 in der Universität zu Köln gewürdigt. *Heiner Meulemann*, Nachfolger *Scheuchs* auf dessen Lehrstuhl, war der Organisator dieser Veranstaltung. *Franz Urban Pappi* und *Heinrich Best* erinnern an *Scheuchs* Bedeutung für die empirische Soziologie in Deutschland sowie als akademischer Lehrer, Vorgesetzter und Kollege.

Anknüpfend an Beobachtungen zur Links-Rechts-Asymmetrie in der Wahrnehmung optisch dargebotener Stimuli geht *Erich Lamp* der Frage nach, ob sich eine Vorzugsrichtung in der Wahrnehmung auch bei Bildblattvorlagen im Interview auswirkt. Fragebogenexperimente mit verschiedenen Skalenmodellen zeigen, dass die Befragten beim Betrachten einer Bildblatt-Skala offenbar einer "natürlichen Blickbahn" folgen, die ihr Antwortverhalten unbewusst beeinflusst.

*Harald Schoen* vergleicht zwei Formen von Internetbefragungen mit persönlichen Interviews. Die Befunde aus einer offenen WWW-Umfrage weichen erwartungsgemäß erheblich von denen einer Face-to-face-Befragung ab. Dagegen scheinen Online-Befragungen für Reliabilitätsanalysen bei der Skalen-Konstruktion sinnvoll zu sein.

*Walter Renner* und *Ingrid Salem* beschreiben die Entwicklungsschritte bei der Konzeption eines neuen österreichischen Werteinventars.

Das Verzerrungspotential durch soziale Erwünschtheit ist Gegenstand der Untersuchungen von *Volker Stocké* und *Christian Hunkler*. Die Autoren vergleichen die Gültigkeit von drei unterschiedlich differenzierten und daher verschieden aufwendigen Operationalisierungen für die Stärke und Richtung der von den Befragten wahrgenommenen Anreize zu sozial erwünschtem Antwortverhalten bei verschiedenen Befragungsthemen.

*Franz Bauske*

## **Begrüßung zur Akademischen Gedenkfeier der Universität zu Köln am 23.4.2004 für Erwin K. Scheuch (\*9.6.1928 †12.10.2003)**

**von Heiner Meulemann<sup>1</sup>**

... *Erwin K. Scheuch* war in vieler Hinsicht ein außergewöhnlicher Soziologe – zuerst als Vertreter des akademischen Fachs. Wohl kaum einer aus der Soziologengeneration, die in den sechziger und siebziger Jahren dem nach dem Krieg wieder etablierten Fach zur öffentlichen Anerkennung verholfen haben, hat das Fach so breit in Forschung und Lehre vertreten wie er. Das wird leicht vergessen, weil er vor allem als Förderer und Fortentwickler der empirischen Sozialforschung bekannt geworden ist. Aber ebenso hat er die zentralen theoretischen Fragen des Faches nach der Entwicklungsrichtung und dem inneren Zusammenhalt moderner, hoch differenzierter Gesellschaften behandelt. Und es gibt kaum einen Bereich des sozialen Lebens, dem er sich nicht in seiner soziologischen Arbeit gewidmet hat – von den großen Systemen Politik und Wirtschaft über die intermediären Instanzen von Verbänden und Vereinen bis zu der – wie er es nannte – „subinstitutionellen Ebene des Alltagslebens“ von Konsum und Freizeit.

*Erwin K. Scheuch* war aber nicht nur ein hervorragender Vertreter seines Fachs. Er hat – so der Titel seiner Autobiographie – „drei Leben als Soziologe“ geführt: nicht nur als Wissenschaftler, sondern auch als Anwender und Publizist. Als Anwender hat er seine sozialwissenschaftliche Expertise zur politischen Entscheidungshilfe eingesetzt; sein größter Erfolg hier war vielleicht die Verhinderung des Leverkusener „Lindwurms“, eines überdimensionierten Wohnhochhauses. Als Anwender hat er zudem die Verbindung der akademischen mit der privatwirtschaftlich geführten Sozialforschung gepflegt. Als Publizist war er sein Lebtag lang, vor allem aber zu Beginn und Ende seiner akademischen Laufbahn tätig. Er hat sich als Volontär bei *Werner Höfer* und als Kölner Lokalreporter das Geld für sein Studium verdient und dabei das Recherchieren gelernt. Und er hat in den letzten Jahren gemeinsam mit seiner Frau Verkrustungen in Politik und Wirtschaft recherchiert und sich nicht ge-

---

<sup>1</sup> Dr. *Heiner Meulemann* ist Professor am Institut für Angewandte Sozialforschung, Universität zu Köln, Greinstraße 2, 50939 Köln.

scheut, institutionelle Regelungen vorzuschlagen, die Anreize zum „Klüngel“ mindern könnten – von der Stärkung des Kartellamts bis zur Beschränkung der Mandatsdauer der Politiker.

Vielleicht sollte man diesen drei sozialwissenschaftlichen Leben ein viertes, wissenschaftspolitisches, hinzufügen: als Institutionenbauer. **Erwin K. Scheuch** hat aus seinen amerikanischen Erfahrungen die Gründung des Zentralarchivs für Empirische Sozialforschung 1962 unterstützt; er war von 1964 bis zu seiner Emeritierung 1993 Wissenschaftlicher Direktor des Zentralarchivs und danach Vorsitzender seines Trägervereins, der Kölner Gesellschaft für Sozialforschung. Er hat also vier Jahrzehnte lang diese erste Institution der sozialwissenschaftlichen Infrastruktur in Deutschland aufgebaut und von dieser Basis aus die beiden weiteren Institutionen der sozialwissenschaftlichen Infrastruktur, das Informationszentrum Sozialwissenschaften (IZ) in Bonn und das Zentrum für Umfragen, Methoden und Analyse (ZUMA) in Mannheim entscheidend mitgeprägt.

Auf der heutigen akademischen Gedenkveranstaltung soll das erste der sozialwissenschaftlichen Leben **Erwin K. Scheuchs** im Zentrum stehen, das des Fachkollegen. Zwei Kollegen, die seit Jahrzehnten mit ihm zusammengearbeitet haben, werden sein Werk wie seine Person aus unterschiedlichem Blickwinkel würdigen. Professor **Franz-Urban Pappi** von der Universität Mannheim war von 1967 bis 1974 Assistent bei **Scheuch** am Zentralarchiv für Empirische Sozialforschung, er hat sich 1977 an dieser Fakultät bei **Scheuch** habilitiert. Er hat mit ihm in der Forschung, insbesondere in der politischen Soziologie, eng kooperiert und wird ihn mit Blick auf sein akademisches Werk würdigen. Professor **Heinrich Best** von der Universität Jena war nach 1978 zwei Jahre Wissenschaftliche Hilfskraft unter **Scheuch**, nach 1979 acht Jahre Wissenschaftlicher Assistent bei **Scheuch** und nach 1987 sieben Jahre sowohl als Professor dieser Universität wie als Wissenschaftlicher Direktor des Informationszentrums Sozialwissenschaften Streiter für die sozialwissenschaftliche Infrastruktur mit **Scheuch**. Sein akademischer Werdegang ist also über lange Zeit mit **Scheuch** verbunden gewesen, und er wird ihn mit Blick auf die Person des Lehrers, Vorgesetzten und Kollegen würdigen.

Aber **Erwin K. Scheuch** war nicht nur Fachkollege, sondern prominentes Mitglied unserer Fakultät und hat mit seinen Arbeiten zur öffentlichen Anerkennung unserer Fakultät viel beigetragen. ...

## **Erwin K. Scheuch und seine Bedeutung für die empirische Soziologie in Deutschland**

**von Franz Urban Pappi<sup>1</sup>**

Ich beginne mit einem Zitat aus der Festschrift für *Erwin K. Scheuch*:

„Die Absicht, das Wirken *Erwin K. Scheuchs* für die Entwicklung der deutschen und der internationalen Soziologie nach 1945 mit einer Festschrift zu würdigen, stand vor dem Dilemma, in einem nahezu unübersehbaren Lebenswerk, das wie kaum ein anderes für sich beanspruchen kann, ‚allgemeine Soziologie‘ zu repräsentieren, einen verbindenden Grundakkord zu finden.“ So lautet der erste Satz der Herausgeber der Festschrift für *Erwin K. Scheuch* zum 65. Geburtstag. An diesen Satz erinnerte ich mich, als mir Herr Kollege *Meulemann* die ehrenvolle Aufgabe antrug, das Œuvre von *Scheuch* im Rahmen der heutigen akademischen Gedenkveranstaltung zu würdigen. Ich schlug spontan das engere Thema vor: *Erwin K. Scheuch* und seine Bedeutung für die empirische Soziologie in Deutschland. Ich spreche bewusst nicht von empirischer Sozialforschung, sondern von empirischer Soziologie und meine damit die erst nach 1945 entstandene empirische Soziologie, die auf der Basis von systematisch erhobenen Daten zu Aussagen über die gesellschaftliche Wirklichkeit gelangt. Jede Soziologie muss als Gesellschaftswissenschaft natürlich empirisch sein und ist es seit den Gründervätern im späten 19. Jahrhundert auch immer gewesen. Was die Entwicklung nach dem Zweiten Weltkrieg kennzeichnet, ist dagegen die Fundierung von Aussagen über die gesellschaftliche Wirklichkeit mit Ergebnissen der empirischen Sozialforschung.

In der Rolle des empirischen Soziologen und empirischen Sozialforschers habe ich *Erwin Scheuch* als Student und Assistent Anfang der 60er Jahre kennen und schätzen gelernt, und in dieser Rolle haben wir später als Kollegen in vielen Gremien zusammen gearbeitet. Besonders bewegt hat mich, als mir vor wenigen Wochen

---

1 Dr. *Franz Urban Pappi* ist Professor am Lehrstuhl für Politische Wissenschaft I Universität Mannheim, Seminargebäude A 5, 68131 Mannheim

Vortrag anlässlich der Akademischen Gedenkfeier zum Gedenken an Prof. Dr. *Erwin K. Scheuch*, am 23.4.04, in der Universität zu Köln

eine eigene Festschrift überreicht wurde, dass *Erwin Scheuch* darin, zusammen mit Ihnen, sehr verehrte Frau *Scheuch*, noch einen Beitrag zu dem gewählten Thema meiner Festschrift „Interdisziplinäre Sozialforschung“ verfasst hat.

*Erwin Scheuch* gehört zu der Generation deutscher Soziologen, die die bisherige Geschichte der Bundesrepublik Deutschland miterlebt und im Teilbereich der Sozialwissenschaft mitgestaltet hat: 1. die Aufbruchphase der empirischen Soziologie in den 50er Jahren, 2. die ersten Erfolge und Bewährungsproben in den 60er Jahren und 3. die Phase der Institutionalisierung der sozialwissenschaftlichen Infrastruktureinrichtungen seit den 70er Jahren. Ich werde die Bedeutung *Scheuchs* für die empirische Soziologie in Deutschland nacheinander für diese drei Phasen herausarbeiten, um in einem 4. Abschnitt das Verhältnis zu und die Nutzung von Ergebnissen der empirischen Sozialforschung durch *Scheuch* selbst in seinen soziologischen Arbeiten zu charakterisieren. 5. werde ich seine Auffassung von empirischer Soziologie skizzieren, die sich nicht auf empirische Sozialforschung reduzieren lässt.

## 1 Die Aufbruchphase der empirischen Soziologie nach dem Zweiten Weltkrieg

Schwerpunkte der Soziologie in den 50er Jahren waren neben Köln Frankfurt mit *Horkheimer* und *Adorno* und Hamburg mit *Schelsky*. Diese Schwerpunkte wurden auch als Schulen bezeichnet und die von *René König* begründete Kölner Schule unterschied sich von der Frankfurter und der *Schelsky*-Schule durch das Verständnis von Soziologie als einer empirischen Einzelwissenschaft, einer – um eine Formulierung von *René König* aufzugreifen – Soziologie, die „nichts als Soziologie“ sein will (vgl. *Lepsius* 1979: 36). Alle drei Schulen erkannten die zentrale Rolle der empirischen Sozialforschung, das war die gemeinsame Basis. Die Theoriekonzeptionen unterschieden sich aber. Auf der einen Seite die kritische Theorie in Frankfurt und auf der anderen Seite eine gewisse Theorielosigkeit im modernen Verständnis der *Schelsky*-Schule. Die Führung im Bereich der empirischen Sozialforschung hat aber in den 50er Jahren eindeutig Köln übernommen. Ich erinnere an die zwei Bände über das „Interview“ und „Beobachtung und Experiment“ in der Reihe Praktische Sozialforschung, die von *René König* herausgegeben wurden und an denen Dipl.-Volkswirt *Erwin Scheuch* und später 1956 Dr. rer. pol. *Erwin Scheuch* mitwirkte. Abgeschlossen wurde diese Phase mit der Herausgabe des ersten Bandes des Handbuchs der empirischen Sozialforschung, in dem *Scheuch* dann prominent mit mehreren zentralen Beiträgen wie z.B. über das Interview oder über Stichprobenverfahren vertreten war. Da sich inzwischen die Soziologie neben diesen Zentren auch an vielen anderen deutschen Universitäten etabliert hatte und man wegen der schulübergreifenden Einsicht in die Bedeutung der empirischen Sozialforschung

im Curriculum jeweils Methoden als eigene Lehrveranstaltungen vorsah, überrascht es nicht, dass die so genannte Kölner Schule, und mit ihr ganz zentral **Erwin Scheuch**, eine zentrale Bedeutung für die Methodenlehre bekam.

## 2 Erste Erfolge und Bewährungsproben in den 60er Jahren

Es folgte die Phase der ersten Erfolge und der Bewährungsproben in den 60er Jahren. Damit komme ich auf die ersten großen empirischen Projekte zu sprechen, die von **Scheuch**, in der Regel zusammen mit Kollegen, geleitet wurden. Hier ist einmal die Kölner Wahlstudie von 1961 zu erwähnen und dann die international vergleichende Zeitbudgetstudie. Mit letzterer musste sich **Scheuch** in einem Feld bewähren, das er für seine zukünftige Arbeit als Kölner Ordinarius seit 1965 als sein professionelles Spezialgebiet ausbaute: Die international vergleichende Sozialforschung. Zu ihr hat er sowohl wichtige methodische Beiträge geleistet als sich auch auf ihre Ergebnisse verlassen, wenn er bestimmte gesellschaftliche Probleme in einer einzelnen Gesellschaft, und hier natürlich insbesondere der deutschen Gesellschaft, analysierte. Die internationale Zeitbudgetstudie war, wenige Jahre nach der ebenfalls international vergleichenden Untersuchung der politischen Kultur durch **Almond** und **Verba**, deswegen bedeutend, weil man sich für den internationalen Vergleich nicht auf bloße Einstellungen verlassen wollte. Es ging vielmehr um die Verwendung der Zeit für bestimmte Kategorien von Alltagsaktivitäten wie Essen, Schlafen, Arbeiten am Arbeitsplatz, mit Freunden zusammen sein. Die Erfassung der für diese Tätigkeiten aufgewendeten Zeit erlaubt eine sichere Grundlage zur Beurteilung des Alltagsverhaltens unter den Bedingungen unterschiedlicher Gesellschaften. Da dies eine der wenigen Studien zur damaligen Zeit war, an denen auch osteuropäische Länder im damals noch real existierenden Sozialismus beteiligt waren, konnte man untersuchen, wie sich das Alltagsverhalten je nach Wirtschaftssystem unterschied. Typisch für die sozialistischen Länder waren eine stärkere Berufszentrierung der Alltagstätigkeiten und damit insgesamt eine geringere Trennung von Beruf und Privatleben als im Westen. Die Auswertung dieser komplexen Daten war mit den damaligen EDV-Möglichkeiten an deutschen Universitäten fast unmöglich. Hier machte es sich bezahlt, dass **Erwin Scheuch** nach seiner Habilitation in Köln einige Jahre als Lecturer in Social Relations an der Harvard University verbringen konnte.

Dieser Aufenthalt von **Scheuch** kam auch der Analyse der Umfragedaten zugute, die im Rahmen der Kölner Wahlstudie 1961 gesammelt worden waren. Zur Bedeutung dieser Studie kann ich **Scheuch** selbst zitieren: „Die Wahlstudie 1961 war der Versuch dreier junger Wissenschaftler (neben **Scheuch**, **Rudolf Wildenmann** und

*Gerhard Baumert*), eine empirische Untersuchung zu planen und durchzuführen, die technisch und intellektuell besten internationalen Maßstäben mindestens entsprechen sollte. Mit dieser Untersuchung sollten zumindest für die Bundesrepublik Standards gesetzt werden für die Sozialwissenschaft auf empirischer Grundlage“ (*Scheuch* 2000: 4). *Scheuch* selbst spricht hier von der Überschätzung als „Folge eines häufigen Ehrgeizes von Nachwuchswissenschaftlern“. Im Nachhinein betrachtet und verglichen mit damaligen amerikanischen Präsidentschaftswahlstudien und Wahlstudien in europäischen Ländern komme ich hier aber zu dem Schluss, dass die Ziele alle erreicht worden sind, wenn auch nicht sehr schnell nach der eigentlichen Datenerhebung.

Neben diesen großen empirischen Untersuchungen war eine weitere Herausforderung in den 60er Jahren die Politikberatung. Kurz nach *Scheuchs* Berufung an die Wirtschafts- und Sozialwissenschaftliche Fakultät der Universität zu Köln kam es in Bonn zum Regierungswechsel von der CDU/FDP-Koalition zur Großen Koalition CDU/CSU und SPD. Ein Ziel dieser neuen Regierung war die Änderung des Wahlrechts. Der damalige Innenminister *Lücke*, selbst ein prinzipientreuer Vertreter des relativen Mehrheitswahlrechts, berief im Februar 1967 einen wissenschaftlichen Beirat für Fragen der Wahlrechtsreform. Dieser aus sieben Wissenschaftlern bestehende Beirat legte im Dezember desselben Jahres einen umfangreichen Abschlussbericht vor. *Scheuch* gehörte zu den sechs Mitgliedern, die sich eindeutig für die Einführung des relativen Mehrheitswahlrechts aussprachen. Zusammen mit zwei weiteren Sachverständigen schlug er aber vor, das Mehrheitswahlrecht in Einerwahlkreisen entweder durch die Einführung von Vorwahlen bei der Kandidatenaufstellung oder durch eine Urwahl unter den Parteimitgliedern jedes Wahlkreises zu ergänzen (vgl. *Schönhoven* 2004: 270). Man erkennt unschwer einen weiteren Tenor in den späteren Arbeiten von *Scheuch*: Das Problem der Verhinderung oder wenigstens Kontrolle eines Elitenkartells in der Demokratie.

### **3 Phase der Institutionalisierung sozialwissenschaftlicher Infrastruktureinrichtungen**

Nach dieser Phase der Herausforderung und Bewährung folgte drittens die Phase der Institutionalisierung einer sozialwissenschaftlichen Infrastruktur in der Bundesrepublik Deutschland. Empirische Soziologie im modernen Verständnis ist auf systematisch erhobene Daten angewiesen. Neben prozessproduzierten Daten ist einer der wichtigsten Datenlieferanten die Umfrageforschung. Hier galt es, nach den Regeln der Kunst erhobene Umfragedaten nach der eigentlichen Primärauswertung für Sekundäranalysen verfügbar zu halten. Das ist die Gründungsidee des Zentralar-

chivs für Empirische Sozialforschung, zu dessen Mitdirektor **Scheuch** nach Übernahme des zweiten Soziologie-Lehrstuhls an der Universität Köln ernannt wurde. Der erste Direktor und Gründer des Zentralarchivs war der Finanzwissenschaftler **Günter Schmölders**, der lange vor der *Behavioral Economics* die verhaltenswissenschaftliche Orientierung in die Wirtschaftswissenschaft eingeführt hatte. Es folgten in relativ schneller Folge weitere Schritte auf dem Weg zur Institutionalisierung, wie die Gründung des Informationszentrums Sozialwissenschaften in Bonn, an dessen ursprünglicher Konzeption und Organisation als Einrichtung der Arbeitsgemeinschaft sozialwissenschaftlicher Institute **Scheuch** ebenfalls maßgeblich beteiligt war und des Zentrums für Umfragen, Methoden und Analysen, kurz ZUMA, 1974 in Mannheim. **Scheuch** gehörte dem Gründungsgremium und später dem ZUMA e.V. als Mitglied bis zu seinem Tode an.

Ein weiterer Institutionalisierungsschritt war die Einrichtung einer allgemeinen Bevölkerungsumfrage für die Sozialwissenschaften, heute als ALLBUS bekannt. Die erste ALLBUS-Erhebung fand 1980 statt und **Scheuch** war im ALLBUS-Beirat vertreten. Diese verschiedenen Institutionalisierungsschritte wurden in den 80er Jahren in der GESIS zusammengeführt, der Gesellschaft für Sozialwissenschaftliche Infrastruktur-Einrichtungen mit den drei Instituten Zentralarchiv in Köln, ZUMA in Mannheim und Informationszentrum Sozialwissenschaften in Bonn.

#### **4 Nutzung von Ergebnissen der empirischen Sozialforschung durch Scheuch selbst**

Wer **Scheuch** kannte, muss jetzt die entscheidende Frage stellen. Wie geht das zusammen, **Erwin K. Scheuch** und die Institutionalisierung? Damit komme ich zu meinem vierten Punkt, nämlich der Charakterisierung des besonderen Verhältnisses von **Scheuch** zu den jetzt von einer solide finanzierten sozialwissenschaftlichen Infrastruktureinrichtung zur Verfügung gestellten Ergebnissen der empirischen Sozialforschung. Hierzu ist zunächst hervorzuheben, dass **Scheuch** nie ausschließlich empirischer Sozialforscher gewesen ist. Er verstand sich immer als empirischer Soziologe und als solcher sah er es als seine Aufgabe an, theoriegeleitete Analysen gesellschaftlicher Probleme auf empirischer Basis als seinen Beitrag zur Wissenschaft und zum öffentlichen Diskurs durchzuführen.

Wir können also die Nutzung von Daten der empirischen Sozialforschung im engeren Sinn, also **Scheuch** als Sozialforscher, unterscheiden von seiner Auffassung von empirischer Soziologie in einem breiteren Sinn. Nur wenn wir letzteres mitberücksichtigen, werden wir **Scheuchs** Bedeutung für die empirische Soziologie in Deutschland gerecht.

Zunächst zum Verhältnis von **Scheuch** zur Sozialforschung im engeren Sinn. **Scheuch** blieb selbst an der Durchführung von großen sozialwissenschaftlichen Datenerhebungen wie insbesondere dem ALLBUS beteiligt und brachte sein umfangreiches Wissen über theoretische Bezüge und Operationalisierungen zentraler Konzepte in die Beratungen des ALLBUS-Ausschusses ein. Daneben blieb er an der Durchführung großer eigener Erhebungen interessiert. Hier erwähne ich als Beispiel die Begleituntersuchung zur Volkszählung 1987.

Er nutzte diese Datenbasis auch in seinen vielen Aufsätzen der letzten Jahre. So fand ich in dem **Scheuch**-Archiv, das am Zentralarchiv für Empirische Sozialforschung verwaltet wird, einen Aufsatz vom Ende der 90er Jahre mit dem Titel „Veränderung in den Beziehungen zwischen Amerika und Deutschland im Spiegel der Umfragen“. Hier untersucht **Scheuch** sehr sorgfältig das Deutschland-Bild der Amerikaner und das Amerika-Bild der Deutschen und zwar im zeitlichen Verlauf und im Vergleich der jeweiligen Einschätzung der Befragten in den beiden Ländern auch für andere Vergleichsgesellschaften. Gleich im ersten Absatz erwähnt er ein wichtiges Motiv, das wieder die Elitenkritik und die Kontrolle der Eliten durch die mit Umfragen erfasste öffentliche Meinung anklingen lässt. Ich zitiere: „Die Veröffentlichungen in den Tagesmedien ... sind selbstverständlich kein getreues Abbild der Befindlichkeit einer Bevölkerung und deren Führungsschichten. Umfragen sind hier verlässlicher. Und auch die in Führungsschichten verbreitete Vorstellung, ein Tenor in der veröffentlichten Meinung müsse schließlich zur öffentlichen Meinung selber werden, erweist sich in vielen Fällen als unzutreffend.“ Charakteristisch für **Scheuch** ist zum einen das Aufgreifen aktueller Probleme, die in der öffentlichen Diskussion eine Rolle spielen, aber gleichzeitig die Fundierung seiner Analyse mit Ergebnissen der empirischen Sozialforschung. Diese Fundierung erlaubte ihm dann auch in öffentlichen Diskussionen oft das eine oder andere Aperçu. Eines, das mich damals vor vielleicht 20 Jahren sehr beeindruckt hat, möchte ich hier wiederholen. In einem Diskussionsforum über die Gefahren der Atomenergie erinnere ich folgenden Satz von **Scheuch**: „Was die Leute wirklich aufregt, ist nicht die Atomenergie, sondern Hundedreck auf der Straße.“

In vielen Aufsätzen, die wohl ursprünglich als Vortragsmanuskripte entstanden sind, werden Ergebnisse der empirischen Sozialforschung als „attention step“ in der Einleitung verwendet, um dann das jeweilige Problem zugespitzt unter einer bestimmten Fragestellung angehen zu können. In anderen Aufsätzen sind Ergebnisse aus Umfragen das zentrale Beweisstück wie z.B. in dem 2002 erschienenen Aufsatz „Selbstaufgabe der deutschen Amtskirchen?“. Was diese Aufsätze kennzeichnet, ist ein konkreter Problembezug und die Verwendung vor allem von Umfragedaten entweder zur Darstellung der Problemlage oder zur entscheidenden Basis der Analyse.

Alle diese Aufsätze können aber die Frage nicht beantworten, warum *Scheuch* einen so wichtigen Beitrag zur Institutionalisierung der sozialwissenschaftlichen Infrastruktureinrichtungen geleistet hat. Hat er die Möglichkeiten, die ihm Zentralarchiv, ZUMA oder GESIS lieferten, ausreichend genutzt?

Diese Frage ist falsch gestellt. *Erwin Scheuch* hat Zeit und Energie in die Institutionalisierung sozialwissenschaftlicher Infrastruktureinrichtungen als Dienst an der empirischen Soziologie in Deutschland, d.h. in ein Kollektivgut der Profession investiert.

### 5 Scheuchs Auffassung von empirischer Soziologie

Was aber ist seine Auffassung von empirischer Soziologie? Sie lässt sich nicht auf die übliche Auffassung eingrenzen, wonach es Aufgabe der Empirie ist, aus Theorien abgeleitete Hypothesen zu testen. Daraus lernt man zwar, ob Theorien sich empirisch bewähren oder nicht. Aber man lernt nicht unbedingt Entscheidendes über die soziale Wirklichkeit einer bestimmten Gesellschaft zu einem bestimmten Zeitpunkt. *Scheuchs* Anliegen war die Gesellschaftsanalyse oder -interpretation. Seine Beiträge leben vom Spannungszustand zwischen bestimmten Gesellschaftsinterpretationen und unabhängig davon gewonnenen Daten oder anderen Quellen über diese Gesellschaft oder ihre Teilsysteme. Nimmt man den üblichen Maßstab von theoretisch abgeleiteten Hypothesen und ihrem Test an Daten über die soziale Wirklichkeit an, kann man zu der Auffassung kommen, *Scheuch* habe die von ihm selbst geschaffenen Daten-Archive als normaler Empiriker unzureichend genutzt. Aber er war kein normaler Empiriker, sondern empirischer Soziologe, der an Gesellschaftsanalysen interessiert war. Dies hatte mehrere Konsequenzen für sein wissenschaftliches Engagement. Ich greife die 5 nach meiner Auffassung wichtigsten heraus:

1. Er beschränkte seine Analysen nicht auf Probleme, für die Daten zur Verfügung standen oder leicht erhoben werden konnten. So blieb er bis zum Schluss an nationalen Eliten und Machtstrukturen interessiert, bezog vorhandene empirische Ergebnisse in seine Analysen ein, kam aber zu wichtigen Schlussfolgerungen z.B. über die „Wirtschaftseliten und Entscheidungsverläufe in der real existierenden Marktwirtschaft Deutschlands“ (*Scheuch* und *Scheuch* 2004) auch auf dem Wege freier Interpretation oder, um einen von seinem Kollegen *Dahrendorf* geprägten Begriff zu verwenden, auf dem Wege „informierter Willkür“. Dabei ist die stimmige Gesamtinterpretation der Plausibilitätstest der Analyse.
2. Ein häufiger Ausgangspunkt seiner Analysen waren die Interpretationen anderer, die er mit vorhandenen Daten konfrontierte. Den von ihm oft gesehenen Gegensatz von veröffentlichter Meinung und öffentlicher Meinung habe ich bereits er-

wähnt. Diese Vorgehensweise konnte er nur wählen, wenn tatsächlich ein unabhängiger Zugang zur jeweiligen sozialen Wirklichkeit mit Hilfe systematisch erhobener Daten der empirischen Sozialforschung möglich war. Bei Aussagen über die öffentliche Meinung stellen allgemeine Bevölkerungsumfragen eine solche Möglichkeit dar und die hat **Scheuch** ausgiebig genutzt. Er ist aber nie in den individualistischen Fehlschluss verfallen und hätte nun seinerseits aggregierte Individualdaten mit strukturellen und institutionellen Sachverhalten einer Gesellschaft verwechselt.

3. An den soziologischen Theorien interessierten ihn weniger die Theorien selbst, sondern, als Wissenssoziologe, die Annahmen dieser Theorien über die jeweilige gesellschaftliche Wirklichkeit. Diese Erkenntnis habe ich seinem erst im Juni 2003 erschienenen zweibändigen Werk über sozialen Wandel entnommen, dessen erster Band den großen soziologischen Theoretikern des sozialen Wandels gewidmet ist. Ihn interessieren diese Theorien nicht als reine Dogmengeschichte, sondern ihn interessiert die wissenssoziologische Deutung im Hinblick auf Veränderungen der realen Situationen. „Worauf in seiner Zeit hat der Autor reagiert, was erschien ihm zu diesem Zeitpunkt erklärungsbedürftig und was sollte von ihm gelernt werden“ (Band 1: 17). Einen knappen Überblick über diese Vorgehensweise bietet sein ebenfalls 2003 in der Zeitschrift für Ethnologie erschienener Aufsatz „Die Entwicklung zum Postmodernismus in der Sozialwissenschaft“.
4. Für die Rekonstruktion sozialer Wirklichkeit verlor **Scheuch** das Alltagsverhalten der Menschen nie aus dem Auge, oder, um es mit einem Wort seines Doktorvaters **René König** zu sagen, „das real gelebte Leben“. Er war ein Soziologe auch des Subinstitutionellen, dessen Vorstellung allerdings war, dass man institutionelle Gegebenheiten einer Gesellschaft aus den Sachverhalten der Mikroebene, wenigstens zum Teil, aufbauen könne. Ich zitiere als Beleg aus dem Band 2 des Sozialen Wandels: „Die Entwicklung in Deutschland zwischen 1944 und 1951 kann als Beispiel dafür dienen, dass nach dem Zerfall der Gesellschaftsordnung von der Mikroebene her schnell größere gesellschaftliche Zusammenhänge rekonstruierbar werden“ (Band 2: 382). Das Ergebnis ist dann eine Gesellschaft als „Geflecht von Widersprüchlichkeiten“, um einen weiteren Aufsatztitel von **Scheuch** zu zitieren (**Scheuch** 2000b).
5. Dieses Verständnis von empirischer Soziologie schließt das übliche Verständnis von Theorie und Empirie nicht aus. Deshalb kann man 5. sagen, dass **Scheuch** immer ein großer Förderer dieser empirischen Arbeiten im engeren Sinn gewesen ist. Paradigmatisch kann man hier seine Initiative und langjährige Leitung des Ausschusses zur Vergabe des Thyssen-Preises für den besten sozialwissenschaftlichen Zeitschriftenaufsatz in deutscher Sprache erwähnen. Für die moderne em-

pirische Sozialwissenschaft ist der Zeitschriftenaufsatz das wichtigste Publikationsforum und nicht mehr die Monographie. Diese Form der Publikation wollte **Scheuch** im deutschsprachigen Raum mit der Einrichtung des Thyssen-Preises stärken. Dies ist ihm gelungen. Der Preis wird seit 1981 vergeben und **Scheuch** war immer an diesen Entscheidungen beteiligt. Im Jahr 1991 ist eine Auswahl der bis dahin ausgezeichneten Aufsätze in französischer Übersetzung von **Scheuch** herausgegeben worden unter dem Titel "Perspectives des sciences sociales en Allemagne aujourd'hui". Es überrascht nicht, dass Arbeiten der empirischen Soziologie in diesem Buch überwiegen.

**Erwin K. Scheuch** hat keine soziologische Schule gegründet. Man kann im Gegenteil sagen, er hat uns eine Infrastruktur der Sozialwissenschaften hinterlassen, mit deren Hilfe man geschlossene Gedankengebilde in Frage stellen, modifizieren, aber gelegentlich, wenn auch nur in Teilbereichen, bestätigen kann. In diesem Sinn haben wir empirische Soziologen und empirische Sozialforscher **Erwin K. Scheuch** zu danken. Wir werden ihm stets ein ehrendes Andenken bewahren.

#### Literatur

**Lepsius, Rainer M.**, 1979 : Die Entwicklung der Soziologie nach dem Zweiten Weltkrieg 1945 bis 1967 . Kölner Zeitschrift für Soziologie und Sozialpsychologie (Sonderheft) 21: 25-70.

**Reuband, Karl-Heinz, Franz U. Pappi** und **Heinrich Best** (Hrsg.) 1995: Die deutsche Gesellschaft in vergleichender Perspektive : Festschrift für Erwin K. Scheuch zum 65. Geburtstag. Opladen : Westdeutscher Verlag.

**Scheuch, Erwin K.**, o.J.: Veränderung in den Beziehungen zwischen Amerika und Deutschland im Spiegel der Umfragen. Universität zu Köln (Unveröffentlichtes Manuskript).

**Scheuch, Erwin K.**, (Hrsg.), 1991: Perspectives des sciences sociales en Allemagne aujourd'hui. Paris: Éditions de la Maison des Sciences de l'Homme.

**Scheuch, Erwin K.**, 2000 a: Die Kölner Wahlstudie zur Bundestagswahl 1961. In: **Klein, Markus, Wolfgang Jagodzinski, Ekkehard Mochmann** und **Dieter Ohr**, (Hrsg.), 50 Jahre empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten. Wiesbaden: Westdeutscher Verlag, S. 41-58.

**Scheuch, Erwin K.**, 2000 b: Deutschland als moderne Gesellschaft: ein Geflecht von Widersprüchlichkeiten. In: **Niedermayer, Oskar** und **Bettina Westle**, (Hrsg.), Demokratie und Partizipation. Wiesbaden: Westdeutscher Verlag, S. 372-383.

**Scheuch, Erwin K.**, 2002: Selbstaufgabe der deutschen Amtskirchen? Religion - Staat - Gesellschaft 3: 165-177.

**Scheuch, Erwin K.**, 2003 a: Sozialer Wandel. Band 1: Theorien sozialen Wandels, Band 2: Gegenwartsgesellschaften im Prozess des Wandels. Wiesbaden: Westdeutscher Verlag.

**Scheuch, Erwin K.**, 2003 b: Die Entwicklung zum Postmodernismus in der Sozialwissenschaft. Zeitschrift für Ethnologie 128: 25-57.

**Scheuch, Erwin K.; Gräf, Lorenz; Kühnel, Steffen**, 1989: Volkszählung, Volkszählungsprotest und Bürgerverhalten: Ergebnisse der Begleituntersuchung zur Volkszählung 1987. Stuttgart: Metzler-Poeschel.

**Scheuch, Erwin K.; Scheuch, Ute**, 2004: Wirtschaftseliten und Entscheidungsverläufe in der real existierenden Marktwirtschaft Deutschlands. In: **Henning, Christian H. C. A.** und **Christian Melbeck**, (Hrsg.), Interdisziplinäre Sozialforschung. Theorie und empirische Anwendungen; Festschrift für **Franz Urban Pappi**. Frankfurt am Main: Campus, S. 115-128.

**Schönhoven, Klaus**, 2004: Das Scheitern der Wahlrechtsreform in der Ära der großen Koalition 1966-1969. In: **Henning, Christian H. C. A.** und **Christian Melbeck**, (Hrsg.), Interdisziplinäre Sozialforschung. Theorie und empirische Anwendungen; Festschrift für **Franz Urban Pappi**. Frankfurt am Main: Campus, S. 265-283.

## **Erwin K. Scheuch als akademischer Lehrer, Vorgesetzter und Kollege**

### **Ansprache zur Akademischen Gedenkfeier für Erwin K. Scheuch an der Universität zu Köln am 23. April 2004**

von **Heinrich Best**<sup>1</sup>

Mir wurde in dieser Gedenkveranstaltung die ehrenvolle, aber auch schwierige Aufgabe zuteil, über „*Erwin K. Scheuch* als akademischer Lehrer, Vorgesetzter und Kollege“ zu berichten. Schwierig ist sie nicht zuletzt deshalb, weil ich hier vor einem Kreis von Experten stehe, von denen viele an *Erwin Scheuch* in diesen Rollen Erinnerungen haben, die frischer als die meinen sind. Auch seine akademische Lehre hat er ja bis zum Ausbruch der Krankheit, die zum Tode führte, fortgesetzt. Die einzige Besonderheit, die mich zum Berichterstatter qualifiziert, ist der Umstand, dass ich *Erwin Scheuch* in allen drei Rollenbezügen zugeordnet war: zuerst seit dem Wintersemester 1967/68 als Student der Soziologie, dann als Assistent am Institut für Angewandte Sozialforschung und zuletzt bis 1992 als frisch habilitierter Professor für Soziologie an der Wirtschafts- und Sozialwissenschaftlichen Fakultät – insgesamt also ein Zeitraum von 25 Jahren, der nahezu mit *Erwin Scheuchs* Zeit als Ordinarius für Soziologie in Köln zusammenfällt. Da gäbe es Stoff genug für eine veritable Wissenschaftlerbiographie, die in weiten Teilen mit der Geschichte unseres Faches und in mehrfacher Hinsicht mit der intellektuellen Geschichte der Bundesrepublik in dieser Zeit zusammenfielen. Offenkundig ein zu großes Thema für die halbe Stunde, die mir in dieser Mittagsveranstaltung gegeben ist. Ich kann hier nicht mehr tun, als Ihnen einige signifikante Episoden aus diesem Erinnerungsausschnitt zu erzählen; unvermeidbar persönlich gefärbte Reminiszenzen, die vielleicht plausibel machen werden, warum *Erwin Scheuch* eine bleibende Spur in den Köpfen und Institutionen hinterlassen hat, mit denen er zu tun hatte.

Episode 1: *Scheuch* als akademischer Lehrer in bewegter Zeit. Meine erste Lehrveranstaltung besuchte ich bei *Scheuch* im „tollen Jahr“ 1968: „Gegenwartsgesell-

---

<sup>1</sup> Dr. *Heinrich Best* ist Professor für Soziologie an der Friedrich-Schiller-Universität Jena, Carl-Zeiss-Str. 2, 07740 Jena

schaften I: Deutsche Gesellschaft der Gegenwart“. Ein pietätvolles Zögern vor der Entsorgung hat mir meine damalige Vorlesungsmitschrift erhalten, die nicht so sehr wegen meiner zahlreichen Notizen, sondern vor allem wegen der Disposition des Dozenten bemerkenswert ist. Deutlich an den „Syllabus“ amerikanischer Universitäten angelehnt, an denen **Scheuch** ja bis 1964 gelehrt hatte, beginnt sie mit einem „Statement of purpose“, den „Zwecken“. Sie sind hier und heute noch mitteilenswert, weil sie in eindeutiger Weise die Spur markieren, von der ich zuvor gesprochen habe, die aber auch erkennen lassen, warum seine Lehre manchen unerträglich war:

„Gegenwartsgesellschaften I:

*Deutsche Gesellschaft der Gegenwart*

Zwecke: (1) *Vermittlung faktischen Wissens über die eigene Umwelt*

*Das Leben in einem Lande führt zu einem falschen Gefühl der Vertrautheit. Kompetenz in der Lösung der eigenen Probleme ist sicher nicht das Gleiche wie das Verständnis eines Systems. Gerade das Eingebundensein in ein System bedingt Selektivität der Erfahrungen und Engagement.*

(2) *Vermittlung einer Schauweise*

*Angestrebt wird als Voraussetzung für selbständiges soziologisches Arbeiten die Verfremdung gegenüber den Selbstverständlichkeiten der vertrauten Umwelt. Diese Distanzierung ist verbunden mit einer speziellen Perspektive: der Frage nach den Wirkungen von Institutionen und Sachverhalten statt nach ihrer Bewertung. Die hier vertretene Perspektive wird außerhalb Kölns als „stark amerikanisch beeinflusste funktionalistisch-empirische Soziologie“ bezeichnet, dies ist übrigens nur teilweise richtig.*

(3) *Einführung in die Begriffssprache der Soziologie*

*Dies ist zugleich als Einführung in die Soziologie im Sinne einer akademischen Disziplin herkömmlicher Art gemeint: einer Disziplin mit formal abgegrenzten Objektbereichen, Begriffen und Lehrsätzen. Objektbereich und Lehrsätze sind ziemlich umstritten. Eher ist noch eine Einigung (bzw. Übersicht) über Begriffssprache möglich. In dieser Einführung soll bewusst auf Ideengeschichte verzichtet werden, stattdessen sollen die Begriffe in der Anwendung erläutert werden“.*

Was hier erkennbar hochgestimmt und selbstgewiss als Lehrzweck mitgeteilt wird, ist eine empirische Soziologie, orientiert an der **Weberschen** Maxime der Wertur-

teilsfreiheit und dadurch auf Aufklärung zielend. Bezeichnend ist die Opposition, die *Scheuch* zwischen den intellektuellen Operatoren „Verständnis“ und „Bewertung“ herstellt. Auch geht es ihm nicht nur darum, soziologisch relevante Sachverhalte zu vermitteln und auf diesem Weg eine Einführung in die Soziologie zu geben, sondern ganz offenkundig auch darum, Soziologen zu erzeugen.

Die Gliederung der Vorlesung gibt zur Vermutung Anlass, dass ich mich im Herbst 1968 an dem Ort in Deutschland befand, an dem am kompetentesten und fachlich modernsten über den damaligen Zustand unserer Gesellschaft Auskunft gegeben wurde. Sie antizipiert die wesentlichen Facetten und Themen der *Scheuchschen* Soziologie kommender Jahrzehnte. Tatsächlich wirkt sie mit Hauptgliederungspunkten wie „Zur Soziologie der Freizeit“, „Führungsgruppen“, „Massenkommunikationssystem“ oder „Kultur als soziale Institution“ wie die Zusammenstellung der inhaltlichen Schwerpunkte seines Forschungsprogramms. Auch die systematischen Elemente der *Scheuchschen* Gesellschaftsanalyse sind bereits präsent, etwa in den Gliederungspunkten „’Linkeage’ – Elemente und mediatisierende Instanzen“ oder „Moderne und traditionale Formen der Organisation von Spannungslinien (cross-cutting cleavages)“ oder „Autonomie und Schwellenwerte als Kriterium eines differenzierten Systems“ – die letzten beiden aber nur „falls Zeit gewonnen würde“, was jedoch leider nicht gelang.

Neben diesen verblüffenden Vorwegnahmen künftiger Entwicklungen des Faches, die im Sommer 1968 konzipiert und ausgearbeitet worden sein müssen, gibt es aber auch der damals aktuellen Harvardschen Theorieproduktion geschuldete Themen wie „funktionalistische Betrachtungen“ des Rechts, der Kultur oder von Religion und Kirchen, die heute bestenfalls einen nostalgischen Appeal haben. Überrascht hat mich der umstandslose – wenn auch durch Anführungszeichen etwas gemilderte – Rückgriff auf die dem *Marx-Hegelschen* Kategorienfundus entstammende Unterbau-Überbau-Metapher als ein Hauptgliederungselement, um die Struktur- und Wertsphären der Gesellschaft und die mit diesen assoziierten Institutionen voneinander abzugrenzen. Völlig ausgeblendet blieb übrigens die DDR oder SBZ, wie es damals noch überwiegend hieß. Dies war, wie ich heute vermute, nicht linksrheinischer Idiosynkrasie geschuldet, sondern eher der *Scheuchschen* Maxime, dass man über das schweigen soll, worüber man nicht empiriegestützt reden kann.

Die Vorlesung selbst verband die inspirierte Exegese empirischen Materials mit verallgemeinernden Aussagen, die gerne das Kontraintuitive hervorkehrten und auch Provokantes nicht umschifften. „Apologetisch“ im Sinne einer Unterschlagung oder Beschönigung gesellschaftlicher Missstände war sie nicht. So wurde etwa unter dem Gliederungspunkt „Einkommens- und Besitzverhältnisse“ ausgiebig über

Ungleichheitsstrukturen berichtet, und ich kann meinen Aufzeichnungen den Sachverhalt entnehmen, dass in Deutschland Mitte der 1960er Jahre rund 2% der Steuerpflichtigen rund 40% der Vermögenswerte besaßen, womit aber Vermögen in Deutschland breiter gestreut waren als in England und Frankreich. Das war zwar keine „Kritische Theorie“, aber doch kritische Empirie, genauer: eine zur Gesellschaftskritik verwendbare Empirie. An dieser Weise des Realitätszugangs änderten auch die Ereignisse nichts, über die ich gleich berichten werde. In seiner Vorlesung über die USA, die ich im Sommersemester 1970 besuchte, wurde etwa ohne jede Beschönigung über die Deprivation und Diskriminierung der afroamerikanischen Bevölkerung und die Konstruktionen zur ideologischen „Bewältigung“ dieses Umstands berichtet. Einzig die seinerzeit gebräuchliche und damals völlig unkontroverse Verwendung der ethnischen Kategorie „Neger“ fällt aus dem Rahmen heutiger Standards des Umgangs mit diesem Thema.

Dies alles lässt sich problemlos mit der Tatsache in Übereinstimmung bringen, dass **Scheuch** damals in der „Neuen Gesellschaft“ – dem theoretischen Organ des Deutschen Gewerkschaftsbundes – publizierte und passt auch gut zu den Umständen unserer allerersten Begegnung, die bereits im Herbst 1967 außerhalb des akademischen Territoriums stattgefunden hatte. Auf meinen Erkundungszügen durch das nächtliche Köln war ich eines abends an einen eher abseitigen Ort geraten, den „Republikanischen Club“, in dem sich bei Schmalzbrot und Kölsch eine gemischte Gesellschaft von Kritikern der bestehenden sozialen und politischen Verhältnisse einfand, die von Linksliberalen und sozialdemokratischen Intellektuellen dominiert wurde, aber auch einige salonkommunistische Einsprengsel aufwies. Wenige Monate später sollte sich diese Versammlungsstätte rheinischer Nonkonformisten dann unversehens in eine der Schaltstellen der Kölner „außerparlamentarischen Opposition“ verwandeln. An besagtem Abend erschien zu später Stunde ein Paar, dessen Auftritt ein Raunen auslöste. Mir wurde bedeutet, dass der jugendlich wirkende Herr mit schwarzrandiger Brille Professor **Scheuch** sei, der einzige Hochschullehrer der Universität Köln, der den RC aufsuche. Zweierlei scheint mir an dieser kleinen Szene, an der sonst weiter nichts berichtenswert ist, bemerkenswert zu sein: zum einen zeigt sie, dass wir uns in einer Zeit befanden, in der der Auftritt eines Universitätsprofessors an einem öffentlichen Ort noch Raunen auslösen konnte, zum anderen zeigte sich **Erwin Scheuch** auf einer Position im politischen Raum, die sich polar zu jener verhielt, die ihm spätestens ein Jahr später zugewiesen wurde<sup>2</sup>.

---

2 „Diese Episode erscheint mir inzwischen in einem neuen Licht, nachdem ich aus einem Interview mit **Heinz Sahner**, das in der soeben erschienenen Festschrift zu seinem 65. Geburtstag erschienen ist, erfuhr, dass **Erwin Scheuch** sogar (Mit)gründer des Kölner Republikanischen Clubs war. Vgl. „**Heinz Sahner** im Interview“, in: **Michael Bayer** u. **Sören Petermann** (Hrsg.),

„Umstritten“ waren beide Positionen, wie fast alle seine Äußerungen als öffentliche Person, doch keine der Bewegungen im politischen Koordinatenfeld hatte für ihn so massive und auch unmittelbar in seine akademische Existenz einwirkende Folgen, wie jene, die sich zwischen 1967 und 1968 vollzog.

Es war damals schwer zu erkennen, ob es sich um einen Positionswechsel **Scheuchs** oder um eine Verschiebung des politischen Koordinatenfeldes insgesamt handelte, die einen libertären Wissenschaftler vom Schlage der amerikanischen „Liberals“ zu einem der Hauptantagonisten der „Neuen Linken“ werden ließ. Doch schon zu Beginn des Wintersemesters 1968/69 muss der Konflikt auch in seiner akademischen Lehre virulent geworden sein, worauf nicht zuletzt **Scheuchs** Vorbemerkung zu seiner Vorlesungsdisposition und hier vor allem sein Hinweis auf die Wahrnehmung der Kölner Soziologie als „amerikanisch beeinflusst“ und „funktionalistisch-empirisch“ hindeutet. Vorausgegangen waren seine Rede auf dem Neumarkt zum Tod **Benno Ohnesorgs**, sein Auftritt auf dem Frankfurter Soziologentag, wo er sich als ein Kritiker der Soziologie der Gastgeber exponiert hatte, und die Veröffentlichung der „Wiedertäufer der Wohlstandsgesellschaft“ – eine Buch, das nach seiner eigenen Mitteilung die „aufklärerische Funktion der Sozialwissenschaften“ gegen eine Bewegung durchsetzen wollte, „die sich selbst als Aufklärungsbewegung mit endzeitlichen Ambitionen“ versteht. Alles das war übrigens an mir vorbeigegangen, der ich gerade stärker damit beschäftigt war, andere als die in engerem Sinne soziologischen Aspekte der damaligen Umbruchsituation wahrzunehmen. Doch war das, was **Scheuch** die „68er Kulturrevolution“ nannte, bald auf unüberhörbare Weise in seinem Hörsaal präsent.

Mein Vorlesungsmanuskript zielt auf einem Blatt, das auf den Anfang des Jahres 1969 zu datieren ist, eine etwas ungelente Zeichnung, deren Mittelpunkt die Darstellung eines Kommissars aus den Tagen des russischen Bürgerkriegs bildet. Umgeben von Rotarmisten und roten Matrosen steht er mit anklagend erhobenem Arm an einem Katheder. Ich bin geneigt, diese Szene als verfremdend kritische Darstellung dessen zu interpretieren, was sich damals vor meinen Augen abspielte: der Versuch, **Scheuchs** Veranstaltungen – und hier insbesondere seine Vorlesungen – als Foren für die Gesellschaftsveränderung und Fundamentalkritik am so genannten Spätkapitalismus zu instrumentieren. Dies sollte durch die eventhafte Erzwingung von Vorlesungsunterbrechungen und ihr „Umfunktionieren“ in ein „Tribunal“ geschehen.

Wenn diese Interventionen eine Wirkung hatten, dann allerdings nicht die intendierte, das Herrschende System in der Gestalt eines seiner prominenten intellektuellen Apologeten zu entlarven und zugleich die studentischen Massen zu mobilisieren. Die saßen im keineswegs „kochenden“ Hörsaal wie in einer Arena und beobachteten aus Halbdistanz das Geschehen, bei dem die Vereitelung der Vorlesung durch „Umfunktionieren“ regelmäßig mit der Vereitelung der Umfunktionierung durch Vorlesung endete. **Scheuch** war – was die Verteidigung seines unverkürzten Rechtes zur akademischen Lehre anging – völlig unnachgiebig und kompromisslos. Eine intellektuelle Herausforderung waren diese Angriffe für ihn nicht, wohl aber ein Anschlag auf die Freiheit seiner Lehre, den er mit entschlossener Behauptung des Katheders, darüber hinaus aber mehr und mehr mit publizistischen und juristischen Mitteln abwehrte. In seinem Porzer Arbeitszimmer hängt hinter Glas ein Bericht des Kölner Stadtanzeigers aus dem Jahr 1977, in dem über den Prozess gegen einen der Störer berichtet wird. Er zeigt, dass sich dieser Konflikt bis tief in die 1970er Jahre fortgesetzt haben muss, als in anderen Bezirken der Universität schon längst wieder Ruhe eingekehrt war. Warum er dies auf sich nahm, wo es doch relativ einfach war, sich durch Konzessionen und Wegtauchen aus der Schusslinie zu bringen, blieb vielen unverständlich. Doch wird man ihm heute in der Deutung folgen können, dass seine *Unerschütterlichkeit* in der Überzeugung begründet war, die offene Wissenschaft gegen den Ansturm von Ideologen, gegen die „Zerstörung der Möglichkeit eines Austausches von Argumenten“<sup>3</sup> – wie er später schrieb – verteidigen zu müssen. Sein *Zorn* über die „angeblichen Rebellen“ hatte aber – denke ich – noch andere, biographische Gründe. **Scheuch** war, anders als die Angehörigen der älteren und die meisten seiner eigenen Soziologengeneration, weder patrizisch-großbürgerlicher noch bildungsbürgerlicher Herkunft. Was er war und schon in jungen Jahren erreicht hatte, hatte er ausschließlich sich selbst, seiner überlegenen Intelligenz und einer immensen Arbeitsfähigkeit zu verdanken. Das schlechte Gewissen und die Selbstzweifel vieler durch Herkunft Privilegierter waren ihm fremd. Dass ihm nun sein wohl erworbenes Recht – die unbeschränkte Lehrbefugnis des Hochschullehrers – von Kindern der Wohlstandsgesellschaft streitig gemacht wurde, die durch Herkunft und Generationslage doppelt bevorteilt waren, war ihm wohl unerträglich.

Über nahezu zehn Jahre hinweg wurde **Scheuchs** Lehre mit schwankenden Intensitätsgraden durch Störungen beeinträchtigt – eine Situation, die ihn nach meiner Ein-

---

3 Dieses und die folgenden Zitate sind – soweit nicht anders im Text vermerkt – entnommen aus **Erwin K. Scheuch**, *Wissenschaft-Anwendung-Publizistik: Drei Leben als Sozialwissenschaftler*, in: **Karl Martin Bolte** u. **Friedhelm Neidhardt**, *Soziologie als Beruf. Erinnerungen westdeutscher Hochschulprofessoren der Nachkriegsgeneration* (Soziale Welt, Sonderband 11), NOMOS: Baden-Baden 1998.

schätzung stärker belastet hat, als es seine etwas abgeklärten biographischen Selbstzeugnisse erkennen lassen, in denen übrigens jeder direkte Hinweis auf diese Ereignisse fehlt. Am Ende hat jedoch seine Lehre ihre „Zwecke“ erfüllt, und dazu beigetragen, Soziologie als eine erfahrungswissenschaftliche akademische Disziplin in Deutschland zu etablieren. Die Verbreitung des Kölner soziologischen Curriculums mit der zum Standard gewordenen Verankerung eines gewichtigen Methodenmoduls im Grundstudium, war dazu ein wichtiges Vehikel. Nicht zuletzt hat er Schüler gewonnen, darunter eine Reihe von Proselyten aus den historischen Disziplinen, die nicht trotz, sondern wegen seines strikten Empirismus die Fahnen wechselten.

Episode II: *Scheuch* als Coach und Teamplayer.

Meine Rekrutierung als *Scheuchs* Mitarbeiter erfolgte Ende 1977 im Anschluss an mein soziologisches Rigorosum, in dem ich von ihm umfassend über seine Sicht auf die Entwicklung der europäischen Familie aufgeklärt worden war. Der Umstand, dass ich im Fach Neuere Geschichte promoviert hatte, war für ihn dabei kein Hinderungs-, sondern ein Einstellungsgrund. Bereits Mitte der 1970er Jahre hatte er damit begonnen, die historische Sozialforschung in der Gestalt einer Gruppe fachlich hybrider Postgraduates zu fördern und ihr am Institut für Angewandte Sozialforschung Heimstatt zu gewähren. Sein Motiv war zum einen ein genuines Interesse an historischen Sachverhalten, vor allem aber die wissenschaftssystematische Erkenntnis, dass – wenn die Soziologie Wissenschaft des sozialen Wandels sein will – sie nicht ausschließlich Kostgängerin der Historie bleiben darf, sondern auch selbst historische Evidenz gewinnen muss, die ihren eigenen Fragestellungen und Ansprüchen an wissenschaftliche Beweismittel genügt. Meine Aufgabe war es zunächst, als Subunternehmer an diesem wissenschaftlichen Großprojekt *Scheuchs* mitzuwirken, das ja erst unmittelbar vor seinem Tod, mit der Veröffentlichung der „Theorien des sozialen Wandels“ abgeschlossen wurde.

Dabei trat die beauftragte Person hinter die Aufgabe zurück: recht lange dauerte es, bis sich *Scheuch* ohne zu zögern meines Namens erinnerte, der doch recht leicht zu merken ist. Ich hatte allerdings keinen Anlass, dies persönlich zu nehmen, denn seine Momente des Zögerns waren das unschuldige Symptom einer extrem gesteigerten Überlastung *Scheuchs*, der damals und noch auf lange Zeit mehrere parallele Leben führte, die nur in seinem Sekretariat kunstvoll verknotet wurden. Seinen Mitarbeitern eröffnete diese Situation weite Handlungsspielräume und Zugänge zu Handlungsfeldern, die Assistenten sonst verschlossen sind. Das galt auch für die Sphäre akademischen Institutionenbaus, in der *Scheuch* von der Mitte der 1960er Jahre bis Ende der 1980er Jahre stark engagiert war.

Wenige Monate nach meinem Dienstantritt erreichte mich zu später Stunde ein Telefonanruf, der mich an den Schauplatz einer Szene rief, die wohl den Gründungsakt der GESIS markiert, wobei ich nicht von ihrer ideellen Existenz rede – entsprechende Überlegungen hatte es schon zuvor gegeben –, sondern von dem Beginn des neunjährigen Implementationsprozesses, der ihrer materiellen Existenz vorausging. Präsent waren die „Großen“ der Kölner Soziologie in der Gestalt von **Renate Mayntz**, **Friedhelm Neidhardt**, **Peter Flora** und **Erwin Scheuch** sowie ihr Major-domus **Ekkehard Mochmann**. Verhandelt wurde der Weg, wie in Deutschland ein arbeitsteiliges Servicesystem zur Datenversorgung etabliert werden könne. Vermutlich war gerade von Frau **Mayntz** – die damals Vorsitzende der „Senatskommission für die empirische Sozialforschung“ war – der Vorschlag gekommen, die fälligen Planungsarbeiten im Kontext der DFG zu koordinieren und zu finanzieren. Ich wurde zum Koordinator der Planungsgruppe bestimmt – ein Angebot, das offenkundig nicht abgelehnt werden konnte. Ich erlebte dann **Scheuch** in den folgenden Jahren in einer Rolle, die ihm gewöhnlich nicht zugeschrieben wird: als teamplayer, wenn auch als ein Mitspieler ungewöhnlicher Art, der während des matches gelegentlich das Spielfeld verließ, häufig erst nach dem Anpfiff auflief oder vor dem Abpfiff verschwand, der aber doch, wenn er präsent war, das engmaschige Zuspiel und auch den Doppelpass verstand; ein Spielzug, der in Köln später ja theoretische Weihe erhielt. Die Mannschaft, die damals antrat, markiert – nicht zuletzt, was ihre institutionelle Durchsetzungsmacht angeht – eine Sternstunde der Kölner Soziologie. Über ein verteiltes System von interlocking directorates and mandates verband sie die sozialwissenschaftlichen Infrastruktureinrichtungen, die DFG und später den Wissenschaftsrat miteinander und besaß zudem einen privilegierten Zugang zu den einschlägigen Bundesministerien. Die Möglichkeiten dieser Konstellation konnten aber nur im abgestimmten Zusammenwirken der Hauptakteure nutzbar werden, das trotz mancher Spannungen und gelegentlich heftigen Entladungen auch in den entscheidenden Momenten zustande kam. Bemerkenswert ist, dass **Scheuch** bei einer völlig unbeirrbar und konfliktbereiten Ablehnung jeglicher „Finalisierung“ der Sozialwissenschaften oder ihrer Verwertung für eine sozialökonomische „Globalsteuerung“, wie sie in den Jahren der sozialliberalen Koalition en vogue war, durchaus mit den ministerialen Exekutoren dieser Politik im Gespräch blieb und von diesen auch respektiert wurde.

Die Zusammenarbeit mit ihm war zwar asymmetrisch, aber unhierarchisch. Er war offen für Ratschläge und nutzte die Kompetenz anderer. Letzteres jedoch nur im freien intellektuellen Austausch: wissenschaftlich habe ich keine Zeile verfasst, die unter seinem Namen erschienen wäre, wohl aber habe ich unzählige Briefe für ihn geschrieben. Im Windschatten seiner Aktivitäten ließen sich dann durchaus wissen-

schaftliche und institutionenpolitische Eigeninteressen verfolgen, was er förderte, wenn sich dies mit seinen Absichten verband und der höheren Ehre der Kölner Soziologie dienlich war. Das größte Problem in den alltäglichen Arbeitsbezügen war – neben seinen schwer antizipierbaren Eruptionen – die Schwierigkeit, ihn für einen kritischen Moment an einen bestimmten Ort zu fixieren – was sich mit einem reichen Anekdotenschatz belegen ließe. Seine Ansprüche an die Verfügbarkeit und Performanz seiner Mitarbeiter waren hoch, mit Lob wurde eher gespart. Er selbst besaß ein starkes Leistungsethos und war strikt auf Berufserfolg orientiert – insofern war er eine wandelnde Widerlegung der *Weberschen* Protestantismusthese. Dies verband sich jedoch erstaunlich widerspruchsarm mit rheinischer Soziabilität, Bonhomie, unpräntiösen Umgangsformen und Lebenszugewandtheit. Sein Netzwerk privater und beruflicher Bezüge war kaum zu überblicken und füllte mühelos die Dampfer der Rheinflotte. Großzügig ließ er auch seine Mitarbeiter an diesem sozialen Kapital teilhaben, das nicht zuletzt in seiner internationalen Dimension außerordentlich wertvoll war. Mit seiner Patronage bei der Platzierung im akademischen Stellenmarkt war dagegen eher nicht zu rechnen; hier war er wohl untergründig auf die durch eigene biographische Erfahrung gestützte Maxime orientiert, dass sich Qualität durchsetzen werde.

Meine Assistentenjahre gingen 1987 zu Ende, übrigens zeitgleich mit der Gründung der GESIS, deren Austragungszeit sich durch Regierungskrisen, Ministerwechsel, Bund-Länder-Querelen und allerlei Armdrücken zwischen den Instituten ungebührlich verlängert hatte. Schließlich kam der Erfolg, der bekanntlich viele Väter hat, aber doch wesentlich *Scheuchs* langjährigem und weit vorausschauendem Bestreben zu verdanken war, die sozialwissenschaftliche Infrastruktur als ein empirisches Sinnesorgan zu entwickeln, das seiner Weise des Wirklichkeitszugangs entsprach. Dem Umstand, dass dazu auch eine historische Tiefendimension gehörte, verdankt das Zentrum für Historische Sozialforschung wesentlich seine Aufnahme in die institutionelle Förderung.

### Episode III: Wertbezüge und Symbolbedeutung

Ich nähere mich dem Schluss und möchte hier unter dem Stichwort „*Scheuch* als Kollege“ einige Beobachtungen zu seinem Verhältnis zu den akademischen Institutionen einfügen. Die wichtigsten institutionellen Spielfelder *Scheuchs* waren die Profession – dies vor allem in ihren internationalen Bezügen – und die deutsche Öffentlichkeit, während seine Fakultät hier doch etwas zurücktrat. Den informellen Zirkeln, die seinerzeit ihre Geschicke bestimmten, gehörte er nicht an, zum Dekan wurde er nie gewählt und hat dieses Amt meines Wissens auch nicht angestrebt. Und doch war die Hochschule für ihn ein mit besonderer Würde ausgestatteter Ort,

auch hatte er ein emphatisches Verhältnis zum Amt des Hochschullehrers, das durchaus den Anspruch auf eine besondere „ständische Ehre“ enthielt. Mir ist dieses Pietätsverhältnis lange Zeit entgangen, obwohl seine entschiedene Verteidigung jenes symbolischen Ortes, der in Österreich bezeichnenderweise „Lehrkanzel“ heißt, bereits deutliche Hinweise darauf enthalten hatte. Welche starken Wertbezüge sein Verhältnis zu den akademischen Institutionen hatte, und wie markant die Symbolbedeutung war, die er dem akademischen Zeremoniell zuwies, wurde mir jedoch erst wenige Monate vor seinem Tode auf anrührende Weise verdeutlicht. Dies geschah, als wir über den Verlauf des Festaktes zur Verleihung der Ehrendoktorwürde an ihn sprachen. Dabei fragte er mich, ob im Jenaer akademischen Zeremoniell der Talar gebräuchlich sei. Ich bejahte – darauf **Scheuch**, offenkundig erfreut: dies sei für ihn kein Problem, er habe noch den alten der Kölner WiSo-Fakultät im Schrank. Hier schließt sich ein Kreis, denn als ich in Köln zu studieren begann, war die akademische Amtstracht dort bereits außer Gebrauch. Über mehr als 35 Jahre und mehrere Umzüge hinweg hat **Erwin Scheuch** also dieses ehrwürdige Relikt vergangenen Zeremoniells aufbewahrt; ein vielsagender Sachverhalt, der andeutet, wie wichtig die akademische Welt für ihn als Quelle von Lebenssinn und Wertgefühl war. Was umgekehrt **Erwin Scheuch** der Hochschule und speziell der akademisch verfassten Sozialwissenschaft zu geben hatte, habe ich versucht, in meinen episodischen Reminiszenzen anzudeuten.

## **Online-Umfragen – schnell, billig, aber auch valide? Ein Vergleich zweier Internetbefragungen mit persönlichen Interviews zur Bundestagswahl 2002**

**von Harald Schoen<sup>1</sup>**

### ***Zusammenfassung***

*Der Beitrag vergleicht die Ergebnisse einer persönlichen Befragung einer Zufallsstichprobe, einer Online-Befragung zufällig ausgewählter Internetnutzer sowie einer offenen WWW-Umfrage zu politischen Einstellungen, die im Umfeld der Bundestagswahl 2002 durchgeführt wurden. Aus Unterschieden in der Rekrutierung der Respondenten und im Befragungsmodus werden Hypothesen abgeleitet und empirisch überprüft. Die Randverteilungen der Online-Befragungen, vor allem aber der offenen WWW-Umfrage, weichen deutlich von jenen in der mündlichen Befragung ab: Männer, junge, hochgebildete, politisch interessierte und informierte Menschen sind darin deutlich überrepräsentiert. Sieht man von Wissensfragen ab, lassen sich zwischen den drei Umfragen keine gravierenden Unterschiede im Messfehler nachweisen. In Bezug auf Reliabilitätsschätzungen treten nur geringe Differenzen zwischen den drei Erhebungen auf. Dagegen resultieren zu Korrelationen zwischen Variablen, die mit der Rekrutierungswahrscheinlichkeit in Online- und Offline-Erhebungen zusammenhängen, in den Online-Erhebungen, vor allem aber der WWW-Umfrage deutlich andere Befunde als in der mündlich befragten Zufallsstichprobe. Gleiches gilt für Korrelationen zwischen parteipolitischen Einstellungen und für die Erklärungskraft politischer Einstellungen in Bezug auf das Wahlverhalten. Online-Erhebungen, insbesondere aber offene Web-Umfragen sind somit weitgehend ungeeignet, valide Aussagen über Randverteilungen und Variablenzusammenhänge in der deutschen Gesamtbevölkerung zu treffen.*

---

<sup>1</sup> Dr. **Harald Schoen** ist Wiss. Mitarbeiter am Institut für Politikwissenschaft der Johannes Gutenberg-Universität Mainz, Saarstraße 21, 55099 Mainz, Email: harald.schoen@politik.uni-mainz.de

## **Abstract**

*Online surveys – fast, cheap, but also valid? A comparison of two internet surveys with personal interviews on the German federal election 2002*

*This article analyses whether internet surveys on political attitudes allow for valid inferences on the electorate as a whole. For this purpose difference hypotheses on marginal distributions and correlations are proposed. They are tested empirically by comparing results of personally interviewed members of a probability sample, an online survey among members of an online panel and results of a survey among voluntary participants in an internet survey. As far as univariate statistics are concerned, the analysis reveals systematic biases in online surveys: Men, young, highly educated and politically interested persons participate in online surveys at a disproportionately high rate. Reliability estimates in online surveys do not diverge from results of personal interview data, but correlations among variables which are correlated with the selection probabilities in online surveys differ clearly from offline survey results; additionally, online estimates of the correlations among political attitudes and the explanatory power of political attitudes on voting behaviour differ from offline results; and this is especially true for the www-survey. Hence, the results show that at the moment online surveys – and especially www-surveys – are of very limited value for the analysis of the electorate as a whole in Germany.*

## **1 Einleitung<sup>2</sup>**

Die Expansion des Internets hat nicht nur der Wirtschaft eine Sonderkonjunktur beschert, sondern auch zur Verbreitung von Online-Befragungen beigetragen (siehe etwa **Pötschke/Simonson** 2001). Mittlerweile gibt es kaum mehr ein renommiertes Meinungsforschungsinstitut, das sich nicht des Internets zur Datenerhebung bediente (siehe etwa **Güllner** 2000: 566). Zudem können Ergebnisse solcher Erhebungen in der Öffentlichkeit mit erheblicher Aufmerksamkeit rechnen (siehe etwa Perspektive Deutschland 2003). Die Beliebtheit von Internetbefragungen lässt sich leicht mit den offenkundigen Vorteilen erklären, die sie gegenüber etablierten Erhebungsformen aufweisen (siehe etwa **Batinic** 2001: 12-14; **Cobanoglu** et al. 2001: 443-444; **Taylor** 2000: 53, 57): In sehr kurzer Zeit können zahlreiche Merkmale vieler Menschen bei einem vergleichsweise kleinen Erhebungsaufwand ermittelt werden. Ferner können anspruchsvolle multimediale Stimuli eingesetzt (siehe **Alvarez** et al.

---

2 Für wertvolle Anregungen danke ich dem anonymen Gutachter.

2003: 23) und besser als bei manchen anderen Befragungsformen Reaktionszeiten gemessen werden (siehe etwa *Bassili* 1995).

Trotz dieser Vorteile werden jedoch Zweifel daran geäußert, dass Onlineumfragen etablierte Erhebungstechniken gleichwertig ersetzen könnten, wenn man valide Aussagen über die Gesamtbevölkerung anstrebt. In der wissenschaftlichen Diskussion wird beispielsweise auf verzerrende Wirkungen unterschiedlicher Web-Browser (siehe *Couper* et al. 2001) und der optischen Gestaltung von Online-Fragebögen (siehe etwa *Gräf* 1999: 159-173; *Dillman* 2000: 352-433) hingewiesen. Als Kernproblem von Online-Erhebungen wurden jedoch starke Selbstselektionseffekte identifiziert (siehe etwa *Bandilla* et al. 2001: 8-11; *Vehovar* et al. 2002: 239). Allerdings konzentrierte sich die Forschung auf diesem Gebiet bislang weitgehend auf Randverteilungen, während Zusammenhänge zwischen Variablen allenfalls am Rande und explorativ untersucht wurden (siehe *Bandilla* et al. 2001: 24-25; siehe aber für die USA *Berrens* et al. 2003). Dies erscheint problematisch, da Zusammenhänge als vergleichsweise robust gegen Stichprobenverzerrungen gelten (siehe *Schnell* 1991a: 133) und Onlineumfragen daher in der Praxis trotz ihrer Defizite geeignet sein könnten, mit geringem Aufwand zumindest über Variablenzusammenhänge in der Gesamtbevölkerung valide Aussagen zu treffen.

Der vorliegende Beitrag vergleicht empirisch die Ergebnisse einer persönlichen Befragung zu politischen Einstellungen mit Resultaten zweier praktisch bedeutsamer Varianten der Online-Umfrage, einer Onlinebefragung eines Online-Panels und einer offenen WWW-Umfrage. Zunächst werden mögliche Verzerrungseffekte theoretisch diskutiert und daraus Hypothesen zu Randverteilungen und Variablenzusammenhängen abgeleitet. Sie werden anschließend anhand von Daten überprüft, die im Umfeld der Bundestagswahl 2002 per Internet und mit persönlichen Interviews erhoben wurden. Abschließend werden die Ergebnisse zusammengefasst und diskutiert.

## 2 Theoretische Überlegungen

Persönliche Interviews, online durchgeführte Interviews zufällig ausgewählter Internetnutzer und offene WWW-Umfragen unterscheiden sich hinsichtlich der Rekrutierung von Respondenten und des Befragungsmodus' (siehe für eine Übersicht **Couper** 2000: 477-490). Aus diesen Unterschieden sollen in diesem Abschnitt Hypothesen darüber abgeleitet werden, wie sich Befunde zu Randverteilungen und Variablenzusammenhängen zwischen den drei Erhebungsformen bei einer Umfrage zu politischen Einstellungen unterscheiden sollten.

In Bezug auf die Rekrutierung der Respondenten kommt das klassische face-to-face-Interview einer echten Zufallsstichprobe recht nahe. Gleichwohl führen Probleme bei der Stichprobenziehung (vgl. **Schnell** 1991a; siehe allgemein **Groves** 1987: S159-S166, 1989: 1-37) und bei der Stichprobenausschöpfung (vgl. **Schnell** 1997: 71-127) zu Verzerrungen zugunsten von Personen mit hoher formaler Bildung und ausgeprägtem Interesse am Gegenstand der Untersuchung (siehe etwa **Esser** 1986a; **Hartmann/Schimpl-Neimanns** 1992; **Koch** 1997).

Ungleich größere Abweichungen vom Ideal einer für die Gesamtbevölkerung repräsentativen Stichprobe treten bei Online-Umfragen unter Mitgliedern eines Online-Panels auf. Die Auswahlgesamtheit umfasst nur Personen, die in einer mündlichen oder telefonischen Befragung angeben, über einen Internetzugang zu verfügen und grundsätzlich an Online-Erhebungen teilnehmen zu wollen, und daher in ein Online-Panel aufgenommen werden. Da Internetnutzer noch immer überdurchschnittlich unter Männern, Personen niedrigen Alters und Höhergebildeten zu finden sind (siehe etwa **Bandilla** 1999: 15; **Bandilla** et al. 2001: 17; **Batinic** 2001: 48-51), sind diese Gruppen auch in der Stichprobe überrepräsentiert. Zudem dürften unter den Internetnutzern die am Befragungsthema interessierten und hoch gebildeten eher bereit sein, sich in ein Online-Panel aufnehmen zu lassen, weshalb die Stichprobe selbst für die Internetnutzer nur eingeschränkt repräsentativ sein dürfte. Die Ausfallprozesse nach der Stichprobenziehung werden – sieht man von internetspezifischen Faktoren wie stillgelegten E-mail-Adressen sowie Telefon- und Providergebühren ab – wie bei Offline-Erhebungen nicht zuletzt vom Interesse am Interview-Gegenstand beeinflusst (siehe etwa **Vehovar** et al. 2002: 235; **Bosnjak** 2002), und zwar vermutlich stärker als bei face-to-face-Interviews, weil keine persönlichen Überzeugungsversuche möglich sind. Insgesamt ist daher bei Online-Umfragen unter zufällig ausgewählten Internetnutzern von einer deutlich stärkeren Verzerrung der Stichprobe als bei mündlichen Umfragen auszugehen.

Bei offenen Online-Befragungen, für die im Internet oder auch in der Offline-Welt geworben wird, ist mit noch gravierenderen Verzerrungen zu rechnen. Nicht einmal im Ansatz wird bei diesem Verfahren aus einer wohldefinierten Auswahlgesamtheit eine Zufallsstichprobe gezogen, vielmehr werden die Respondenten willkürlich ausgewählt (siehe etwa **Bandilla** 1999: 11). Die Selbstselektionseffekte dieses Rekrutierungsprozesses liegen auf der Hand: Soweit die Erhebung im Internet beworben wird, werden bevorzugt Personen mit Internetzugang und einer hohen Nutzungsfrequenz dieses Mediums überhaupt auf die Befragung aufmerksam (siehe **Bandilla** 1999: 12; **Bandilla/Hauptmanns** 1998: 37-41; **Hauptmanns** 1999: 24-29). Zudem werden Hinweise auf eine Erhebung – ob nun im Internet oder in der Offline-Welt – nicht vollkommen wahllos platziert, sondern vor allem dort, wo am ehesten teilnahmebereite Personen zu vermuten sind. Folglich ist die Wahrscheinlichkeit, auf eine offene Web-Umfrage aufmerksam zu werden und überhaupt über die Teilnahme entscheiden zu können, asymmetrisch zugunsten hochgebildeter und am Befragungsthema interessierter Personen verteilt. Unter den informierten Personen hängt die Beteiligungsentscheidung von ähnlichen Faktoren wie bei anderen Online-Befragungen ab: eine Person muss teilnehmen können, also Zugang zum Internet haben, und wollen, also das Ausfüllen des Fragebogens für so wichtig halten, dass sie dafür Zeit und Geld zu investieren bereit ist; deshalb werden sich von den angesprochenen Personen bevorzugt formal hochgebildete und am Gegenstand der Befragung interessierte Menschen an der Erhebung beteiligen (siehe etwa **Bosnjak/Batinic** 1999: 146-149). Infolge der kumulativen Wirkung der drei Schwellen – Internetzugang, Aufmerksamkeit, Teilnahme – ist somit für die offene WWW-Umfrage eine stärkere Verzerrung des Befragtenkreises in Richtung hoher formaler Bildung und inhaltlichen Interesses zu erwarten als für eine Umfrage unter zufällig ausgewählten Internetnutzern.

Aus den Unterschieden in der Rekrutierung der Respondenten lässt sich für die hier betrachteten Umfragen zu politischen Einstellungen Hypothese 1 ableiten: Männer, junge Menschen und hoch Gebildete sollten in Online-Erhebungen generell überrepräsentiert sein, wegen der zusätzlichen Aufmerksamkeitsschwelle dürfte die Verzerrung bei offenen WWW-Umfragen noch größer ausfallen als für Online-Umfragen unter Mitgliedern eines Online-Panels. Daneben sind Unterschiede in der Verteilung politischer Merkmale zu erwarten: Eine hohe Formalbildung hängt zusammen mit politischem Interesse, Wissen und Engagement (siehe etwa **Delli Carpini/Keeter** 1996: 154-174; **Maier** 2000: 144-155), und diese Faktoren begünstigen zusätzlich eigenständig die Aufmerksamkeit für eine offene Online-Umfrage sowie – bei gegebener Information – die Teilnahme daran. Daher ist für Onlineinterviews

generell mit der Überschätzung der politischen Involvierung und des politischen Wissens zu rechnen, vor allem aber bei einer offenen WWW-Umfrage (Hypothese 2).

Die vermuteten Verzerrungen von Randverteilungen können auch dazu führen, dass Zusammenhänge der betreffenden Merkmale in Online-Erhebungen von den in einer mündlich befragten Zufallsstichprobe gemessenen abweichen. In Online-Befragungen kommen bestimmte Merkmalsausprägungen, etwa niedrige formale Bildung oder geringes politisches Interesse, deutlich seltener oder überhaupt nicht vor. Im Extremfall könnte der Wertebereich – gemessen an der Verteilung in einer repräsentativen Stichprobe – nicht vollständig abgebildet werden, in jedem Fall dürfte jedoch die Varianz systematisch unterschätzt werden. In der Folge könnte sich der Zusammenhang zwischen Merkmalen, die mit der Rekrutierungswahrscheinlichkeit zusammenhängen, in einer verzerrten Online-Auswahl anders darstellen als in einer repräsentativen Stichprobe, und zwar in der Regel schwächer ausfallen (siehe *Heckman* 1979; *Berk* 1983; *Groves* 1989: 90-95). Folglich ist für eine Online-Umfrage, vor allem aber eine offene WWW-Umfrage, mit einer Unterschätzung des tatsächlichen Zusammenhangs zu rechnen. Diese Überlegungen werden exemplarisch an den Zusammenhängen zwischen der formalen Bildung, dem politischen Interesse und der subjektiven politischen Kompetenz untersucht, da deren Randverteilung in Online-Erhebungen verzerrt sein sollte (Hypothese 3).

Die verzerrenden Rekrutierungsmechanismen könnten aber auch Zusammenhänge zwischen Merkmalen beeinträchtigen, die nicht unmittelbar mit der Auswahlwahrscheinlichkeit zusammenhängen. In Internetumfragen dürften formal hochgebildete, politisch involvierte und informierte Personen überrepräsentiert sein, die vergleichsweise stark kristallisierte politische Einstellungen und klar strukturierte Einstellungssysteme aufweisen (siehe *Bartle* 1997, 2000; *Converse* 1964: 208; *Delli Carpini/Keeter* 1996: 227-254; *Feldman* 1989; *Zaller* 1990). Abstrahiert man von Messfehlern, sollten daher unter den Teilnehmern an Online-Befragungen, vor allem aber an offenen WWW-Umfragen stärkere Zusammenhänge festgestellt werden als in einer repräsentativen Stichprobe. Folglich dürfte die verzerrte Rekrutierung die Gültigkeit von Reliabilitätsprüfungen von Likertskalen in Online-Erhebungen beeinträchtigen: wegen der stärker ausgeprägten Einstellungsstrukturen bei Online-Teilnehmern könnte die Reliabilität eines Instruments überschätzt werden, so dass es in einem Online-Panel, erst recht aber in offenen WWW-Umfragen als reliabel gelten könnte, ohne in einer repräsentativen Stichprobe die gängigen Anforderungen zu erfüllen. Diese Frage soll für drei Skalen empirisch untersucht werden, eine Rechtsextremismus-Skala (vgl. *Falter* 1994: 136-140) sowie Skalen zur Messung politischen Kompetenzgefühls (internal efficacy) und Responsivitätsgefühls (external efficacy) (vgl. *Vetter* 1997: 113-138) (Hypothese 4).

Hochgradig politisch involviert sind nicht zuletzt Parteianhänger, also Personen, die sich mit einer Partei identifizieren, daher parteipolitisch außerordentlich stark strukturierte politische Einstellungen besitzen, das politische Geschehen verzerrt im Sinne ‚ihrer‘ Partei betrachten und politische Akteure unterschiedlicher Couleur überdurchschnittlich stark als Gegner wahrnehmen (siehe etwa *Falter* et al. 2000: 251-255). Sollten unter Respondenten in (offenen) Online-Umfragen überproportional viele politisch Interessierte und Parteianhänger zu finden sein, müssten daher in Online-Umfragen, vor allem aber in einer offenen WWW-Erhebung die Einstellungen zu Parteien und Kandidaten stärker zusammenhängen als in einem repräsentativen Querschnitt aus den Wahlberechtigten (Hypothese 5).

Gestützt auf diese Überlegung, kann man zwischen den drei Erhebungen auch Unterschiede erwarten, wenn man versucht, Wahlverhalten mit dem klassischen sozialpsychologischen Modell der Michigan-Schule zu erklären, das die Stimmenscheidung auf Parteiidentifikation, Kandidaten- und Sachfragenorientierungen zurückführt (vgl. *Campbell* et al. 1954, 1960). Es ist davon auszugehen, dass die wahlrelevanten Einstellungen – die Parteiidentifikation, die Sachfragen- und die Kandidatenorientierungen – in einer Online-Erhebung, vor allem aber in einer offenen WWW-Umfrage stärker als in einer repräsentativen Stichprobe zusammenhängen, was größere Multikollinearitätsprobleme mit sich bringt (Hypothese 6). Zugleich sollten jedoch wegen der stärkeren Strukturierung der Einstellungen in dieser Gruppe die Attitüden enger mit dem Wahlverhalten zusammenhängen, so dass das klassische Drei-Variablen-Modell in einer Online-Befragung, vor allem aber einer offenen WWW-Erhebung eine höhere Erklärungsleistung erzielen dürfte als in einer repräsentativen Stichprobe (Hypothese 7).

Bislang wurde implizit unterstellt, infolge der unterschiedlichen Rekrutierungsprozesse träten zwischen den drei betrachteten Erhebungsformen nur Unterschiede in den ‚wahren‘ Werten der Variablen auf, nicht aber in den Messfehlern. Dies muss aber nicht zwangsläufig zutreffen, da theoretisch Unterschiede in der Rekrutierung der Respondenten und im Befragungsmodus zu Unterschieden in den Messfehlern führen können.

Eine Wirkung von Unterschieden in der Rekrutierung auf die Häufigkeit von Messfehlern erscheint durchaus plausibel. Denn mit jedem Schritt von der mündlichen hin zur offenen WWW-Umfrage wächst die Verzerrung der Respondenten in Richtung hoher Bildung und inhaltlichen Interesses. Damit nimmt aber die Wahrscheinlichkeit zu, dass die Befragten auskristallisierte Einstellungen besitzen und keine Messfehler auftreten (siehe etwa *Converse* 1964: 208; *Sudman/Bradburn* 1974: 71; *DeMaio* 1984: 273; *Narayan/Krosnick* 1996; siehe aber auch *Reuband* 2000,

2001). Der Verzerrungseffekt bei der Respondentenrekrutierung spricht somit für eine vergleichsweise hohe Datenqualität von Online-Erhebungen. Allerdings hängt die Neigung zu Falschantworten empirisch nur schwach mit der formalen Bildung und Interesse am Interviewinhalt zusammen (siehe etwa *Katosh/Traugott* 1981). Daher sollte die aus dem Rekrutierungsprozess resultierende Verzerrung in der Zusammensetzung der Respondenten in der Regel allenfalls zu geringfügigen Unterschieden in der Messgüte zwischen einer mündlichen Befragung und Online-Erhebungen führen.

Unterschiede in der Güte der Messung könnten auch aus Unterschieden im Befragungsmodus persönlich und online durchgeführter Interviews resultieren. Zum einen werden beim face-to-face-Interview die Fragen mündlich gestellt und nur im Einzelfall mit visuellen Materialien unterstützt, während die Online-Befragung durchgängig mit optischen Stimuli arbeitet. Das kann Respondenten durchaus, wenngleich nicht generell zu unterschiedlichen Antworten veranlassen. Zum anderen kommt es beim mündlichen Interview zu einer persönlichen Interaktion zwischen Befragtem und Interviewer, wohingegen der Respondent einen Online-Fragebogen vollkommen selbständig ausfüllt. Bei der Online-Befragung entfällt daher die Möglichkeit, dass ein Interviewer feststellt, welche Person die Fragen beantwortet,<sup>3</sup> und Unklarheiten und Missverständnisse während des Ausfüllens des Fragebogens ausräumt. Zugleich kann er aber nicht selbst als Störfaktor wirken: erstens ist ausgeschlossen, dass er bewusst oder unbewusst nicht die Äußerungen der Respondenten protokolliert; zweitens kann die Anwesenheit eines Interviewers nicht die Antworten der Befragten beeinflussen und die Tendenz zu als sozial erwünscht wahrgenommenen Antworten verstärken (siehe etwa *Esser* 1986b: 319-321). Da Online-Befragungen zudem keine sofortige Antwort fordern und es den Respondenten erlauben, ihre Antworten gründlicher zu bedenken, könnte man für Online-Umfragen eine vergleichsweise hohe Datenqualität erwarten (siehe etwa *Nicholson* et al. 1998; *Sassenberg/Kreutz* 1999; *Batinic* 2001: 58; *King/Miles* 1995; *Stanton* 1998; *Ilieva* et al. 2002: 368, 374). Dem wird jedoch entgegengehalten, die anonyme Online-Atmosphäre lasse die Hemmschwelle sinken, etwa in langen Fragebatterien unabhängig vom Inhalt bestimmte Antwortmuster zu wählen oder auf andere Weise reine ‚Spaßantworten‘ zu geben (siehe etwa *Batinic* 2001: 57). Alle Teilargumente zusammen betrachtet, ist daher im Allgemeinen kein Unterschied in den vom Befragungsmodus verursachten Messfehlern zu erwarten (siehe *Vehovar* et al. 2002: 239). Da eine sehr ähnliche Folgerung bereits zu den Messfehlerdifferenzen, die aus Un-

---

3 Dies ist beim mündlichen Interview zwar möglich, doch ist diese Technik nicht gegen teilweise oder vollständig gefälschte Interviews gefeit (siehe *Reuband* 1990; *Schnell* 1991b).

terschieden in der Befragtenrekrutierung resultieren, gezogen wurde, ist insgesamt zwischen den drei Erhebungsformen mit keinen Unterschieden in den Messfehlern zu rechnen (Hypothese 8).

Damit sind jedoch bei bestimmten Fragen gravierende Messfehlerdifferenzen nicht ausgeschlossen. Dies gilt nicht zuletzt für Wissensfragen: Befragte finden bei einer Internetumfrage weitaus bessere Möglichkeiten vor, etwaige Wissenslücken unbeobachtet mit einem Blick in Nachschlagewerke – ob nun online oder offline – zu schließen, weshalb das Wissen der Respondenten in Interneterhebungen systematisch überschätzt werden sollte. Mit anderen Worten, der in Hypothese 2 postulierte Wissensvorsprung der Mitglieder des Online-Panels und der Teilnehmer der offenen WWW-Umfrage dürfte zu einem Teil auf einen Mode-Effekt zurückzuführen sein (Hypothese 9).

### 3 Datenbasis

Die vorgestellten Hypothesen werden empirisch anhand dreier Datensätze aus dem DFG-Projekt „Politische Einstellungen, politische Partizipation und Wählerverhalten im vereinigten Deutschland“ überprüft.<sup>4</sup> Der erste Datensatz entstammt einer persönlich-mündlichen Befragung durch das Inra-Institut, Mölln, in der vom 10.8.2002 bis zum 15.11.2002 3263 Personen interviewt wurden. Dem zweiten Datensatz liegt eine Online-Befragung desselben Instituts zugrunde, in deren Rahmen vom 13.9.2002 bis zum 1.10.2002 insgesamt 1065 Personen aus dem „inr@sample-Panel“ der Fragebogen aus der persönlichen Befragung in praktisch identischer Form per Internet vorgelegt wurde. Dieses Access-Panel besteht aus Personen, die in face-to-face- oder telefonischen Umfragen dieses Instituts angeben, einen privaten Internet-Zugang zu besitzen und zur Teilnahme an Online-Befragungen bereit zu sein. Um eine Erhebung durchzuführen, werden aus diesem Personenkreis, dessen Zusammensetzung mit Hilfe von Gewichtungen rechnerisch an die Struktur der privaten Internetnutzer angepasst wurde, per Zufall Personen ausgewählt und zur Teilnahme aufgefordert (siehe für Einzelheiten Inra 2002).

Der dritte Datensatz entstammt der „wahnumfrage2002“, einer offenen Online-Umfrage ohne Teilnahmebeschränkung. Sie wurde auf dreierlei Weise beworben: Erstens wurde versucht, Hinweise in den Medien, etwa in einer populären Polit-Talkshow, zu platzieren; zweitens wurde die Umfrage im Internet an möglichst vielen prominenten Multiplikatorstellen, etwa Politik-Portalen, Mailing-Listen oder

---

4 Projektnehmer sind *Jürgen W. Falter*, *Oscar W. Gabriel* und *Hans Rattinger*.

Newslettern erwähnt oder verlinkt; drittens war mit dem Ausfüllen des Fragebogens die Teilnahme an einem Gewinnspiel verbunden, und es bestand für Teilnehmer die Möglichkeit, die Umfrage weiterzuempfehlen, wobei Empfehlungen die Gewinnchancen steigerten (vgl. *Faas* 2003a: 4-6). Die WWW-Befragung stand in der Zeit vom 20.8. bis zum 22.9.2002 offen, und es konnten insgesamt 34098 – wenigstens teilweise ausgefüllte – Fragebögen registriert werden. Insgesamt entspricht dieses Erhebungsdesign einer willkürlichen Auswahl, wie sie bei vielen offenen Internetumfragen üblich ist.

Um die aus den unterschiedlichen Rekrutierungsmechanismen resultierenden Verzerrungen möglichst unverfälscht abzubilden, wird in der empirischen Analyse auf Gewichtungen verzichtet, die im übrigen die Ergebnisse nur leicht verändern würden, wie Validierungsanalysen zeigen. Lediglich die persönlich-mündlich erhobenen Daten werden ost-west-gewichtet, weil in diesem Fall eine regional disproportional geschichtete Stichprobe gezogen wurde. Da sich die vorliegenden Analysen auf die Wahlberechtigten als Grundgesamtheit beziehen, werden in allen drei Datensätzen nur die deutschen Staatsbürger ab 18 Jahren berücksichtigt.

## **4 Empirische Ergebnisse**

### **4.1 Befunde zu Randverteilungen**

In soziodemographischer Hinsicht entsprechen die empirischen Ergebnisse den in Hypothese 1 formulierten Erwartungen. Wie ein Vergleich der mündlich Befragten mit den zufällig ausgewählten Internet-Nutzern zeigt, surfen bevorzugt Männer, Menschen jüngeren Alters und mit höherer Bildung im WWW – die bis zum Jahr 2002 deutlich gewachsene Verbreitung des Mediums Internet scheint die soziale Verzerrung im Nutzerprofil zwar etwas abgeschwächt, aber nicht beseitigt zu haben (siehe Tabelle 1). Der Schritt hin zur offenen WWW-Umfrage verstärkt die soziodemographische Verzerrung ganz erheblich. Denn unter den WWW-Respondenten finden sich nur 22% Frauen, liegt das mittlere Lebensalter bei nur 33 Jahren und haben drei Viertel der Personen das Abitur absolviert, was in der Flächenstichprobe nur auf ein Viertel der Respondenten zutrifft. Die Rekrutierungsprozesse bei offenen WWW-Umfragen – insbesondere die nur sehr selektive Verbreitung der Information über die Befragung – sorgen für erhebliche Abweichungen des Sozialprofils der Teilnehmer von den Wahlberechtigten insgesamt, aber auch von den Mitgliedern des Online-Panels.

**Tabelle 1** Verteilung soziodemographischer Merkmale, politischer Involvierung, subjektiver politischer Kompetenz und politischen Wissens in den drei Befragengruppen (Standardfehler in Klammern)

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
<b>Soziodemographie</b>			
Frauen (in %)	51,8 (0,9)	40,9 (1,0)	22,1 (0,3)
Lebensalter (aM)	47,5 (0,3)	37,8 (0,4)	33,11 (0,8)
Formale Bildung (aM)	0,42 (0,01)	0,67 (0,01)	0,85 (0,02)
- Varianz	0,17	0,13	0,08
N min.	3079	1067	20678
<b>Politische Involvierung</b>			
Politisches Interesse (aM)	0,53 (0,01)	0,58 (0,01)	0,77 (0,001)
- Varianz	0,06	0,05	0,04
Partei Anhänger (in %)	69,1 (0,9)	82,4 (1,2)	85,2 (0,2)
Stärke Parteibindung (aM)	0,50 (0,01)	0,63 (0,01)	0,70 (0,002)
Parteimitglieder (in %)	5,3 (0,4)	6,7 (0,8)	23,7 (0,3)
Wahlbeteiligung (in %)	91,9 (0,5)	95,4 (0,6)	98,2 (0,1)
N min.	2946	1075	21726
<b>Subjektive politische Kompetenz</b>			
Verständnis (aM)	0,67 (0,004)	0,68 (0,01)	0,81 (0,02)
- Varianz	0,06	0,05	0,03
Aktive Rolle (aM)	0,44 (0,01)	0,55 (0,01)	0,71 (0,03)
- Varianz	0,11	0,08	0,03
Kompliziertheit (aM)	0,42 (0,01)	0,34 (0,01)	0,19 (0,002)
- Varianz	0,09	0,06	0,05
N min.	3110	1086	9534
<b>Politisches Wissen</b>			
Wahlsystem (in %)	48,7 (0,9)	69,0 (1,4)	ne
Bundesländer (in %)	53,2 (0,9)	67,6 (1,5)	ne
Links-Rechts-Position (aM)	0,69 (0,01)	0,69 (0,01)	0,90 (0,002)
N min.	3105	1022	8539

aM: arithmetisches Mittel; Wertebereich: 0-1;

ne: nicht erhoben.

Siehe für die Operationalisierungen den Anhang. Für die offene WWW-Umfrage werden in dieser und den folgenden Tabellen Standardfehler angegeben, obwohl keine Zufallsstichprobe vorliegt.

Die Teilnehmer an beiden Online-Erhebungen sind im Einklang mit Hypothese 2 deutlich stärker politisch involviert und fühlen sich subjektiv merklich kompetenter als die Mitglieder der mündlich befragten Zufallsstichprobe. Im Vergleich zu dieser Gruppe geben die Mitglieder des Online-Panels ein größeres politisches Interesse, häufiger und stärkere Parteibindungen und eine höhere Wahlbeteiligung an. Ebenso

sehen sie sich seltener von politischen Fragen kognitiv überfordert und fühlen sich eher in der Lage, eine aktive politische Rolle zu übernehmen. Lediglich in Bezug auf die Parteimitgliedschaft und ein Item zum Verständnis politischer Fragen lassen sich keine Unterschiede zwischen beiden Gruppen erkennen. Die Teilnehmer an der offenen WWW-Umfrage übertreffen – unabhängig vom verwendeten Indikator – die Mitglieder des Online-Panels deutlich in punkto politischer Involvierung, und damit natürlich erst recht die Zufallsstichprobe. Besonders ins Auge fallen die eklatanten Unterschiede in der subjektiven politische Kompetenz, im subjektiven politischen Interesse sowie in Bezug auf die Parteimitgliedschaft; der hohe Anteil von beinahe einem Viertel deutet auf bewusste Mobilisierungsversuche von Parteien hin, die die natürlichen Selbstrekrutierungsmechanismen verstärken und für wahlbezogene offene Online-Umfragen nicht untypisch sein dürften.

Die Befunde zum politischen Wissen bestätigen ebenfalls die theoretischen Erwartungen. Entsprechend ihrem ausgeprägten politischen Selbstbewusstsein lassen die Teilnehmer der offenen WWW-Umfrage die beiden Vergleichsgruppen deutlich hinter sich, wenn man das ‚Wissen‘ um die Standpunkte der Parteien auf der Links-Rechts-Achse – den einzigen für alle drei Erhebungen verfügbaren, aber problematischen, da nicht unbedingt objektives Wissen messenden Indikator – als Kriterium heranzieht. Die Mitglieder des Online-Panels und der mündlich befragten Zufallsstichprobe unterscheiden sich zwar nicht in dieser Hinsicht, doch geben erstere deutlich häufiger als die letzteren richtige Antworten auf die Fragen nach der Funktionsweise des Wahlsystems und der Zahl der Bundesländer. Allerdings spiegelt die Differenz nicht nur echte Wissensunterschiede zwischen beiden Gruppen wider, sondern auch einen Unterschied im Erhebungsmodus: Da die Internetnutzer unter den mündlich Befragten, die sich soziodemographisch und in der politischen Involvierung praktisch nicht vom online befragten Access-Panel unterscheiden, die Fragen nach dem Wahlsystem und der Zahl der Bundesländer um 6 bzw. 11 Prozentpunkte seltener korrekt beantworten als ihre online befragten Pendanten (nicht tabellarisch ausgewiesen), ist zu folgern, dass – in Einklang mit Hypothese 9 – die Möglichkeit, unbeobachtet mit einem Griff zu einem Lexikon oder ein paar Mausklicks zu einer Internetsuchmaschine Wissenslücken zu schließen, zur Überschätzung von Wissen in Online-Befragungen beiträgt.<sup>5</sup>

---

5 Die Vergleichbarkeit der Ergebnisse könnte zusätzlich dadurch beeinträchtigt werden, dass die repräsentative Online-Erhebung nur kurz vor und nach der Wahl stattfand, die mündliche Befragung aber rund sieben Wochen vor der Wahl begann und etwa sieben Wochen danach endete, weshalb die politische Involvierung, Teilnahmebereitschaft sowie subjektive und objektive Kompetenz infolge des Wahlkampfes überschätzt werden könnten. Diese Möglichkeit wurde untersucht, indem die mündliche Stichprobe in Befragte aufgeteilt wurde, die in der Feldphase des Online-Panels befragt wurden, und solche, die außerhalb dieses Intervalls interviewt wur-

Die Erwartung, Messfehler sollten in den drei Erhebungen in ähnlichem Maße auftreten, wird von der empirischen Evidenz unterstützt. Wie sich Tabelle 2 entnehmen lässt, kommen Antwortmuster, die zwar nicht ausschließlich, aber häufig Zufallsangaben indizieren, in der mündlichen und den beiden Online-Umfragen praktisch gleich häufig vor: in Bezug auf die Nennung der Mittelkategorie ist dies offensichtlich; in Bezug auf die Antwortmonotonie deuten die Befunde zwar auf eine etwas niedrigere Fehlerneigung in der offenen Online-Umfrage hin, doch fällt eine Differenz von 0,05 bei einem Wertebereich von 0 bis 1 praktisch kaum ins Gewicht. Wenngleich hier ungeklärt bleiben muss, ob die von der Messprozedur und der Respondentenrekrutierung verursachten Messfehlerkomponenten je für sich über die drei Umfragen hinweg stabil sind oder sich Unterschiede in den Komponenten gegenseitig neutralisieren, unterstützt die empirische Evidenz Hypothese 8. Folglich spiegeln die Unterschiede in den vorangegangenen Auswertungen nur Differenzen in den ‚wahren‘ Werten, und für die Analyse von Variablenzusammenhängen ist nicht damit zu rechnen, dass Messfehlerdifferenzen eine Rolle spielen.<sup>6</sup>

**Tabelle 2** Indikatoren für Zufallsantworten in den drei Befragengruppen (Standardfehler in Klammern)

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
Mittelkategorie (aM)	0,22 (0,005)	0,22 (0,006)	0,20 (0,001)
N	926	863	8219
Monotone Antworten (aM)	0,34 (0,004)	0,34 (0,005)	0,29 (0,001)
N	1576	959	8439

aM: arithmetisches Mittel; Wertebereich: 0-1. Siehe für die Operationalisierungen den Anhang.

den, und anschließend die hier betrachteten Variablen für beide Gruppen verglichen wurden. Die empirischen Befunde stehen im Einklang mit der Vermutung einer wahlkampfbedingten Überschätzung politischer Kompetenz und Beteiligungsbereitschaft und parteipolitischer Loyalitäten, nicht jedoch des politischen Interesses; in jedem Fall sind die in diesem Vergleich gemessenen Differenzen in den politischen Einstellungen und Verhaltensabsichten aber deutlich kleiner als die Unterschiede zwischen Online- und Offline-Stichprobe. Folglich sind die in Tabelle 1 erkennbaren Unterschiede zwischen beiden Gruppen wegen der inkongruenten Befragungsintervalle etwas inflationiert, doch bleiben sie auch unter Kontrolle dieser Verzerrung erhalten.

- 6 Diese Ergebnisse entkräften auch den Einwand, der empirisch gemessene Vorsprung der Respondenten in beiden Online-Erhebungen in punkto politisches Interesse und Engagement sei dadurch inflationiert, dass die überdurchschnittlich politisch interessierten Online-Nutzer in Interviews ihr Engagement stärker als Durchschnittsbefragte übertrieben; gegen dieses Artefaktargument sprechen zudem hier nicht berichtete Analysen, die nur schwache Zusammenhänge zwischen Falschantworten und politischer Involvierung nachweisen.

Insgesamt lässt sich damit festhalten, dass die in den Online-Befragungen gemessenen Randverteilungen soziodemographischer und politischer Variablen deutlich von den Ergebnissen der mündlichen Befragung abweichen (siehe zu Unterschieden in der parteipolitischen Komposition *Faas* 2003b). Die Diskrepanz ist dabei im Falle der offenen WWW-Umfrage wegen des Rekrutierungsprozesses erheblich größer als für die Online-Befragung eines Access-Panels. Allerdings scheinen die Verzerrungseffekte die Häufigkeit von Messfehlern unberührt zu lassen.

#### 4.2 Befunde zu Zusammenhängen

In Bezug auf Zusammenhänge zwischen Variablen wurde vermutet, dass die starke Verzerrung in Richtung hoher formaler Bildung, politischer Kompetenz und Involvierung dazu führt, dass Zusammenhänge solcher mit der Teilnahmewahrscheinlichkeit korrelierter Merkmale wegen der Varianzeinschränkung in Online-Umfragen in der Regel schwächer ausfallen als in einer repräsentativen Stichprobe. Wie Tabelle 1 zeigt, nimmt die Varianz der formalen Bildung, des politischen Interesses und der Indikatoren zur subjektiven politischen Kompetenz von der mündlichen Befragung über das Online-Panel bis hin zur offenen WWW-Umfrage erwartungsgemäß etwas ab. Betrachtet man die Korrelationen in Tabelle 3, ist für den Zusammenhang zwischen dem politischen Interesse und dem subjektiven politischen Verständnis kein Unterschied zwischen den drei Erhebungen festzustellen. Der im Vergleich zur mündlichen Befragung in der WWW-Umfrage deutlich stärkere Zusammenhang zwischen politischem Interesse und dem Gefühl, eine aktive politische Rolle spielen zu können, spricht zwar für die Vermutung, dass die Art der Erhebung die Stärke des Zusammenhangs beeinflusst; jedoch widerspricht er ihr der Richtung nach, da eine schwächere Korrelation erwartet wurde. In allen anderen Fällen ist zumindest der in der offenen WWW-Umfrage gemessene Zusammenhang dem Betrage nach schwächer als in der mündlichen Befragung; in drei Fällen trifft dies auch auf die Online-Befragung des Access-Panels zu. Hypothese 3 wird somit eingeschränkt bestätigt.<sup>7</sup>

---

7 Die Varianzeinschränkung resultiert aus dem überproportionalen Ausfall wenig gebildeter und politisch wenig involvierter Personen. Da – wie in Abschnitt 2 gezeigt – die Verzerrung zugunsten hoher Bildung und starker politischer Involvierung zu einer Überschätzung von Zusammenhängen beitragen kann, wurden ergänzende Analysen durchgeführt. Sie zeigen, dass die in Tabelle 3 für die beiden Online-Erhebungen berichteten Korrelationen infolge der skizzierten Verzerrungseffekte leicht erhöht sind. Folglich wird Hypothese 3 unter Berücksichtigung dieses Arguments von der empirischen Evidenz noch deutlicher unterstützt.

**Tabelle 3** Korrelation zwischen formaler Bildung, politischem Interesse und subjektiver politischer Kompetenz in den drei Befragtengruppen (Unter- und Obergrenze der 95%-Konfidenzintervalle um Pearsons  $r$ )

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
Bildung – Interesse	0,24; 0,31	0,11; 0,23	0,12; 0,14
Interesse – Aktivität	0,38; 0,44	0,43; 0,52	0,50; 0,53
Interesse – Verstehen	0,43; 0,49	0,46; 0,54	0,47; 0,50
Interesse – Kompliziertheit	-0,45; -0,39	-0,49; -0,39	-0,40; -0,37
Bildung – Aktivität	0,25; 0,32	0,09; 0,21	0,14; 0,18
Bildung – Verstehen	0,19; 0,26	0,15; 0,26	0,13; 0,17
Bildung – Kompliziertheit	-0,35; -0,29	-0,22; -0,11	-0,18; -0,14
N min	3073	1049	9262

Die Erwartung, in (offenen) Online-Erhebungen würde die Reliabilität von Skalen zur Messung politischer Attitüden überschätzt, findet nur wenig Rückhalt in der empirischen Evidenz. Im Fall der Rechtsextremismusskala nach *Falter* (1994) unterscheiden sich die Reliabilitätsschätzungen zwischen den Erhebungen praktisch nicht. Für die Skala zur Messung von internal efficacy (vgl. *Vetter* 1997: 113-138) wird in der mündlichen Befragung sogar eine etwas höhere Reliabilität als in den Online-Umfragen ermittelt (siehe Tabelle 4), was mit der soeben diskutierten Varianzeinschränkung zusammenhängen dürfte. Nur in Bezug auf die external efficacy (vgl. *Vetter* 1997: 113-138), also das Responsivitätsgefühl, übertrifft Cronbachs  $\alpha$  unter den selbstrekrutierten Teilnehmern an der Wahlumfrage 2002 die entsprechenden Reliabilitätskennwerte in den anderen beiden Befragtengruppen; aber selbst in diesem Fall ist der Unterschied nur moderat. Im Ergebnis ist Hypothese 4 kaum mit der empirischen Evidenz vereinbar. Mit kleinen Einschränkungen scheinen folglich selbst offene WWW-Umfragen für Pretests geeignet, in denen die Reliabilität von Messinstrumenten geprüft werden soll.

**Tabelle 4** Reliabilitäten verschiedener Skalen in den drei Befragtengruppen (Cronbachs  $\alpha$ )

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
Rechtsextremismus	0,77	0,79	ne
Internal efficacy	0,71	0,70	0,66
External efficacy	0,62	0,59	0,68
N min.	2333	1050	8245

Mit dem verwendeten Programmpaket Stata 8.0 lassen sich für Cronbachs  $\alpha$  keine Konfidenzintervalle berechnen; deshalb werden nur die Punktschätzungen angegeben.  
ne: nicht erhoben.

Hypothese 5 postuliert, der vergleichsweise hohe Anteil an Parteihängern unter den Teilnehmern an (offenen) Online-Befragungen (siehe Tabelle 1) Sorge dafür, dass in Internetumfragen politische Objekte relativ stark entlang parteipolitischer Linien wahrgenommen und beurteilt würden. Diese Vermutung wird von der empirischen Evidenz unterstützt, wie die Zusammenhänge zwischen den Urteilen über die beiden Kanzlerkandidaten sowie deren Parteien belegen (siehe Tabelle 5). Denn die Befragten aus dem Online-Panel nehmen die Kanzlerkandidaten *Schröder* und *Stoiber* erheblich stärker als Antipoden wahr als die Teilnehmer an der mündlichen Erhebung, werden darin aber nochmals von den Teilnehmern an der WWW-Umfrage übertroffen. Ganz ähnlich werden die beiden großen Parteien SPD und CDU/CSU von ihnen polarisierter beurteilt. Beinahe folgerichtig erscheint es vor diesem Hintergrund, dass bei den online befragten Personen die Sympathiebewertungen eines Kandidaten und der Partei seines Rivalen um das Amt des Regierungschefs stärker negativ miteinander korrelieren als unter den Offline-Respondenten. Im Vergleich dazu nimmt die Korrelation zwischen den Sympathien für einen Kanzlerkandidaten und seiner Partei beim Schritt zu den Teilnehmern an der offenen WWW-Umfrage nur schwach zu; nicht zuletzt dürfte dies daran liegen, dass das Potential für einen Anstieg in diesen Fällen nur klein ist, da diese Bewertungen bereits in der mündlich befragten Stichprobe mit über 0,7 stark korrelieren. Ein ähnliches Muster ist zu beobachten, wenn man die Bewertungen der Kanzlerkandidaten *Schröder* und *Stoiber* auf verschiedenen Einzeldimensionen mit einem summarischen Urteil über beide korreliert: Die Korrelationen sind zwischen persönlichem Interview und dem Online-Panel mit einer Ausnahme statistisch nicht signifikant verschieden, in den meisten Fällen bringt der Schritt zur offenen WWW-Umfrage jedoch einen deutlichen Anstieg. Verglichen mit dem Zuwachs der Polarisierung, der in den Wahrnehmungen der verschiedenen ‚Parteilager‘ festzustellen ist, erscheint auch er allerdings als ausgesprochen moderat. Die Teilnehmer an den beiden Online-Umfragen, vor allem aber die Respondenten der offenen Internetbefragung, besitzen folglich zwar nicht generell, aber zumindest *parteipolitisch* überdurchschnittlich stark strukturierte Einstellungssysteme, weshalb eine Verallgemeinerung von Befunden aus (offenen) Online-Umfragen zu einer systematischen Überschätzung der Zusammenhänge zwischen parteipolitischen Attitüden führte.<sup>8</sup>

---

8 Gegen diese Feststellung könnten mehrere Artefaktargumente vorgebracht werden. Erstens könnte die im Vergleich zur mündlichen Erhebung etwas variierte optische Darstellung des Fragestimulus starke Zusammenhänge in den Online-Befragungen etwas begünstigen; wie hier nicht tabellarisch ausgewiesene Analysen zeigen, ist dieser Effekt nur schwach ausgeprägt und kann die empirischen Differenzen nicht erklären. Zweitens könnte man einwenden, für die stärkeren Korrelationen unter den zufällig ausgewählten Internetnutzern sei nicht die verzerrte Auswahl, sondern der Befragungszeitpunkt in unmittelbarer Nähe der Wahl verantwortlich, der

**Tabelle 5** Korrelationen zwischen Kandidaten- und Parteisympathien sowie zwischen der summarischen Kanzlerbewertung und den Urteilen auf vier Dimensionen der Kandidatenbewertung in den drei Befragtengruppen (Unter- und Obergrenze der 95%-Konfidenzintervalle um Pearsons r)

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
<b>Korrelation: Kandidaten-/Parteisympathie</b>			
Schröder – Stoiber	-0,40; -0,34	-0,55; -0,46	-0,61; -0,59
Schröder – SPD	0,71; 0,74	0,69; 0,75	0,85; 0,86
Schröder – CDU/CSU	-0,44; -0,38	-0,56; -0,47	-0,61; -0,59
Stoiber – CDU/CSU	0,76; 0,78	0,79; 0,83	0,90; 0,91
Stoiber – SPD	-0,39; -0,33	-0,54; -0,46	-0,60; -0,59
SPD – CDU/CSU	-0,36; -0,32	-0,51; -0,42	-0,58; -0,56
N min.	2981	1096	22737
<b>Korrelation: Gesamt-/Teilurteile zu Kandidaten</b>			
<i>Schröder</i>			
Wirtschaft	0,67; 0,71	0,70; 0,75	0,78; 0,80
Sympathie	0,68; 0,71	0,69; 0,75	0,77; 0,79
Vertrauen	0,71; 0,74	0,74; 0,79	0,83; 0,84
Tatkraft	0,62; 0,67	0,64; 0,71	0,74; 0,76
N min.	2992	1078	8835
<i>Stoiber</i>			
Wirtschaft	0,64; 0,68	0,67; 0,73	0,72; 0,74
Sympathie	0,71; 0,75	0,76; 0,80	0,82; 0,83
Vertrauen	0,71; 0,74	0,70; 0,76	0,79; 0,80
Tatkraft	0,56; 0,61	0,56; 0,64	0,62; 0,64
N min.	2886	1078	8815

infolge des stärkeren Wahlkampfeinflusses zur Messung parteipolitisch stärker strukturierter Einstellungen führen müsse. Dieses Gegenargument wurde überprüft, indem die für die mündlich Befragten die betrachteten Korrelationen – wie in Fußnote 4 beschrieben – getrennt nach dem Befragungszeitpunkt berechnet wurden. Dabei bestätigt sich der Verdacht teilweise, doch kann der Zeitpunkt der Erhebung vor allem in Bezug auf die in Tabelle 6 eingetragenen Zusammenhänge die Unterschiede zwischen Spalte 2 und 3 erklären. Drittens könnte man kritisieren, die WWW-Umfrage sei nur vor der Wahl und daher nur unter dem direkten Einfluss des Wahlkampfes durchgeführt worden, weshalb überhöhte Korrelationen gemessen würden. Berechnet man in der mündlichen Befragung die Korrelationen getrennt für die Vor- und die Nachwählerhebung, resultieren für die Vorwahlinterviews jedoch zumindest tendenziell schwächere Zusammenhänge als für die Nachwählerhebung, so dass das hier gewählte Analysedesign den Unterschied der Ergebnisse aus der WWW-Befragung zu den anderen Erhebungen eher unter- als überschätzt. Schließlich könnte es sich bei dem Anstieg der Kandidaten-Parteien-Korrelationen beim Übergang zur offenen WWW-Umfrage (siehe Tabelle 5) um ein Methodenartefakt handeln, da in der „Wahlumfrage2002“ die Kandidaten-Sympathieskalometer, anders als in den Vergleichsstudien, unmittelbar nach den Parteiskalometern eingesetzt wurden. Doch lässt sich auch dieser Einwand entkräften, da in allen Erhebungen die Parteiskalometer in einem Block eingesetzt wurden, zwischen den drei Befragtengruppen aber auch hinsichtlich der darauf bezogenen Korrelationen erhebliche Unterschiede erkennbar sind. Bei aller grundsätzlichen Berechtigung der Einwände dürften daher die dargestellten Unterschiede in der parteipolitischen Strukturierung der Einstellungen kein Artefakt sein.

**Tabelle 6** Korrelationen zwischen Parteibindung, Kandidaten- und Sachfragenorientierung (Unter- und Obergrenze der 95%-Konfidenzintervalle um Pearsons  $r$ ) und deren Wirkung auf das Wahlverhalten in den drei Befragtenengruppen (unstandardisierte Logitkoeffizienten; Standardfehler in Klammern)

	Mündliches Interview	Online-Panel	Offene WWW-Umfrage
<b>a) Korrelationen</b>			
Parteibindung – Issueorientierung	0,70; 0,74	0,78; 0,82	0,82; 0,83
Parteibindung – Kandidatenpräferenz	0,67; 0,71	0,71; 0,77	0,74; 0,75
Issue-/Kandidatenpräferenz	0,68; 0,72	0,70; 0,75	0,76; 0,77
N min.	2716	1019	21852
<b>b) Erklärungsmodell</b>			
Konstante	-0,30 (0,12)	-0,08 (0,23)	-0,19 (0,06)
Parteibindung	2,26 (0,16)	1,58 (0,19)	1,80 (0,05)
Kandidatenorientierung	1,18 (0,13)	1,38 (0,24)	1,46 (0,06)
Issueorientierung	1,29 (0,13)	1,23 (0,20)	1,76 (0,05)
Pseudo-R <sup>2</sup> (x 100)	75,0	77,5	83,3
N	2047	801	19887

Im oberen Teil von Tabelle 6 sind die Korrelationen zwischen Parteiidentifikation, Sachfragen- und Kandidatenorientierungen zusammengestellt, also zwischen den drei Größen, auf die im Michigan-Modell Wahlverhalten zurückgeführt wird. Empirisch übertrifft eine von drei Korrelationen im Online-Panel ihr Pendant in der mündlichen Befragung. In allen drei Fällen liegt die in der WWW-Umfrage gemessene Korrelation deutlich über jener in der mündlichen Befragung. Hypothese 6 kann somit zumindest in Bezug auf den Vergleich zwischen mündlicher Befragung und WWW-Umfrage als bestätigt gelten. Folglich fallen unter den Teilnehmern an der offenen Online-Erhebung beim Versuch, mit dem Michigan-Modell Wahlverhalten zu klären, Multikollinearitätsprobleme etwas stärker ins Gewicht als in der Offline-Umfrage.

Wie der untere Teil von Tabelle 6 zeigt, nimmt im Einklang mit Hypothese 7 die Gesamterklärungskraft des Michigan-Modells um 2,5 Punkte zu, wenn man von der mündlichen Befragung zum Access-Panel übergeht; beim Schritt zur offenen Web-Umfrage wächst die Differenz sogar auf über 8 Punkte an. Darüber hinaus zeichnet sich eine Gewichtsverlagerung zwischen den drei Wahlverhaltensdeterminanten ab: Der Parteibindungseffekt ist in der WWW-Umfrage signifikant schwächer als in der mündlichen Befragung, während für den Sachfrageneinfluss das Gegenteil gilt. Die

offene Web-Umfrage liefert somit ein merklich anderes Bild von den Bestimmungsgründen des Wahlverhaltens als die mündlich befragte Zufallsstichprobe.<sup>9</sup>

Insgesamt werden somit die Hypothesen zu Variablenzusammenhängen in mündlich und in online durchgeführten Befragungen empirisch größtenteils bestätigt. Zwar finden die Bedenken gegen Online-Daten als Grundlage für Reliabilitätsanalysen nur wenig Unterstützung. Doch in vielen anderen Fällen weichen zumindest die Befunde aus der offenen WWW-Umfrage erheblich von den Ergebnissen der persönlichen Befragung ab.

## 5 Schluss

Der vorliegende Beitrag vergleicht empirisch die Ergebnisse einer persönlichen Befragung zu politischen Themen mit Resultaten zweier praktisch bedeutsamer Varianten der Online-Umfrage, einer Online-Befragung eines Online-Panels und einer offenen WWW-Umfrage. Dazu wurden Unterschiedshypothesen zu Randverteilungen und Variablenzusammenhängen entwickelt. Sie wurden empirisch überprüft, indem die Ergebnisse einer persönlichen Befragung einer Zufallsstichprobe, einer Online-Befragung von zufällig ausgewählten Internetnutzern sowie einer offenen WWW-Umfrage zur Bundestagswahl 2002 einander gegenübergestellt wurden.

Die empirische Analyse belegt systematische Verzerrungen univariater Häufigkeitsverteilungen in Online-Erhebungen: Männer, Menschen jüngeren Alters und mit hoher formaler Bildung sind überrepräsentiert. Ebenso sind Internetnutzer generell politisch interessierter, besser informiert und aktiver als die Gesamtbevölkerung. In besonderem Maße gilt dies für die Teilnehmer an der offenen WWW-Umfrage; gerade diese verbreitete Erhebungsform erreicht offenkundig vor allem eine kleine politische Elite aus hochgradig interessierten, informierten und engagierten Parteilanhängern, nicht aber einen repräsentativen Querschnitt der Gesamtbevölkerung. Sieht man von Wissensfragen ab, lassen sich in den Messfehlern keine Differenzen zwischen den Erhebungen feststellen. Zusammenhänge zwischen Variablen werden in den beiden Online-Erhebungen im Vergleich zur mündlich befrag-

---

9 In der offenen WWW-Umfrage wurden die drei Komponenten des Michigan-Modells zeitlich weniger entzerrt erhoben als in den beiden anderen Befragungen, weshalb die Variation der Fragereihenfolge für eine Inflationierung der betrachteten Zusammenhänge sorgen könnte. Da in der offenen Online-Umfrage – wie bereits gesehen – jedoch auch Zusammenhänge zwischen Variablen, die in allen drei Befragungen in identischer Reihenfolge gemessen wurden oder in der WWW-Umfrage weiter voneinander entfernt sind, merklich erhöht sind, dürften die in Bezug auf das Michigan-Modell gemessenen Differenzen nur zu einem kleinen Teil auf die Variation der Fragereihenfolge zurückgeführt werden können.

ten Zufallsstichprobe ebenfalls verzerrt. Zwar unterscheiden sich die betrachteten Reliabilitätsschätzungen zwischen den verschiedenen Interviewvarianten nur geringfügig. Jedoch variieren Korrelationen zwischen Variablen, die mit der Rekrutierungswahrscheinlichkeit in Online- und Offline-Erhebungen zusammenhängen, etwa die formale Bildung und das politische Interesse, merklich zwischen den Umfragen. Zusammenhänge zwischen parteipolitischen Einstellungen und die Erklärungskraft politischer Einstellungen in Bezug auf das Wahlverhalten fallen in Online-Umfragen, vor allem aber in der offenen WWW-Erhebung, deutlich stärker aus als in der mündlich befragten Zufallsstichprobe.

Insgesamt sprechen die erheblichen empirischen Abweichungen dafür, dass Online-Erhebungen, insbesondere aber offene Web-Umfragen zumindest in Bezug auf die hier betrachteten Merkmale kein zutreffendes Bild von den Verhältnissen in der Gesamtbevölkerung liefern. Daher können Online-Befragungen, vor allem aber offene WWW-Umfragen in der Praxis der Gesellschaftsbeobachtung kaum als adäquater Ersatz für etablierte Erhebungsverfahren dienen. Dies gilt nicht nur, wenn man Randverteilungen von Variablen untersucht, sondern auch dann, wenn man Variablenzusammenhänge betrachtet. Interessiert man sich nicht nur für Internetnutzer, so scheinen Online-Erhebungen momentan, wenn überhaupt, am ehesten noch zur Kalibrierung von Instrumenten in Pretests geeignet. Die weitere Forschung sollte der Frage nachgehen, ob diese Befunde auf andere Themenfelder, etwa andere Einstellungen, übertragen werden können. Ebenso erscheint es sinnvoll, zu untersuchen, inwieweit Fortschritte der Internetpenetration in Deutschland die Aussagekraft von Online-Befragungen steigern.

## 6 Anhang: Operationalisierung der verwendeten Variablen

*Formale Bildung:* Codierung: 0: maximal Hauptschulabschluss; 0,5: Mittlere Reife; 1: (Fach-)Hochschulreife.

*Politisches Interesse:* „Wie stark interessieren Sie sich für Politik? Würden Sie sagen sehr stark, ziemlich stark, mittelmäßig, weniger stark oder überhaupt nicht?"; Codierung: 0: „überhaupt nicht“ oder „weniger stark“; 0,5: „mittelmäßig“; 1: „ziemlich stark“ oder „sehr stark“.

*Parteiidentifikation:* „Viele Leute in der Bundesrepublik neigen längere Zeit einer bestimmten Partei zu, obwohl sie auch ab und zu mal eine andere Partei wählen. Wie ist das bei Ihnen: Neigen Sie – ganz allgemein gesprochen – einer bestimmten Partei zu? Wenn ja, welcher?"; Codierung: Parteiläufer: 0: nein, 1: ja; Stärke der Parteiidentifikation: 0: keine Parteiidentifikation; 0,5: sehr schwach, schwach, mittelmäßig; 1: ziemlich stark oder sehr stark; Richtung der Parteibindung: -1: Präferenz CDU/CSU-FDP; 0: keine/andere Partei; 1: SPD-B90/Grüne.

*Rechtsextremismus:* „Ich bin stolz ein Deutscher zu sein.“, „Wir sollten endlich wieder Mut zu einem starken Nationalgefühl haben.“, „Gruppen- und Verbandsinteressen sollten sich bedingungslos dem Allgemeinwohl unterordnen.“, „Unter bestimmten Umständen ist eine Diktatur die bessere Staatsform.“, „Der Nationalsozialismus hatte auch seine guten Seiten.“, „Ohne die Judenvernichtung würde man Hitler heute als einen großen Staatsmann ansehen.“, „Die Bundesrepublik ist durch die vielen Ausländer in einem gefährlichen Maß überfremdet.“, „Ausländer sollten grundsätzlich ihre Ehepartner unter ihren eigenen Landsleuten auswählen.“, „Auch heute noch ist der Einfluss von Juden zu groß.“, „Die Juden haben einfach etwas Besonderes und Eigentümliches an sich und passen daher nicht so recht zu uns.“

### *Internal efficacy*

*Aktivität:* „Ich traue mir zu, in einer Gruppe, die sich mit politischen Fragen befasst, eine aktive Rolle zu übernehmen.“

*Kompliziertheit:* „Die ganze Politik ist so kompliziert, dass jemand wie ich nicht versteht, was vorgeht.“ (für Reliabilitätsanalyse umgepolt)

*Verständnis:* „Wichtige politische Fragen kann ich gut verstehen und einschätzen.“

(jeweils Fünferskala: Codierung für arithmetisches Mittel: 0: stimme überhaupt nicht zu; 0,25: stimme eher nicht zu; 0,5: stimme teils zu/teils nicht zu; 0,75: stimme eher zu; 1: stimme voll und ganz zu).

*External efficacy*

„Politiker kümmern sich darum, was einfache Leute denken.“, „Die Bundestagsabgeordneten bemühen sich um einen engen Kontakt zur Bevölkerung.“, „Die Parteien wollen nur die Stimmen der Wähler, ihre Ansichten interessieren sie nicht.“ (Antwortvorgaben: Fünfer-Ratingskalen).

*Politisches Wissen*

- Wahlsystem: „Bei der Bundestagswahl ist es so, dass der Wähler zwei Stimmen hat, eine Erststimme und eine Zweitstimme. Was ist eigentlich die wichtigere Stimme, wodurch wird letztlich über die Stärke der Parteien im Bundestag entschieden: durch die Erststimme oder durch die Zweitstimme oder sind Erst- und Zweitstimme hierfür gleich wichtig?

- Zahl der Bundesländer: „Wissen Sie, wie viele Bundesländer die Bundesrepublik Deutschland heute insgesamt hat, alte und neue Bundesländer zusammengerechnet?“

- Links-Rechts-Positionen der Parteien: es wurde für alle Parteien, zu denen Links-Rechts-Einstufungen der Respondenten vorliegen (SPD, CDU/CSU, B90/Grüne, FDP, PDS, REP) paarweise Vergleiche vorgenommen und überprüft, ob die relative Einordnung („Partei A ist links von Partei B positioniert“) eines Individuums mit der Einordnung in der Gesamtstichprobe übereinstimmt; dies wurde für alle möglichen Partei-paare wiederholt und anschließend der Anteil der Übereinstimmungen zwischen individueller und kollektiver Einordnung für jeden einzelnen Respondenten errechnet.

*Mittelkategorie:* Es wurden die seitens der Befragten wahrgenommenen Standpunkte von SPD, CDU/CSU, FDP, B90/Grüne und PDS (REP wurden in der WWW-Umfrage nicht erhoben) zu drei Positionsissues betrachtet und für jeden einzelnen Respondenten ermittelt, wie häufig er die Mittelkategorie angab. Issues: „Sollte die Kernenergie weiter ausgebaut werden oder sollten alle Kernkraftwerke abgeschaltet werden?“, „Sollten die Zuzugsmöglichkeiten für Ausländer erleichtert oder eingeschränkt werden?“, „Sollte die europäische Einigung so vorangetrieben werden, dass es bald eine gemeinsame europäische Regierung gibt, oder geht die europäische Einigung schon jetzt viel zu weit?“ (die wahrgenommene Parteiposition wurde jeweils einer 7er-Ratingskala gemessen).

*Antwortmonotonie:* Für die gleichen drei Positionsissues wurde ermittelt, welche Antwortkategorie ein Befragter bei den insgesamt 15 Positionseinschätzungen am

häufigsten wählte; anschließend wurde die absolute Häufigkeit dieser Kategorie durch 15 dividiert.

*Sympathie für Parteien und Kandidaten:* Was halten Sie – ganz allgemein – von der SPD, CDU/CSU, Gerhard Schröder, Edmund Stoiber? (Antwortvorgaben: -5 bis +5).

*Dimensionen der Kandidatenbewertung:* „Er hat vernünftige Vorstellungen, um die Wirtschaft anzukurbeln.“, „Er ist menschlich sympathisch.“, „Er ist politisch vertrauenswürdig.“, „Er ist tatkräftig.“ (Antwortvorgaben: Fünfer-Ratingskala: „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“).

*Sachkompetenz:* „Was sind Ihrer Meinung nach die wichtigsten Probleme, die es heute in der Bundesrepublik zu lösen gilt?“ - „Das für Sie wichtigste Problem ist .... Welche Partei ist Ihrer Meinung nach am besten geeignet, dieses Problem zu lösen?“; Codierung: -1: Präferenz CDU/CSU-FDP; 0: keine/andere Partei; 1: SPD-B90/Grüne.

*Kandidatenpräferenz:* Vergleich der Sympathiewerte für Gerhard Schröder und Edmund Stoiber; höherer Wert Stoiber: -1; Gleichstand: 0; höherer Wert Schröder: 1.

*Wahlentscheidung:* Wahlabsicht (Vorwahl) oder rückerinnerte Wahlentscheidung (Nachwahl) bezogen auf den 22.9.2002.

## Literatur

- Alvarez, R. Michael; Sherman, Robert P.; VanBeselaere, Carla*, 2003: Subject Acquisition for Web-Based Surveys, in: *Political Analysis* 11, 23-43.
- Bandilla, Wolfgang*, 1999: WWW-Umfragen – Eine alternative Datenerhebungstechnik für die empirische Sozialforschung?, in: *Batinic, Bernad; Werner, Andreas; Gräf, Lorenz; Bandilla, Wolfgang* (Hrsg.), *Online Research*, Göttingen: Hogrefe, 9-19.
- Bandilla, Wolfgang; Hauptmanns, Peter*, 1998: Internetbasierte Umfragen als Datenerhebungstechnik für die Empirische Sozialforschung?, in: *ZUMA-Nachrichten* 43, 36-53.
- Bandilla, Wolfgang; Bosnjak, Michael; Altdorfer, Patrick*, 2001: Effekte der Erhebungsverfahren? Ein Vergleich zwischen einer Web-basierten und einer schriftlichen Befragung zum ISSP-Modul Umwelt, in: *ZUMA-Nachrichten* 49, 7-28.
- Bartle, John*, 1997: Political awareness and heterogeneity in models of voting: Some evidence from recent British election studies, in: *Pattie, Charles J.; Denver, David; Fisher, Justin; Ludlam, Steven* (Hrsg.), *British Elections Parties Review* 7, London: Frank Cass, 1-22.
- Bartle, John*, 2000: Political awareness, opinion constraint and the stability of ideological positions, in: *Political Studies* 48, 467-484.
- Bassili, John N.*, 1995: On the psychological reality of party identification, in: *Political Behavior* 17, 339-358.
- Batinic, Bernad*, 2001: Fragebogenuntersuchungen im Internet, Aachen: Shaker.
- Berk, Richard A.*, 1983: An introduction to sample selection bias in sociological data, in: *American Sociological Review* 48, 386-398.
- Berrens, Robert P.; Bohara, Alok K.; Jenkins-Smith, Hank; Silva, Carol; Weimer, David*, 2003: The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples, in: *Political Analysis* 11, 1-22.
- Bosnjak, Michael*, 2002: (Non)Response bei Web-Befragungen, Aachen: Shaker.
- Bosnjak, Michael; Batinic, Bernad*, 1999: Determinanten der Teilnahmebereitschaft an internet-basierten Fragebogenuntersuchungen am Beispiel E-Mail, in: *Batinic, Bernad; Werner, Andreas; Gräf, Lorenz; Bandilla, Wolfgang* (Hrsg.), *Online Research*, Göttingen: Hogrefe, 145-157.
- Campbell, Angus; Gurin, Gerald; Miller, Warren E.*, 1954: *The Voter Decides*, Evanston: Row, Peterson and Company.
- Campbell, Angus; Converse, Philip E.; Miller, Warren E.; Stokes, Donald E.*, 1960: *The American Voter*, New York: Wiley.
- Cobanoglu, Cihan; Warde, Bill; Moreo, Patrick J.*, 2001: A comparison of mail, fax and web-based survey methods, in: *International Journal of Market Research* 43, 441-452.
- Converse, Philip E.*, 1964: The Nature of Belief Systems in Mass Publics, in: *Apter, David E.* (Hrsg.): *Ideology and Discontent*, New York: Free Press, 206-261.
- Couper, Mick P.*, 2000: Web Surveys, in: *Public Opinion Quarterly* 64, 464-494.
- Couper, Mick P.; Traugott, Michael W.; Lamias, Mark J.*, 2001: Web Survey Design and Administration, in: *Public Opinion Quarterly* 65, 230-253.
- Delli Carpini, Michael X.; Keeter, Scott*, 1996: *What Americans Know about Politics and Why It Matters*, New Haven: Yale University Press.

- DeMaio, Theresa J.**, 1984: Social Desirability and Survey Measurement: A Review, in: **Turner, Charles F.; Martin, Elizabeth** (Hrsg.): *Surveying Subjective Phenomena*, Band 2, New York: Russell Sage Foundation, 257-281.
- Dillman, Don A.**, 2000: *Mail and Internet Surveys*, 2. Auflage, New York: Wiley.
- Esser, Hartmut**, 1986a: Über die Teilnahme an Befragungen, in: *ZUMA-Nachrichten* 18, 38-47.
- Esser, Hartmut**, 1986b: Können Befragte lügen? in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38, 314-336.
- Faas, Thorsten**, 2003a: [www.wahlumfrage2002.de](http://www.wahlumfrage2002.de). Analysen und Ergebnisse, Bamberg.
- Faas, Thorsten**, 2003b: Umfragen im Umfeld der Bundestagswahl 2002: Offline und Online im Vergleich, in: *ZA-Information* 52, 120-135.
- Falter, Jürgen W.**, 1994: *Wer wählt rechts?* München: C.H. Beck.
- Falter, Jürgen W.; Schoen, Harald; Caballero, Claudio**, 2000: Dreißig Jahre danach: Zur Validierung des Konzepts ‚Parteiidentifikation‘ in der Bundesrepublik, in: **Klein, Markus; Jagodzinski, Wolfgang; Mochmann, Ekkehard; Ohr, Dieter** (Hrsg.): *50 Jahre Empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten*, Wiesbaden: Westdeutscher Verlag, 235-271.
- Feldman, Stanley**, 1989: Reliability and stability of policy positions: Evidence from a five-wave panel, in: *Political Analysis* 1, 25-60.
- Gräf, Lorenz**, 1999: Optimierung von WWW-Umfragen: Das Online Pretest-Studio, in: **Batinic, Bernad; Werner, Andreas; Gräf, Lorenz; Bandilla, Wolfgang** (Hrsg.), *Online Research*, Göttingen: Hogrefe, 159-177.
- Groves, Robert M.**, 1987: Research on Survey Data Quality, in: *Public Opinion Quarterly* 51, S156-S172.
- Groves, Robert M.**, 1989: *Survey Errors and Survey Costs*, New York: Wiley.
- Güllner, Manfred**, 2000: Methodische Entwicklungen in der Empirischen Wahlforschung, in: **Klein, Markus; Jagodzinski, Wolfgang; Mochmann, Ekkehard; Ohr, Dieter** (Hrsg.), *50 Jahre Empirische Wahlforschung in Deutschland*, Wiesbaden: Westdeutscher Verlag, 564-583.
- Hartmann, Peter H.; Schimpl-Neimanns, Bernhard**, 1992: Sind Sozialstrukturanalysen mit Umfragedaten möglich? in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 44, 315-340.
- Hauptmanns, Peter**, 1999: Grenzen und Chancen von quantitativen Befragungen mit Hilfe des Internet, in: **Batinic, Bernad; Werner, Andreas; Gräf, Lorenz; Bandilla, Wolfgang** (Hrsg.), *Online Research*, Göttingen: Hogrefe, 21-38.
- Heckman, James J.**, 1979: Sample Selection Bias as a Specification Error, in: *Econometrica* 47, 153-161.
- Ilieva, Janet; Baron, Steve; Healey, Nigel M.**, 2002: Online surveys in marketing research: pros and cons, in: *International Journal of Market Research* 44, 361-376.
- Inra, 2002: Anlage und Methode der Untersuchung „Bundestagswahl 2002“ für die Otto-Friedrich-Universität Bamberg von INRA Deutschland GmbH, Mölln: Inra.
- Katosh, John P.; Traugott, Michael W.**, 1981: The Consequences of Validated and Self-reported Voting Measures, in: *Public Opinion Quarterly* 45, 519-535.
- King, W. C.; Miles, E.W.**, 1995: A quasi-experimental assessment of the effect of computerizing noncognitive paper-and-pencil measurements: a test of measurement equivalence, in: *Journal of Applied Psychology* 80, 643-651.
- Koch, Achim**, 1997: Teilnahmeverhalten beim Allbus 1994, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 49, 98-122.

- Maier, Jürgen**, 2000: Politisches Interesse und politisches Wissen in Ost- und Westdeutschland, in: **Falter, Jürgen W.; Gabriel, Oscar W.; Rattinger, Hans** (Hrsg.): *Wirklich ein Volk? Die politischen Orientierungen von Ost- und Westdeutschen im Vergleich*, Opladen: Leske und Budrich, 141-171.
- Narayan, Sowmya; Krosnick, Jon A.**, 1996: Education Moderates Some Response Effects in Attitude Measurement, in: *Public Opinion Quarterly* 60, 58-88.
- Nicholson, Thomas; White, John B.; Duncan, David**, 1998: Drugnet: A pilot study of adult recreational drug use via the WWW, in: *Substance Abuse* 19, 109-121.
- Perspektive Deutschland, 2003: *Kurzbericht 2003*, Berlin.
- Pötschke, Manuela; Simonson, Julia**, 2001: Online-Erhebungen in der empirischen Sozialforschung: Erfahrungen mit einer Umfrage unter Sozial-, Markt- und Meinungsforschern, in: *ZA-Information* 49, 6-28.
- Reuband, Karl-Heinz**, 1990: Interviews, die keine sind – ‚Erfolge‘ und ‚Mißerfolge‘ beim Fälschen von Interviews, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 42, 706-733.
- Reuband, Karl-Heinz**, 2000: ‚Pseudo-Opinions‘ in Bevölkerungsumfragen. Wie die Bürger fiktive Politiker beurteilen, in: *ZA-Information* 46, 26-38.
- Reuband, Karl-Heinz**, 2001: Politische Ignoranz und vorgetäushtes Wissen. Über die Bewertung von Politikern in allgemeinen Bevölkerungsumfragen, in: *Zeitschrift für Parlamentsfragen* 32, 812-821.
- Sassenberg, Kai; Kreutz, Stefan**, 1999: Online Research und Anonymität, in: **Batinic, Bernad; Werner, Andreas; Gräf, Lorenz; Bandilla, Wolfgang** (Hrsg.), *Online Research*, Göttingen: Hogrefe, 61-75.
- Schnell, Rainer**, 1991a: Wer ist das Volk? in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 43, 106-137.
- Schnell, Rainer**, 1991b: Der Einfluß gefälschter Interviews auf Survey-Ergebnisse; in: *Zeitschrift für Soziologie*, 20, 25-35.
- Schnell, Rainer**, 1997: *Nonresponse in Bevölkerungsumfragen*, Opladen: Leske und Budrich.
- Stanton, Jeffrey M.**, 1998: An empirical assessment of data collection using the Internet, in: *Personnel Psychology* 51, 709-726.
- Sudman, Seymour; Bradburn, Norman M.**, 1974: *Response Effects in Surveys*, Chicago: Aldine.
- Taylor, Humphrey**, 2000: Does internet research work? in: *International Journal of Market Research* 42, 51-63.
- Vehovar, Vasja; Batagelj, Zenel; Manfreda, Katja Lozar; Zaletel, Metka**, 2002: Nonresponse in Web Surveys, in: **Groves, Robert M.; Dillman, Don A.; Eltinge, John L.; Little, Roderick J.A.** (Hrsg.), *Survey Nonresponse*, New York: Wiley, 229-242.
- Vetter, Angelika**, 1997: *Political Efficacy – Reliabilität und Validität*, Wiesbaden: Deutscher Universitätsverlag.
- Zaller, John R.**, 1990: Political awareness, elite opinion leadership, and the mass survey response, in: *Social Cognition* 8, 125-153.

## Die angemessene Erfassung der Stärke und Richtung von Anreizen durch soziale Erwünschtheit<sup>1</sup>

von Volker Stocké und Christian Hunkler<sup>2</sup>

### *Zusammenfassung*

*Der Beitrag untersucht die soziale Erwünschtheit von vier häufig in Umfragen durch Selbstbeschreibungsfragen erfassten Merkmalen. Bei diesen handelt es sich um die Intensität des Fernseh- und Alkoholkonsums, die Häufigkeit von Blutspenden in der Vergangenheit und das Ausmaß des Lebensglücks der Befragten. Es wird die Gültigkeit von drei unterschiedlich differenzierten und daher verschieden aufwendigen Operationalisierungen für die Stärke und Richtung der von den Befragten wahrgenommenen Anreize zu sozial erwünschtem Antwortverhalten verglichen. Das Standardverfahren erfasst nur die gesellschaftliche Erwünschtheit einer starken Ausprägung des betreffenden Merkmals, die zweite Operationalisierung Erwünschtheitsdifferenzen einer starken und schwachen Merkmalsausprägung und das dritte Verfahren berücksichtigt zusätzlich die Bewertung einer Merkmalsausprägung mittlerer Stärke. Die Ergebnisse der empirischen Untersuchung mit einer lokalen Zufallsstichprobe zeigen erstens, dass ein substantieller Teil der Befragten bei den Themen „Alkohol-“ und „Fernsehkonsum“ umgekehrt u-förmige Erwünschtheitsverläufe über das Merkmalskontinuum wahrnimmt. Bei diesen Themen ist ausschließlich der dritte, differenzierteste Indikator eine gültige Operationalisierung des Verzerrungspotentials durch soziale Erwünschtheit. Das zweite Ergebnis ist, dass bei den Themen „Blutspende“ und „Lebensglück“ mit ausschließlich monotonen Erwünschtheitsverläufen vor allem der mit globalen Differenzwerten verbundene Mehraufwand zu einer starken Verbesserung der Validität der prognosti-*

---

1 Wir danken ganz herzlich **Hartmut Esser** und den Gutachtern dieser Zeitschrift für hilfreiche Kommentare und Anregungen. **Diana Schirowski** und **Jason Rosow** waren eine große Unterstützung bei der Erstellung des Manuskriptes. Der Beitrag wurde durch finanzielle Förderung der Deutschen Forschungsgemeinschaft an den Sonderforschungsbereich 504 der Universität Mannheim ermöglicht.

2 Dr. **Volker Stocké** ist Hochschulassistent und **Christian Hunkler** Mitarbeiter am Sonderforschungsbereich 504 „Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung“ der Universität Mannheim, L13, 15; D-68131 Mannheim.  
E-Mail: vstocke@rumms.uni-mannheim.de.

zierten Anreizstärke führt. Das in der derzeitigen Forschung annähernd ausschließlich verwendete Standardverfahren hat sich dagegen bei keinem der untersuchten Befragungsthemen als angemessener Indikator der Stärke und Richtung des Verzerrungspotentials durch soziale Erwünschtheit erwiesen.

### **Abstract**

*The study investigates the social desirability of four characteristics often asked as self-descriptive questions in surveys. These are the intensity of watching television, the consumption of alcohol, the frequency of past blood donations and the degree of happiness in life. We compare the validity of three differently elaborated operationalizations of the magnitude and direction of the perceived incentives for social desirable response behaviour. The standard method only accounts for the social desirability of a strong occurrence of the respective attribute, however, the second procedure takes differences between strong and weak occurrences into account, and the third implementation additionally considers the evaluation of a medium occurrence. Using data from a local survey based on a random probability sample, we first find a substantial proportion of the respondents reporting u-shaped desirability profiles for the consumption of alcohol and intensity of watching television. For these topics only the third and most sophisticated operationalisation is an appropriate measure of the potential distortions caused by social desirability. Secondly, monotonic desirability profiles are observed for past blood donations and happiness in life. Here, the additional effort necessary for the global difference scores as the second most elaborated measure improves considerably the magnitude of incentives from social desirability which can be identified. Contemporary research has almost exclusively applied the standard method; however, our investigation suggests that this is under no conditions an adequate measure of the strength and the direction of potential social desirability bias.*

## **1 Einleitung**

Antwortverzerrungen durch soziale Erwünschtheit sind eine erhebliche Bedrohung der Gültigkeit von Umfragedaten. Belege für die Wirksamkeit dieser Art von Response-Effects wurden beispielsweise in Umfragen über die Kirchgangshäufigkeit und die psychische Gesundheit der Befragten sowie über deren Drogen- und Alkoholkonsum gefunden (*Aquilino* 1998; *Epstein* et al. 2001; *Jobe* et al. 1997; *Presser* und *Stinson* 1998). Da es sich beim „social desirability bias“ (SD-Bias) um systematische Antwortverzerrungen in eine bestimmte Richtung handelt, muss als Ergebnis mit verzerrten Randverteilungen der abgefragten Einstellungen und fakti-

schen Merkmale der Befragten gerechnet werden. Diese Antwortverzerrungen können dann durch unterschiedliche statistische Verfahren nachträglich korrigiert werden, wenn ihre Stärke und inhaltliche Ausrichtung möglichst genau und verlässlich vorhergesagt werden können.

Im Rahmen der Rational-Choice-Theorie des Befragtenverhaltens, aber auch in anderen Ansätzen, die den SD-Bias als das Ergebnis einer an die soziale Umwelt gerichteten „Impression Management“-Strategie erklären, werden die Erwünschtheitswahrnehmungen der Befragten als eine zentrale Entstehungsbedingung des SD-Bias angesehen (*Esser* 1991, *Stocké* 2004). Demnach verwenden die Befragten ihre Wahrnehmungen über die Erwünschtheit oder „Trait Desirability“ (TD) eines Merkmals als Information darüber, welche Wertungen und Sanktionen durch die soziale Umwelt im Interview bei einer Selbstzuschreibung dieses Merkmals durch ihr Antwortverhalten durchschnittlich erwartet werden müssen. Die Erwünschtheitswahrnehmungen im konkreten Befragungskontext ergeben sich einerseits durch bewertende Reaktionen der Interviewer im vorangegangenen Teil der Befragung und durch auf der Grundlage von Interviewermerkmalen antizipierten Erwartungshaltungen. Andererseits verwenden die Befragten ihr Wissen über gesellschaftliche Normen, um Erwartungen über die Konsequenzen ihres Antwortverhaltens zu generieren. Aus der Perspektive der Rational-Choice-Theorie streben Befragte durch ihre Antworten Verhaltensbestätigung, die Vermeidung negativer Sanktionen und insgesamt soziale Anerkennung an. Die Erwünschtheitswahrnehmungen sind eine zentrale *kognitive* Grundlage für die von den Befragten angenommene Eignung verschiedener Antwortalternativen zur Realisierung dieser Ziele. Eine valide Erfassung dieser Wahrnehmungen ist damit eine notwendige, aber nicht hinreichende Vorbedingung für eine genaue Prognose des SD-Bias.

Aus der theoretischen Logik der Rational-Choice-Theorie prognostizieren die Erwünschtheitswahrnehmungen die Stärke und Richtung von Antwortverzerrungen durch soziale Erwünschtheit nur dann, wenn die Antworten durch anwesende Personen wahrgenommen werden können und außerdem ein hinreichend starkes Bedürfnis nach sozialer Anerkennung vorliegt (*Stocké* 2004). Es wird also von einer nichtkompensatorischen, multiplikativen Verknüpfung dieser drei Entstehungsbedingungen ausgegangen. Liegen keine externen Validierungsdaten vor, so kann die Erklärungskraft dieser Entstehungsbedingungen des SD-Bias für das Antwortverhalten und insbesondere das prognostizierte interaktive Zusammenspiel dieser Bedingungen als Beleg für die Existenz von Antwortverzerrungen herangezogen werden.

Eine besonders schwerwiegende Beeinträchtigung der Verlässlichkeit von Umfragedaten liegt dann vor, wenn sich der SD-Bias verschieden stark oder gar in unter-

schiedlicher Richtung auf das Antwortverhalten durch bestimmte Merkmale definierter Teilgruppen der Befragten auswirkt. Unter diesen Umständen muss damit gerechnet werden, dass beobachtete Zusammenhänge zwischen dem betreffenden Erhebungsmerkmal und der entsprechenden Gruppenzugehörigkeit reine Artefakte der (unterschiedlichen) Wirksamkeit sozialer Erwünschtheit sind. Eine mögliche Ursache für eine gruppenspezifische Differenzierung in der Stärke und Richtung des SD-Bias ist, dass sich die Erwünschtheitswahrnehmungen entlang der Gruppengrenzen unterscheiden. Der Grund für solche Unterschiede kann in der Einbettung in unterschiedliche normative Bezugsgruppen und deren Prägekraft für die als gültig angesehenen Bewertungskriterien gesehen werden (*Singer* 1981). So kann etwa das empirische Ergebnis, wonach Männer in Umfragen bei der Angabe der Anzahl ihrer Geschlechtspartner übertreiben, während Frauen in dieser Hinsicht zu geringe Angaben machen, als Hinweis auf die Existenz geschlechtsspezifischer Normen über das angemessene Ausmaß sexueller Promiskuität gewertet werden (*Smith* 1992). Ob diese und andere Unterschiede in der Richtung des SD-Bias tatsächlich auf Gruppenunterschiede in der Erwünschtheitswahrnehmung zurückgehen, wurde jedoch bisher kaum untersucht. Eine solche Analyse erfordert ein gültiges und zuverlässiges Instrument zur Erfassung der von den Befragten subjektiv wahrgenommenen Erwünschtheit des betreffenden Merkmals.

Bei der bisherigen Erforschung der sozialen Erwünschtheit unterschiedlicher, in Umfragen durch Selbstbeschreibungsfragen erfasster Merkmale wird in der Regel die antizipierte Bewertung der Anwesenheit des betreffenden Merkmals erfasst. So haben beispielsweise die Befragten in der Untersuchung von *Wyner* (1980) angegeben, für wie sozial erwünscht oder unerwünscht sie verschiedene Straftaten, wie etwa Raub oder Einbruch, ansehen. Diese Vorgehensweise bezieht sich bei Merkmalen mit einer kontinuierlich variierenden Ausprägungsstärke – wie etwa dem Ausmaß des Alkoholkonsums – auf die wahrgenommene gesellschaftliche Bewertung einer starken Merkmalsausprägung. Diese Standardoperationalisierung des TD-Konstruktes setzt implizit voraus, dass eine schwache bzw. fehlende Ausprägung des untersuchten Merkmals von den Informanten als neutral bewertet angesehen wird (vgl. hierzu noch Abschnitt 2.1). Diese Annahme wurde in der bisherigen Forschung aber kaum empirisch geprüft. Zu diesem Zweck ist es notwendig, auch die wahrgenommene Erwünschtheit der Abwesenheit bzw. einer schwachen Ausprägung des betreffenden Merkmals zu erfassen. Bei dieser differenzierteren Operationalisierung des TD-Konstruktes wird bei Merkmalen mit kontinuierlich variierender Ausprägungsstärke eine weitere und unserem Wissen nach bisher völlig ungetestete Annahme gemacht. Diese Annahme besteht darin, dass die Erwünschtheitswahrnehmungen der Befragten mit zunehmender Intensität der Merkmalsaus-

prägung einen streng monoton ansteigenden oder abnehmenden Verlauf nimmt. Im Widerspruch hierzu sind jedoch soziale Normen denkbar, wonach einer mittleren Merkmalsstärke oder gar den beiden extremen Merkmalsausprägungen das höchste Ausmaß an sozialer Erwünschtheit zukommt. Unter diesen Umständen ist weder die Standardoperationalisierung noch die differenzierte Erfassung der TD mit zwei Messpunkten eine angemessene Repräsentation der Stärke und Richtung der Anreize durch soziale Erwünschtheit. Ob dies der Fall ist, kann nur durch eine noch weiter differenzierte Erfassung der Erwünschtheitswahrnehmungen für eine schwache, mittlere und starke Merkmalsausprägung überprüft werden.

Der vorliegende Beitrag analysiert bei vier häufig in Umfragen erfassten Merkmalen, wie differenziert bei einer angemessenen Erfassung der jeweils vorliegenden Stärke und Richtung der Anreize zu sozial erwünschtem Antwortverhalten vorgegangen werden muss. Die untersuchten Merkmale sind das Ausmaß des Lebensglücks und die Häufigkeit von Blutspenden der Befragten, sowie die Intensität ihres Alkohol- und Fernsehkonsums. Es wird analysiert, ob und bei welchen dieser Befragungsthemen den drei möglichen Operationalisierungen des TD-Konzeptes konvergente Validität zukommt.

## 2 Forschungsstand

Im folgenden Abschnitt werden zuerst die unterschiedlichen Versionen des TD-Konzeptes sowie verschiedene Möglichkeiten ihrer Operationalisierung genauer diskutiert. Im zweiten Abschnitt wird dann die verfügbare Evidenz für die Prognosekraft dieser Indikatoren der wahrgenommenen sozialen Erwünschtheit für die Entstehung des SD-Bias dargestellt. Der dritte Abschnitt beschreibt den Kenntnisstand über die Stärke und Richtung des zu erwartenden SD-Bias bei den im vorliegenden Beitrag untersuchten Befragungsthemen.

### 2.1 Operationalisierungen der Trait-Desirability

Die soziale Erwünschtheit eines Merkmals und die damit verbundene Sanktionserwartung der Befragten kann entweder in der Form „persönlicher“ oder „extrinsisch evaluierter“ Bewertungen operationalisiert werden. Im ersten Fall geben die Informanten an, für wie erwünscht sie persönlich bestimmte Eigenschaften oder Verhaltensweisen einer Person ansehen (vgl. beispielsweise: *Meleddu* und *Guicciardi* 1998; *Smith* und *Smith* 1993). Die aggregierten Angaben der Informanten über ihre eigenen Erwünschtheitsurteile können dann als Indikator für das normative Klima in der befragten Stichprobe angesehen werden. Dieser „objektive“ Indikator für die

von einem Befragten bei der Selbstzuschreibung des betreffenden Merkmals zu erwartenden positiven oder negativen Sanktionen ist dann valide, wenn die so erfassten Urteile mit den von den Befragten subjektiv erwarteten Bewertungsreaktionen übereinstimmen. Dies ist empirisch nachweisbar nicht immer der Fall (**Crott** und **Roßrucker** 1974). Die zweite mögliche Art der Operationalisierung des TD-Konstruktes, die „extrinsisch“ evaluierte oder fremdeingeschätzte soziale Erwünschtheit, zielt direkt auf die von den Befragten subjektiv wahrgenommenen, in der Gesellschaft dominanten Bewertungsstandards (vgl. beispielsweise: **Crott** und **Roßrucker** 1974; **Edwards** 1957; **Nicotera** 1996). Die Befragten werden hierbei um ihre Einschätzung darüber gebeten, welche Bewertungen oder Reaktionen ein bestimmtes Merkmal typischerweise hervorruft.

Bei der empirischen Erfassung beider Varianten der Operationalisierung des TD-Konstruktes sind weiterhin unterschiedlich differenzierte und damit verschieden aufwendige Vorgehensweisen denkbar. Bei der *ersten* und in der derzeitigen Forschung fast ausschließlich angewandten Standardmethode wird die von den Befragten wahrgenommene Erwünschtheit des Vorliegens oder, im Falle kontinuierlich variierender Eigenschaften, einer starken Ausprägung des untersuchten Merkmals erfasst. Es werden typischerweise bipolare Antwortskalen herangezogen, deren Mittelpunkt eine neutrale Erwünschtheitswahrnehmung repräsentiert. Das Ausmaß und die Richtung der Abweichung der antizipierten Merkmalsbewertung vom neutralen Mittelpunkt der Skala wird als Indikator für die bestehenden Anreize zur Selbstzuschreibung bzw. Leugnung einer starken Merkmalsausprägung angesehen. Aus der Perspektive der Rational-Choice-Theorie ist diese Vorgehensweise nur dann angemessen, wenn die schwache Merkmalsausprägung von den Befragten als neutral bewertet angesehen wird. Dies lässt sich aus der Annahme begründen, dass die Befragten die Differenz in der Bewertung der extremen Pole des Merkmalskontinuums als Grundlage für ihre Wahrnehmungen der Anreize durch soziale Erwünschtheit heranziehen (**Stocké** 2004). Einer positiven oder negativen Erwünschtheitseinschätzung einer starken Merkmalsausprägung alleine kommt demnach nur dann Konstruktvalidität zu, wenn diese Einschätzung mit der Differenz der bewerteten Erwünschtheit einer starken und schwachen Merkmalsausprägung übereinstimmt. Ist dies wegen der Verletzung der Neutralitätsannahme nicht erfüllt, so repräsentiert die Standardoperationalisierung eine Über- oder Unterschätzung der Stärke und im schwerwiegendsten Fall eine Fehleinschätzung der Richtung der tatsächlich vorliegenden Anreize durch soziale Erwünschtheit. Eine Unterschätzung liegt dann vor, wenn die Erwünschtheitswahrnehmung einer starken und schwachen Merkmalsausprägung in unterschiedlicher Richtung von einer neutralen Bewertung abweicht: die Erwünschtheitsunterschiede zwischen den beiden Antwortmöglichkeiten sind dann

in Realität größer als die bei einer starken Merkmalsausprägung erfassten Abweichungen vom neutralen Mittelpunkt der Skala. Weicht dagegen die soziale Erwünschtheit der schwachen Merkmalsausprägung in die gleiche Richtung wie die einer starken Ausprägung von einer neutralen Bewertung ab, so führt die Standardmethode entweder zu einer Überschätzung des Verzerrungspotentials durch soziale Erwünschtheit oder zu falschen Schlussfolgerungen über die Richtung des Verzerrungspotentials durch den SD-Bias. Eine Überschätzung liegt dann vor, wenn eine schwache Merkmalsausprägung weniger erwünscht bzw. unerwünscht als eine starke Ausprägung eingeschätzt wird: die realen Erwünschtheitsunterschiede zwischen der starken und schwachen Merkmalsausprägung sind dann geringer als die bei einer starken Ausprägung beobachtete Abweichung vom neutralen Skalenmittenpunkt. Wird die schwache Merkmalsausprägung dagegen sogar als erwünschter bzw. unerwünschter angesehen als dies bei der erfassten Bewertung einer starken Ausprägung der Fall ist, so nehmen die Befragten den Implikationen der Standardoperationalisierung entgegengesetzte Anreize zu sozial erwünschtem Antwortverhalten wahr.

Bei der *zweiten*, in der empirischen Forschung bisher kaum angewandten Vorgehensweise wird die Annahme aufgegeben, dass eine schwache Merkmalsausprägung notwendigerweise neutral bewertet wird. Bei dieser differenzierteren Operationalisierung des TD-Konstruktes wird neben der Bewertung einer starken auch die einer schwachen Merkmalsausprägung explizit erfasst und die Differenz zwischen beiden Erwünschtheitswahrnehmungen als Indikator für die Stärke und Richtung der Anreize zum SD-Bias herangezogen (als Beispiel vgl. *Stocké* 2004). Bei dieser Operationalisierung mit Differenzwerten wird, ebenso wie bei der Standardoperationalisierung des TD-Konstruktes, vorausgesetzt, dass die Erwünschtheitswahrnehmungen mit steigender Ausprägungsstärke der Merkmale monoton zunehmen oder abnehmen. Ist diese Annahme nicht erfüllt, so führen die damit verbundenen u-förmigen oder umgekehrt u-förmigen Erwünschtheitsprofile nicht zu Antwortverzerrungen in eine bestimmte Richtung, sondern zu einer Konzentrierung der Angaben entweder am Mittelpunkt oder an beiden Extrempolen des Merkmalskontinuums. Derartige nicht-monotone Verläufe der sozialen Erwünschtheit über das Merkmalskontinuum lassen sich nur durch die zusätzliche Erfassung der antizipierten Bewertung einer mittelstarken Merkmalsausprägung identifizieren. Da uns keine empirische Anwendung dieser *dritten*, bereichsspezifischen und daher differenzier-testen Operationalisierung des TD-Konstruktes bekannt ist, ist derzeit unklar, ob der hierbei zusätzlich notwendige empirische Aufwand gerechtfertigt ist.

Insgesamt kann mit einer vollständigen konvergenten Validität der drei Operationalisierungen nur dann gerechnet werden, wenn sowohl die Neutralitätsannahme als

auch die Monotonieannahme bei allen Befragungsthemen empirisch erfüllt sind. Ist die Neutralitätsannahme verletzt, so kommt der Differenz- sowie der bereichsspezifischen Operationalisierung weiterhin konvergente Validität zu, die Standardoperationalisierung führt dann allerdings zu einer fehlerhaften Prognose der vorliegenden Stärke und eventuell auch der Richtung von Anreizen zu sozial erwünschtem Antwortverhalten. Liegen dagegen bei einem relevanten Teil der Befragten nicht-monotone Erwünschtheitsverläufe vor, so sind weder die Standardoperationalisierung noch die Verwendung von globalen Differenzwerten eine angemessene Erfassung der von den Befragten wahrgenommenen Anreize durch soziale Erwünschtheit. Bei gleichzeitiger Verletzung der Neutralitäts- als auch der Monotonieannahme liegt somit das geringste Ausmaß an konvergenter Validität der drei Messkonzepte vor. Für die vier in dieser Arbeit analysierten Befragungsthemen wurde unseres Wissens nach bisher keine der beiden Annahmen getestet.

Als aggregierte Maßzahl für den bei einem bestimmten Befragungsthema drohenden SD-Bias kann bei der Standardoperationalisierung des TD-Konstruktes und bei der Verwendung globaler Differenzwerte der Durchschnitt der individuellen Skalenwerte der Informanten herangezogen werden. Bei der Verwendung bereichsspezifischer Differenzwerte ist diese Maßzahl durch die Summe der aggregierten Erwünschtheitsunterschiede zwischen einer starken und mittleren sowie mittleren und schwachen Merkmalsausprägung definiert. Bei allen drei Operationalisierungen gibt der Betrag des resultierenden Indikators die Stärke und das Vorzeichen die Richtung des auf der Aggregatebene erwartbaren SD-Bias an. Eine auf der Individual-ebene der Befragten möglicherweise vorliegende Heterogenität hinsichtlich der Richtung wahrgenommener Anreize durch soziale Erwünschtheit bleibt hierbei jedoch unberücksichtigt. Da sich solche Unterschiede in der Richtung der Anreizwahrnehmungen auf der Aggregatebene gegenseitig aufheben, kommt es unter diesen Umständen zu einer Unterschätzung des Einflusspotentials durch SD-Bias. Obwohl eine solche Unterschätzung für die beobachtete Randverteilung der untersuchten Merkmale ohne Konsequenzen ist, können Unterschiede in der Richtung von Antwortverzerrungen auf der Ebene von Subgruppen der Befragten zu Abweichungen der beobachteten von den Zusammenhängen zwischen dem untersuchten Merkmal und den relevanten Gruppierungsmerkmalen bewirken. Ob eine solche Heterogenität hinsichtlich der Richtung der Anreizwahrnehmungen vorliegt, kann bei jeder der drei Operationalisierungen des SD-Konstruktes durch eine nach dem Vorzeichen der individuellen Anreizwahrnehmungen differenzierten Analyse kontrolliert werden.

## 2.2 Bedeutung der Erwünschtheitswahrnehmungen für die Erklärung des SD-Bias

In einer Reihe von Studien wurde die soziale Erwünschtheit unterschiedlicher Merkmale mit der Absicht einer Vorhersage von Antwortverzerrungen durch soziale Erwünschtheit erfasst. So haben vorliegende Studien mit britischen und australischen Studenten gezeigt, dass apathische oder arrogante Personen als negativ, Loyalität und Vorurteilslosigkeit dagegen als sehr positiv angesehen werden (*Bochner* und *van Zyl* 1984; *Hampson* et al. 1987). In beiden Studien wurde nur die Bewertung des Vorliegens der jeweiligen Eigenschaft auf bipolaren Skalen abgefragt. In der Untersuchung von *Hampson* und Kollegen wurde erfasst, wie erwünscht die betreffende Eigenschaft für eine Person im allgemeinen ist, während in der Studie von *Bochner* und *van Zyl* gefragt wurde, für wie wünschbar die Informanten die analysierte Eigenschaft bei Familienmitgliedern, Freunden und Bekannten halten. Bei beiden Studien wird also die extrinsisch evaluierte Erwünschtheit mithilfe der Standardoperationalisierung eines Messpunktes erfasst.

Weitere Forschungsarbeiten zeigen, dass sich die soziale Erwünschtheit bei einigen Befragungsthemen deutlich zwischen unterschiedlichen Befragtengruppen unterscheidet. *Mecklenbräuker* und Kollegen (1994) ließen die Eigenschaften „listig“ und „willig“ auf einer bipolaren Skala mit neutralem Mittelpunkt von einer Stichprobe deutscher Studenten bewerten. Dabei unterschieden sie zwischen der Einschätzung der Wünschbarkeit bei einem gleichaltrigen Mann bzw. bei einer gleichaltrigen Frau. Zusätzlich wurden die Bewertungen nach dem Geschlecht der Informanten differenziert berichtet. Es zeigte sich, dass die Eigenschaft „listig“ bei Personen jeglichen Geschlechts von Männern als erwünschter eingeschätzt wird als von Frauen. Frauen bewerten das Merkmal „willig“ bei einer anderen Frau als negativ, bei einem gleichaltrigen Mann aber tendenziell als positiv. Auch in dieser Untersuchung wurde jeweils nur eine starke Ausprägung der jeweiligen Eigenschaft bewertet.

Dass sich die Erwünschtheitswahrnehmungen auch auf das Antwortverhalten der Befragten auswirken, wurde für eine Reihe von Befragungsthemen überprüft. Es konnte beispielsweise gezeigt werden, dass sich Befragte zunehmend durch ihr Antwortverhalten als argumentative Persönlichkeit charakterisieren, wenn sie diese Eigenschaft gesellschaftlich immer mehr als erwünscht einschätzen (*Nicotera* 1996). Die Teilnehmer der Studie haben die Items der „*Infante* and *Rancer* Argumentativeness Scale“ auf einer fünf-stufigen Antwortskala bewertet, deren Endpunkte mit „nicht gesellschaftlich akzeptiert“ und „gesellschaftlich akzeptiert“ verankert waren. Die resultierenden fremdeingeschätzten Erwünschtheitswahr-

nehmungen haben sich als signifikant mit dem Antwortverhalten der Befragten auf der „Argumentativeness Scale“ korreliert erwiesen ( $r = .26$ ).

Ähnliche Zusammenhänge zwischen den Erwünschtheitswahrnehmungen und dem Antwortverhalten finden sich in Befragungen über das Vorliegen von Symptomen psychischer Störungen und bei der Erfassung einer großen Anzahl unterschiedlicher Persönlichkeitseigenschaften. In allen diesen Studien wird ebenfalls die Standardoperationalisierung des TD-Konstruktes herangezogen und somit nur das Vorliegen des jeweils analysierten Symptoms bewertet. *Phillips* und *Clancy* (1970) sowie *Edwards* (1957) lassen die Wünschbarkeit der Symptome auf einer unipolaren neun-stufigen Skala einschätzen: je höher die Erwünschtheit, desto höher der Skalenwert. *Phillips* und *Clancy* (1970) haben einen signifikant positiven Zusammenhang zwischen der Erwünschtheitseinstufung von Symptomen psychischer Gesundheit und der Anzahl berichteter Symptome des „22-Item Mental Health“-Indexes festgestellt: je höher die geschätzte Erwünschtheit, desto mehr Symptome werden berichtet. *Gove* und *Geerken* (1977) verwenden eine bipolare, durchgängig verbal verankerte Antwortskala, die zwischen extrem, ziemlich und etwas (un-) erwünscht differenziert. Auch hier wurde ein positiver Zusammenhang zwischen den Erwünschtheitsurteilen und der Anzahl der berichteten psychiatrischen Symptome gefunden. *Huang* und Kollegen (1998) finden vergleichbare Ergebnisse über den Zusammenhang zwischen den Erwünschtheitswahrnehmungen der Befragten und deren Antwortverhalten. In einer weiteren Studie kann gezeigt werden, dass nicht nur ein Zusammenhang zwischen den Erwünschtheitswahrnehmungen und dem Antwortverhalten besteht, sondern dass diese tatsächlich Antwortverzerrungen vorhersagen. So wurde durch Abgleich der Befragtenangaben über die Anzahl ihrer bisherigen Verhaftungen mit Informationen aus Polizeiakten gezeigt, dass die Umfrageteilnehmer bei der Angabe der Häufigkeit dieser Ereignisse in der Vergangenheit teilweise über- und teilweise untertrieben haben (*Wyner* 1980). Die Stärke und vor allem die Richtung dieser Antwortverzerrungen konnten durch die Unterschiede der von den Befragten wahrgenommenen gesellschaftlichen Bewertung des untersuchten Merkmals vorhergesagt werden.

Im Rahmen der Theorie rationalen Befragtenverhaltens wird vorhergesagt, dass sich die Erwünschtheitswahrnehmungen in Interaktion mit dem Bedürfnis nach sozialer Anerkennung und der Privatheit der Antwortsituation auf die Angaben der Befragten auswirken. Diese Annahme wurde in einer Studie über die Affektivität der Befragten überprüft (*Chen* et al. 1997). Hierbei wurde ein signifikanter Interaktionseffekt zwischen der wahrgenommenen Erwünschtheit positiver und negativer Affektivität und dem durch die *Marlowe-Crowne* SD-Skala operationalisierten Anerkennungsbedürfnis der Befragten gefunden: Die Angaben der Probanden über ihre ei-

gene Affektivität erwies sich zunehmend stärker in Übereinstimmung mit dem, was sie in der Gesellschaft als anerkannt ansahen, wenn es sich um Befragte mit starkem statt mit schwachem Anerkennungsbedürfnis handelte. Hierbei wurde ebenfalls die Standardoperationalisierung des TD-Konstruktes herangezogen. Diese Interaktion zwischen den mit der Standardvorgehensweise operationalisierten Erwünschtheitswahrnehmungen und der Stärke des Anerkennungsbedürfnisses konnte dagegen in einer anderen Studie mit Fragen etwa aus dem Bereich der Ehezufriedenheit und über die ethnischen Vorurteile der Befragten nicht festgestellt werden: Beide Faktoren wirken sich in der von *Phillips* und *Clancy* (1972) durchgeführten Untersuchung unabhängig voneinander auf das Antwortverhalten aus. In einer Studie von *Stocké* (2004) wurden die Erwünschtheitswahrnehmungen der Befragten über verschiedene Einstellungen zu Ausländern mit globalen Differenzwerten operationalisiert. Dabei wurde die extrinsisch evaluierte Erwünschtheit der Zustimmung und Ablehnung der insgesamt 10 untersuchten Einstellungsfragen erfasst. Die Ergebnisse haben gezeigt, dass sich die Erwünschtheitswahrnehmungen der Befragten in interviewer-administrierten Befragungen signifikant stärker auf das Antwortverhalten auswirken, verglichen mit der Bedingung einer sichergestellten Privatheit in selbst-administrierten Interviews. Außerdem hat sich dieser Zusammenhang für Befragte mit hohem Anerkennungsbedürfnis als ebenfalls signifikant stärker erwiesen. Die teilweise widersprüchlichen Ergebnisse sind möglicherweise das Ergebnis der unterschiedlich differenzierten Operationalisierung der Erwünschtheitswahrnehmungen in den vorliegenden Studien und der je nach untersuchtem Befragungsthema verschieden starken Verletzung der Neutralitäts- und Monotonieannahme.

### **2.3 Antwortverzerrungen durch soziale Erwünschtheit bei den untersuchten Befragungsthemen**

Die Ergebnisse einiger Studien erbringen Belege für die Existenz, Stärke und Richtung von Antwortverzerrungen durch soziale Erwünschtheit bei den Themen „Alkoholkonsum“ und „Lebensglück“. Dagegen sind uns für das Ausmaß des Fernsehkonsums und der Häufigkeit des Blutspendens keine Untersuchungen über den hier zu erwartenden SD-Bias bekannt. In einer Studie von *Blair* und Kollegen (1977) wurde der auf der Grundlage einer für die Vereinigten Staaten bevölkerungsrepräsentativen Befragung berechnete Alkoholkonsum mit Daten über den landesweiten Alkoholverkauf verglichen. Nach den Ergebnissen der Studie wird der tatsächliche Alkoholkonsum in der Bevölkerung durch Umfragedaten deutlich unterschätzt. Die Autoren ziehen die Schlussfolgerung, dass Befragungsdaten durch Einflüsse sozialer Erwünschtheit in Richtung auf zu geringe Angaben über den Alkoholkonsum verzerrt sind. In einer anderen Studie wurden Soldaten in einer isolierten militäri-

schen Polarstation über das Ausmaß ihres Alkoholkonsums befragt (*Embree* und *Whitehead* 1991). Der auf diesen Angaben beruhende Gesamtkonsum konnte mit den offiziellen Verkaufsdaten der Basis verglichen werden. Je nach Art der Frageformulierung in der Studie wurden auch der Einfluss unterschiedlicher Referenzperioden und der Einfluss einer Vignettenteknik überprüft. Haben die Befragten einen täglichen Alkoholkonsum von .96 bis 2.66 Getränken angegeben. Bei einem tatsächlichen Konsum von 3.36 Getränken täglich, muss von einer Antwortverzerrung in Richtung auf zu geringe Angaben ausgegangen werden.

Bei der Untersuchung der Stärke und Richtung des SD-Bias wird häufig, mangels objektiver Validierungsdaten, auf indirekte Evidenz zurückgegriffen. Hierbei wird geprüft, ob und in welchem Ausmaß sich theoretisch prognostizierte Determinanten des SD-Bias auf die Wahrscheinlichkeit einer Selbstzuschreibung der untersuchten Merkmale auswirken. So wird in einigen Studien die Privatheit der Antwortsituation experimentell variiert und aus den beobachteten Antwortunterschieden Rückschlüsse auf das Ausmaß und die Richtung des SD-Bias gezogen. Da Antwortverzerrungen durch das Bedürfnis nach positiver Selbstdarstellung und Verhaltensbestätigung durch die soziale Umwelt („other deception“) nur bei Wahrnehmbarkeit und daher Sanktionierbarkeit der Angaben erwartet werden, können die Angaben in privaten Interviewsituationen als eine bessere Annäherung an die „wahren“ Werte der Befragten angesehen werden.<sup>3</sup>

In einer Untersuchung wurden unterschiedliche Aspekte des Alkoholgenusses, wie etwa das Ausmaß des regelmäßigen Konsums oder ob die Befragten jemals ein „Alkoholproblem“ hatten, entweder durch interviewer- oder selbst-administrierte Befragungen erhoben (*Mangione* et al. 1982). Bei den insgesamt 14 untersuchten Aspekten des Alkoholkonsums hat sich nur bei zwei Fragestellungen die Privatheit der Antwortsituation signifikant auf die Randverteilungen der Angaben ausgewirkt. Ähnlich geringe Einflüsse der mit dem Erhebungsmodus verbundenen Privatheit finden sich in einer Studie von *Aquilino* (1994), in der der Alkoholkonsum für verschiedene Referenzperioden und die Häufigkeit der Trunkenheit im vergangenen Jahr durch Verwendung selbst-administrierter Fragebögen, interviewer-administrierter persönlicher Befragungen oder durch Telefoninterviews erfasst wurden. Der Vergleich des Antwortverhaltens zwischen den Modi hat gezeigt, dass in selbst-administrierten Interviews der höchste Alkoholkonsum im vergangenen Monat und am häufigsten eine Trunkenheit von mehr als einmal die Woche angegeben wurden.

---

3 Für eine Diskussion der Unterscheidung zwischen der extrinsischen, auf die soziale Umwelt gerichteten Motivation zur Fremdtäuschung („other deception“) und der intrinsisch motivierten Tendenz zur positiven Selbsttäuschung („self deception“), vergleiche *Paulhus* und *Reid* 1991.

Insgesamt hat die Antwortprivatheit allerdings in der Gesamtstichprobe nur einen schwachen Effekt, so dass die Autoren nur von einem schwachen SD-Bias in Richtung einer Unterschätzung des Alkoholkonsums ausgehen. Die gleiche Studie findet auch, dass die Angaben von Afroamerikanern, verglichen mit dem Rest der Stichprobe, signifikant stärker durch die Privatheitsbedingungen beeinflusst wurden. In einer anderen Studie mit einer bevölkerungsrepräsentativen Stichprobe wurden ebenfalls die Angaben über das Ausmaß des Alkoholkonsums unter der Bedingung interviewer- und selbst-administrierter Interviews verglichen (*Turner* et al. 1992). Hier fanden sich ebenfalls nur geringe und statistisch nicht signifikante Unterschiede nach der Privatheit der Antwortsituation. In einer weiteren Studie vergleichen *Turner* und Kollegen (1998) mit einer bevölkerungsrepräsentativen Stichprobe die Angaben darüber, ob die Befragten Alkohol trinken oder nicht, bei einer selbst-administrierten Audio-CASI Prozedur (computer-administrierte Interviews, bei denen der aufgezeichnete Fragetext über Kopfhörer vorgespielt wird) und einem selbst-administrierten Papierfragebogen. Der anwesende Interviewer kann zwar im Nachhinein den Papierfragebogen einsehen, nicht aber die in den Laptop eingegebenen Angaben. Somit wird hier nur die Anonymität der Antworten variiert. Es finden sich für den Alkoholkonsum im letzten Jahr keine und für den wöchentlichen Alkoholkonsum nur marginal signifikante Effekte der Experimentalbedingungen. Die Studie von *Locke* und *Gilbert* (1995) hat untersucht, ob sich die Angaben einer Stichprobe von Studierenden über ihren Alkoholkonsum, erfasst durch den „Drinking Habits Questionnaire“, dann unterscheidet, wenn die Daten entweder durch selbst-administrierte Papierfragebögen, selbst-administrierte computergestützte Interviews oder mit einer mündlich-persönlichen Befragung erfasst wurden. Es hat sich gezeigt, dass sich der selbstberichtete Alkoholkonsum nicht nach dem Modus der Befragung unterscheidet. In der Studie von *Waterton* und *Duffy* (1984) wurden im Gegensatz zu den vorher dargestellten Studien in selbst-administrierten Interviews signifikant höhere Angaben über den Konsum von Bier, Wein und Spirituosen beobachtet.

Die Privatheit der Antwortsituation unterscheidet sich auch zwischen Interviews unter Verwendung des so genannten „Randomized Response“ (RR)-Verfahrens und konventionellen interviewer-administrierten Befragungen. Bei diesem Verfahren kodieren die Befragten ihre Antworten nach einem vorgegebenen Zufallsprozess, so dass sich deren Gesamtverteilung zwar nachträglich ermitteln lässt, die individuellen Antwortinhalte durch die Interviewer jedoch nicht wahrnehmbar sind. In einer Meta-Analyse berichten *Umesh* und *Peterson* (1991) von drei Untersuchungen, in denen die im Rahmen des RR-Verfahrens rekonstruierten Antwortverteilungen über den Alkoholkonsum der Befragten signifikant höher liegen, verglichen mit jenen,

die bei einer öffentlichen interviewer-administrierten Befragungsart beobachtet wurden.

Ein anderer, häufig verwendeter Indikator für die Richtung und Stärke des vorliegenden SD-Bias ist die Korrelation zwischen dem mit unterschiedlichen „social desirability“- oder kurz SD-Skalen erfassten Bedürfnis der Befragten nach sozialer Anerkennung und dem Antwortverhalten beim untersuchten Befragungsgegenstand: je stärker sich Befragte mit zunehmendem Anerkennungsbedürfnis das betreffende Merkmal selbst zuschreiben bzw. dessen Existenz leugnen, desto stärkere Antwortverzerrungen werden angenommen. Dabei ist das Vorzeichen der beobachteten Zusammenhänge der Indikator für die Richtung des SD-Bias. In einer Studie von **Cox** und Kollegen (1994) wird das Antwortverhalten der Befragten beim „Michigan Alcoholism Screening Test“ mit den durch die **Marlow-Crowne** SD-Skala erfassten individuellen Unterschieden im Bedürfnis nach sozialer Anerkennung in Beziehung gesetzt. Die Ergebnisse der Studie zeigen, dass die Befragten mit steigendem Anerkennungsbedürfnis immer geringere Tendenzen zum Alkoholmissbrauch berichten: je nach Geschlecht der Befragten werden negative Zusammenhänge zwischen  $r = -0.07$  und  $r = -0.46$  beobachtet. Auch der Zusammenhang zwischen dem Bedürfnis der Befragten nach sozialer Anerkennung und deren Antwortverhalten über das eigene Lebensglück wurde empirisch untersucht (**Kozma** und **Stones** 1988).<sup>4</sup> Dabei wurden positive Korrelationen zwischen  $r = .21$  und  $r = .40$  zwischen den Werten der Befragten auf der **Marlow-Crowne** SD-Skala und den drei untersuchten Instrumenten zur Erfassung des Lebensglücks gefunden. Demnach ist bei diesem Befragungsthema mit SD-Bias in Richtung auf zu hohe Angaben über das Lebensglück zu rechnen. Auch eine Studie von **Phillips** und **Clancy** (1972) findet eine signifikante Korrelation zwischen den Antworten auf die Frage „Alles zusammengenommen, würden Sie sagen, dass Sie im Moment sehr glücklich, ziemlich glücklich oder nicht allzu glücklich sind?“ und den Werten der Befragten auf einer Kurzversion der **Marlow-Crowne** SD-Skala. Dieses Ergebnis wurde auch im deutschsprachigen Raum repliziert (**Költringer** et al. 1989). Im Gegensatz zu den bisher berichteten Ergebnissen fand eine Studie von **Laicardi** und Kollegen (2001) keine Hinweise auf die Wirksamkeit des SD-Bias. Das mit der Eysenck-Lügenskala operationalisierte Bedürfnis nach sozialer Anerkennung hat sich als unabhängig von den durch die **Salamon** und **Conte** „Life Satisfaction Scale“ erfassten Angaben der Befragten über ihre Lebenszufriedenheit erwiesen.

---

4 Bei den drei verwendeten Skalen handelt es sich um die „Philadelphia Geriatric Center Moral Scale“, die „Memorial University of Newfoundland Scale of Happiness“ und den „Life Satisfaction Index-Z“.

In einigen Studien wurde das so genannte „Bogus Pipeline“-Verfahren (BP-Verfahren) zur Analyse der Stärke und Richtung des bei verschiedenen Befragungsthemen wirksamen SD-Bias herangezogen. Hierbei wird den Befragten durch eine aufwendige Apparatur die Fähigkeit des Experimentators zur Kontrolle des Wahrheitsgehaltes der Antworten glaubhaft gemacht. Wegen der Abschreckungswirkung der somit für die Probanden subjektiv hohen Entdeckungswahrscheinlichkeit von Falschangaben wird unter „Bogus Pipeline“-Bedingungen von wahrheitsgemäßerem Antwortverhalten ausgegangen (*Jones* und *Sigall* 1971). So haben etwa *Tourangeau* und Kollegen (1997) das Antwortverhalten der Probanden über ihren Alkoholkonsum unter BP-Bedingungen mit dem bei einer persönlichen Befragung verglichen. Es hat sich gezeigt, dass die Befragten bei fingierter Wahrheitskontrolle signifikant häufiger angeben, mehr als der Durchschnitt zu trinken (21% vs. 4%) und eher einen übermäßigen Alkoholgenuss zugeben (34% vs. 16%). Eine Meta-Analyse über den selbstberichteten Alkoholkonsum unter BP-Bedingungen findet dagegen sehr inkonsistente Ergebnisse (*Aguinis* et al. 1995). Manche der analysierten Studien fanden unter BP-Bedingungen deutlich höhere Angaben über den Alkoholkonsum, während in anderen Untersuchungen weniger Alkoholgenuss als in konventionellen Befragungen berichtet wurde.

In einigen Studien wurden die Erwünschtheitswahrnehmungen bei den Themen „Alkoholkonsum“ und „Lebensglück“ explizit erforscht und teilweise mit der Selbstzuschreibung dieser Merkmale durch das Antwortverhalten in Beziehung gesetzt. In der Studie von *Bradburn* und Kollegen (1978) haben die Befragten einer bevölkerungsrepräsentativen Stichprobe angegeben, ob sich die Mehrheit der Menschen in der Gesellschaft unbehaglich fühlen würde, Fragen über ihren Alkoholkonsum zu beantworten. Hiervon sind 10 Prozent der Informanten überzeugt, die selbst auch deutlich geringere Angaben über ihren Bier- und Weinkonsum sowie über den Genuss hochprozentiger alkoholischer Getränke gemacht haben. Dies führt die Autoren zur Schlussfolgerung, dass zumindest bei einem Teil der Befragten ein SD-Bias in Richtung auf zu geringe Konsumangaben vorlag.

Im Rahmen der Wortnormenforschung haben Probanden die persönliche Erwünschtheit der Eigenschaften „glücklich“ und „unglücklich“ auf einer Skala von -100 (außerordentlich wenig erwünschte Eigenschaft) bis +100 (außerordentlich erwünschte Eigenschaft) bewertet (*Mecklenbräuker* et al. 1994). Dabei zeigte sich, dass die studentischen Informanten eine glückliche Person mit einem durchschnittlichen Skalenwert von +10.8 deutlich positiver einschätzen als eine unglückliche Person mit einem Durchschnittswert von -4.2. In einer anderen Studie wurde die fremdeingeschätzte Erwünschtheit der Merkmale „glücklich“ und „unglücklich“ auf einer Skala von 1 (extrem unerwünscht) bis 9 (extrem erwünscht) erfasst (*Hampson*

et al. 1987). Nach den Ergebnissen haben die Befragten eine glückliche Person mit einem mittleren Skalenwert von 8.0 als von anderen Menschen in der Gesellschaft sehr positiv bewertet wahrgenommen. Dagegen wurde bei einer unglücklichen Person im Durchschnitt eine relativ negative gesellschaftliche Bewertung von 2.3 antizipiert. Der Zusammenhang zwischen dem Antwortverhalten der Befragten und der eingeschätzten sozialen Erwünschtheit des Glücklichseins wurde ebenfalls untersucht (*Phillips* und *Clancy* 1972). Es wurde hierbei nur die Bewertung des Merkmals „happy“ auf einer Skala von 1 bis 9 herangezogen. Die Ergebnisse haben gezeigt, dass sich Befragte mit einer hohen Erwünschtheitswahrnehmung des Lebensglücks auch mit hoher Wahrscheinlichkeit selbst als „glücklich“ bezeichnet haben.

Die vorliegenden Studien finden bezüglich des Alkoholkonsums und des Lebensglücks häufig, allerdings keineswegs immer Hinweise auf das Vorliegen von Antwortverzerrungen durch soziale Erwünschtheit. Diese inkonsistenten Befunde könnten einerseits darauf zurückgeführt werden, dass die für die Entstehung des SD-Bias relevanten Faktoren immer nur unvollständig berücksichtigt wurden. Andererseits könnten die Ursache hierfür auch in einer unangemessenen Erfassung der Erwünschtheitswahrnehmungen gesehen werden: in allen berichteten Studien wird ausschließlich die wenig differenzierte Standardoperationalisierung herangezogen.

### **3 Empirische Untersuchung**

#### **3.1 Stichprobe und Repräsentativität**

Die Teilnehmer unserer empirischen Studie entstammen einer lokalen, mehrstufigen Zufallsstichprobe der Wohnbevölkerung einer süddeutschen Großstadt. Die Grundgesamtheit der im Spätjahr 1999 durchgeführten Befragung bildeten Bürgerinnen und Bürger mit Lebensmittelpunkt im Stadtgebiet, die zum Befragungszeitpunkt mindestens 18 Jahre alt waren und die deutsche Staatsbürgerschaft besaßen. Durch die Mitarbeiter des Forschungsprojektes wurde mit einer Zufallsbegehung die Haushaltsstichprobe generiert und den Interviewern vorgegeben (Address-Random). Innerhalb der Haushalte wurde immer die Person für eine Befragung ausgewählt, die vor dem Kontaktzeitpunkt als letzte Geburtstag hatte.

Mit 143 verwertbaren Interviews ließ sich nur eine relativ schlechte Ausschöpfungsquote von 33.3 Prozent realisieren. Ein Vergleich der Zusammensetzung der in der Nettostichprobe repräsentierten Befragten mit den teilweise bekannten und teilweise durch Daten auf einer höheren Gebietsebene angenäherten sozialstrukturellen

Merkmale des Auswahlgebietes ergibt die üblicherweise in Umfragedaten vorliegenden Abweichungen.<sup>5</sup>

### 3.2 Vorgehensweise

Die Befragung erfolgte in allen Fällen als computergestütztes, persönliches Interview bei den Befragten zuhause. Die Interviewer lasen die Fragen laut vor, legten die passende Antwortskala in Papierform vor und gaben dann die Antworten der Probanden in den Laptop-Computer ein. Die Erhebung der Erwünschtheitswahrnehmungen bei den vier untersuchten Merkmalen ist Teil eines durchschnittlich 60-minütigen Interviews, wobei die für die vorliegende Untersuchung relevanten Fragen nach der Erhebung von Erwünschtheitswahrnehmungen über Einstellungen zu Ausländern und dem Demographieteil am Ende des ersten Drittels der Interviews gestellt wurden. Um die Belastung der Befragten zu minimieren, wurde die Gesamtstichprobe mit einem Pseudo-Zufallsgenerator des Befragungsprogramms in zwei gleich große Gruppen aufgeteilt. Jede der Teilgruppen des Populationssplits hat insgesamt 15 Fragen über die soziale Erwünschtheit unterschiedlicher Merkmale in einer randomisierten Reihenfolge beantwortet. Die Befragten des ersten Splits wurden neben den hier berichteten Wahrnehmungen über die soziale Erwünschtheit des Alkoholkonsums auch über die antizipierte Bewertung von Vorurteilen und sportlicher Betätigung sowie jener des Führerscheinentzugs und von Vorstrafen befragt (N=71). Die zweite Teilgruppe von Befragten hat neben der im vorliegenden Artikel untersuchten Fragen über die gesellschaftliche Bewertung des Fernsehkonsums, des Blutspendens und des Lebensglücks auch Fragen über die Bewertung einer Parteihängerschaft sowie von Prüfungsmisserfolgen beantwortet (N=72). Die Auswahl der in unserer Untersuchung hier dargestellten Befragungsthemen erfolgte danach, wie häufig entsprechende Fragen in der Umfrageforschung gestellt werden.

---

5 So sind in der Stichprobe Personen mit Abitur um 22 und Angestellte um 10 Prozentpunkte überrepräsentiert. Dies führt zu einer Unterrepräsentanz der jeweils anderen Bildungs- und Erwerbsstatusgruppen. Im Hinblick auf das Einkommen der Befragten sind in der Stichprobe Befragte mit hohem Einkommen überrepräsentiert, wohingegen Personen mit niedrigeren Ausbildungsabschlüssen (Lehre) um 13 und ledige Befragte um 10 Prozentpunkte unterrepräsentiert sind. Dagegen liegen hinsichtlich des Alters, des Geschlechts und der Konfession nur geringe Abweichungen von der Zusammensetzung der Grundgesamtheit vor. Die Vergleichsdaten für die Bildungs- und Ausbildungsabschlüsse, den Familienstand und die repräsentative Einkommensverteilung beziehen sich mangels spezifischer Daten auf die Bevölkerung Westdeutschlands im Dezember 1996 (*Statistisches Bundesamt* 1998). Die Aussagen über die Repräsentativität der im Datensatz vertretenen Geschlechts-, Alters- und Konfessionsstruktur beruhen auf spezifischen Zahlen der Stadt, welche die Grundgesamtheit definiert (*Stadt Mannheim* 1998).

### 3.3 Operationalisierung der verschiedenen TD-Versionen

Die Befragten haben die verschiedenen Merkmale im Sinne ihrer fremdeingeschätzten sozialen Erwünschtheit beurteilt (vgl. Abschnitt 2.1, oben). Sie wurden aufgefordert ihre Überzeugung darüber zu äußern, in welcher Weise eine Person in der Gesellschaft bewertet wird, die eine schwache, mittlere oder starke Ausprägung auf jeder der vier untersuchten Merkmalsdimensionen aufweist (vgl. die genaue Frageformulierung im Anhang). Die Angaben wurden auf einer bipolaren Antwortskala mit einem Wertebereich zwischen -4 (gesellschaftlich sehr negativ bewertet) bis +4 (gesellschaftlich sehr positiv bewertet) erfasst. Die Befragten haben angegeben, wie eine fiktive Person bewertet wird, die „sehr oft“, „gelegentlich“ oder „nie“ zum Blutspenden geht, ein „sehr“, „teilweise“ oder „überhaupt kein“ glückliches Leben führt, „sehr viel“, „mittelmäßig viel“ oder „nie“ Zeit vor dem Fernseher verbringt und wie eine Person bewertet wird, die „sehr häufig“, „gelegentlich“ oder „nie“ Alkohol trinkt.<sup>6</sup> Die *erste* und unaufwendigste Standardoperationalisierung des TD-Konstruktes besteht in der von den Informanten antizipierten gesellschaftlichen Bewertung einer starken Ausprägung der unterschiedlichen Merkmalsdimensionen (TD<sub>STARK</sub>). Als *zweite* Operationalisierung werden auf der Grundlage der Bewertung einer starken und schwachen Ausprägung der verschiedenen Merkmale Differenzwerte gebildet. Zu diesem Zweck werden die Werte beider Rohskalen zuerst auf einen Wertebereich von 0 (gesellschaftlich sehr negativ bewertet) bis +8 (gesellschaftlich sehr positiv bewertet) transformiert und dann die Erwünschtheitsurteile

---

6 Die hier gewählte Erfassung der Erwünschtheitswahrnehmungen bezieht sich auf konkrete Frageformulierungen, die in unterschiedlichen Umfragen in der Vergangenheit verwendet wurden. So wurde etwa im Eurobarometer 1994 (ZA-Studiennummer 2490) die Frage gestellt: „How often have you given blood? Only once/a few times/many times“. Eine ähnliche Frage wurde auch in der Studie „Aktionsgrundlagen der Bundeszentrale für gesundheitliche Aufklärung“ im Jahr 1995 (ZA-Studiennummer 3075) und im National Health Interview Survey 1993 (United States Department of Health and Human Services) gestellt. Im „Quality of Life Survey“ von **Campbell, Converse** und **Rodgers** im Jahr 1971 wurde nach dem Lebensglück der Befragten gefragt: „Asking all things together, how would you say things are these days? Would you say you're very happy, pretty happy or not too happy these days?“ Eine entsprechende Frage zum Lebensglück findet sich auch in der Untersuchung „Promaterielle und postmaterielle Lebensstile“ aus dem Jahr 1991 (ZA-Studiennummer 2421). Auch die Erwünschtheitswahrnehmungen des Fernsehkonsums entsprechen Frageformulierungen in früheren Umfragen. So wurde in der Vorwahlstudie des Jahres 1994 (ZA-Studiennummer 2599) gefragt: „Wie oft schauen Sie Fernsehen? Häufig/Manchmal/Nie“. Die gleiche Frage wurde auch in der Untersuchung „Gleichberechtigung von Frauen und Männern“ 1995 (ZA-Studiennummer 3028) gestellt. Nach dem Alkoholkonsum wurde im „National Health Interview Survey, 1988: Alcohol Supplement“ des United States Department of Health and Human Services ohne Vorgabe von Antwortkategorien gefragt: „On average, how often do you drink any alcoholic beverage?“. Das Eurobarometer 1992 (ZA-Studiennummer 1961) fragte mit sehr detaillierten Antwortkategorien nach dem Alkoholkonsum „How many units do you drink on average per week day from the time you get up till the time you go to bed?“. Es wurde auch, vergleichbar zu den anderen Themen, danach gefragt, ob die Probanden „häufig“, „gelegentlich“ oder „keinen“ Alkohol konsumieren.

einer schwachen Merkmalsausprägung von denen einer starken Ausprägung abgezogen ( $TD_{DIFF} = TD_{STARK} - TD_{SCHWACH}$ ). Die resultierend globalen Differenzwerte haben einen Wertebereich von -8 (starke Anreize für die Selbstzuschreibung einer schwachen Merkmalsausprägung) und +8 (starke Anreize für die Nennung einer starken Merkmalsausprägung). Bei der dritten und informationsmäßig anspruchsvollsten Umsetzung des TD-Konstruktes werden die Differenzwerte in der wahrgenommenen gesellschaftlichen Bewertung einer starken und mittleren Ausprägung der Merkmalsdimension einerseits und zwischen einer mittleren und schwachen Ausprägung andererseits berechnet ( $TD_{ST-MI} = TD_{STARK} - TD_{MITTEL}$ ;  $TD_{MI-SCH} = TD_{MITTEL} - TD_{SCHWACH}$ ). Jeder der beiden bereichsspezifischen Differenzwerte hat ebenfalls einen möglichen Wertebereich von -8 (starke Anreize zugunsten der Nennung einer mittleren bzw. schwachen Merkmalsausprägung) bis +8 (starke Anreize zur Nennung einer starken bzw. mittleren Merkmalsausprägung). Auch hier wurden die Skalenwerte vor der Differenzbildung auf einen Wertebereich von 0 bis 8 transformiert. Da sich die für das gesamte Merkmalskontinuum prognostizierten Anreize durch soziale Erwünschtheit aus der Summe der beiden Teilanreize ergeben, liegt der insgesamt resultierende Wertebereich dieser Operationalisierung zwischen -8 (starke Anreize für die Nennung einer schwachen Merkmalsausprägung) und +8 (starke Anreize für die Nennung einer starken Merkmalsausprägung). Hinsichtlich der Summe der Beträge der beiden bereichsspezifischen Messungen - hierbei handelt es sich dann um eine Operationalisierung des Verzerrungspotentials durch soziale Erwünschtheit ohne Berücksichtigung der Richtung der Anreize - können Gesamtwerte zwischen 0 (keine Anreize durch SD) und 16 (maximales Verzerrungspotential) realisiert werden. Bei den beiden anderen Operationalisierungen stimmt der Wertebereich des erfassten Verzerrungspotentials dagegen mit dem Betrag der richtungsgebundenen Werte überein ( $TD_{STARK}$ : [0/4];  $TD_{DIFF}$ : [0/8]). Dieser Unterschied geht darauf zurück, dass die bereichsspezifischen Differenzwerte die Möglichkeit nicht-linearer Erwünschtheitsverläufe über das Merkmalskontinuum erfassen.

Die drei Operationalisierungen des TD-Konstruktes haben auch, unter Berücksichtigung der Richtung des prognostizierten SD-Bias, formal unterschiedliche Wertebereiche:  $TD_{STARK}$  [-4/+4],  $TD_{DIFF}$  [-8/+8] und  $(TD_{ST-MI} + TD_{MI-SCH})$  [-8/+8]. Diese Unterschiede gehen inhaltlich darauf zurück, dass die möglichen Anreize durch soziale Erwünschtheit verschieden vollständig erfasst werden. Ausgehend von der Standardoperationalisierung erweitert sich der Wertebereich bei den globalen Differenzwerten daher von +4 bis -4 auf +8 bis -8, weil hier auch die Möglichkeit einbezogen wird, dass eine schwache Merkmalsausprägung extrem positiv und eine starke Ausprägung extrem negativ beurteilt wird, und umgekehrt. Aus der Logik der Standardoperationalisierung werden dagegen nur Abweichungen vom neutralen Nullpunkt in eine bestimmte Richtung als möglich angesehen. Dies gilt auch für

den erweiterten Wertebereich der bereichsspezifischen Differenzwerte. Die bei den Operationalisierungen resultierenden unterschiedlichen Wertebereiche sind somit das Resultat dessen, dass die Neutralitätsannahme und die Monotonieannahme bei den beiden differenzierteren Operationalisierungen aufgegeben werden. In diesen Unterschieden drückt sich somit *nicht* eine mangelnde Vergleichbarkeit der Konstrukte, sondern gerade die erweiterte Fähigkeit der Operationalisierungen zur Erfassung der Erwünschtheitswahrnehmungen aus.

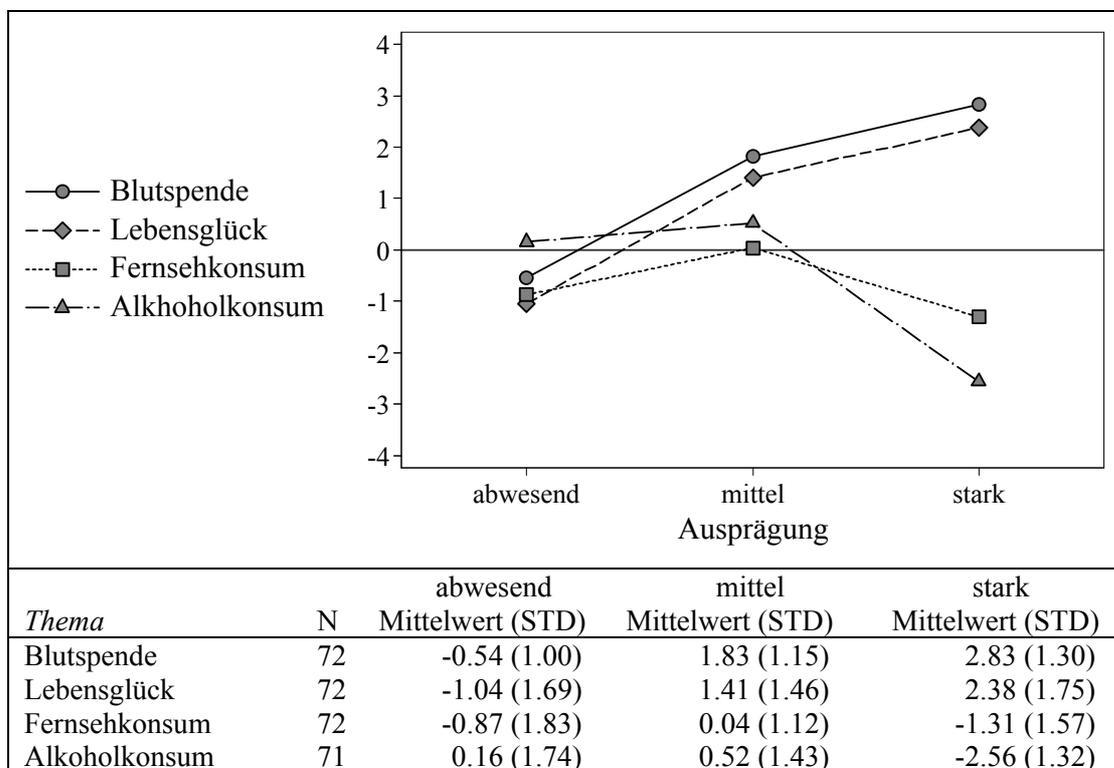
### 3.4 Ergebnisse

Die Präsentation der Ergebnisse unserer empirischen Analyse erfolgt in vier Schritten. Zuerst werden die durch die drei Operationalisierungen des TD-Konstruktes auf der Aggregatebene der Gesamtstichprobe für die untersuchten Merkmalsdimensionen prognostizierten Anreize durch soziale Erwünschtheit dargestellt. Im zweiten Schritt wird dann die der Standardoperationalisierung zugrunde liegende Annahme einer neutralen Bewertung der schwachen Merkmalsausprägungen empirisch überprüft und die Konsequenzen etwaiger Verstöße gegen diese Neutralitätsannahme analysiert. Im dritten Schritt wird dann getestet, ob sich die bei Verwendung der verschiedenen TD-Operationalisierungen jeweils vorhergesagte Richtung der Anreize durch soziale Erwünschtheit unterscheidet. Dabei wird untersucht, ob die im Rahmen der Standardoperationalisierung und dem Konzept der globalen Differenzwerte gemachte Monotonieannahme der Erwünschtheitsverläufe empirisch erfüllt ist. Im vierten und abschließenden empirischen Teil wird dann gezeigt, ob und bei welchen Befragungsthemen sich für die drei Indikatoren verschiedene Prognosen über die Gesamtstärke des SD-Bias ergeben.

#### 3.4.1 Beschreibung der mit verschiedenen Operationalisierungen des TD-Konzeptes prognostizierten Anreize auf der Aggregatebene

In Abbildung 1 werden die für die unterschiedlichen Ausprägungsstärken der Merkmalsdimensionen beobachteten Rohwerte der sozialen Erwünschtheitswahrnehmungen für die Gesamtstichprobe dargestellt. Dabei wird als erstes Ergebnis deutlich, dass auf der Aggregatebene eine schwache Ausprägung aller untersuchten Merkmale als gesellschaftlich relativ ähnlich und neutral bewertet wahrgenommen wird: Eine Person ohne Lebensglück wird im Durchschnitt mit einem Skalenwert von -1.0, eine Person ohne Fernsehkonsum mit einem Wert von -0.9, eine Person ohne Blutspenden mit einem Wert von -0.5 und eine Person mit völliger Alkoholabstinenz mit einem Wert von +0.2 bewertet. Die Neutralitätsannahme der Standardoperationalisierung scheint somit weitgehend erfüllt zu sein.

**Abbildung 1** Wahrgenommene soziale Erwünschtheit unterschiedlich starker Ausprägungen der Merkmalsdimensionen



Wertebereich der Antwortskalen: -4, „sehr negativ bewertet“ bis +4, „sehr positiv bewertet“.

Es zeigt sich weiterhin, dass ausgehend von der im Aggregat relativ neutralen Bewertung einer fehlenden Merkmalsausprägung ein sehr hohes Lebensglück und insbesondere sehr häufiges Blutspenden als gesellschaftlich deutlich erwünschter wahrgenommen werden. Die durchschnittlich von den Befragten antizipierte Bewertung liegt bei einer starken Ausprägung dieser Merkmale bei +2.4 bzw. +2.8. Dagegen erwarten die Befragten, dass Personen mit intensivem Fernseh- und vor allem hohem Alkoholkonsum eine deutlich negativere Bewertung erfahren. Hier werden durchschnittliche Skalenwerte von -1.3 und -2.6 beobachtet. Während sich die Erwünschtheitswahrnehmungen beim Thema Blutspende und Lebensglück auf der Aggregatebene annähernd linear ansteigend entwickeln, trifft dies für die Verläufe bei den beiden anderen Merkmalsdimensionen nicht zu. Beim Thema „Alkoholkonsum“ lassen sich sowohl für Befragte, die häufig und solche, die nie Alkohol trinken, Anreize für Angaben feststellen, die ein mittleres Konsumniveau anzeigen. Auch beim Fernsehkonsum zeigen die aggregierten Kennzahlen gewisse Anreize zur Angabe einer mittelstarken Ausprägung des Merkmals.

In Tabelle 1 ist die auf Grundlage der drei unterschiedlichen Operationalisierungen des TD-Konstruktes prognostizierte Stärke und Richtung der Anreize zu sozial er-

wünschtem Antwortverhalten auf der Aggregatebene der Gesamtstichprobe dargestellt. Bei den dargestellten Kennzahlen handelt es sich um Populationsmittelwerte der verschiedenen Anreizskalen bei den unterschiedlichen Befragungsthemen. Ein Vergleich der Standardoperationalisierung ( $TD_{STARK}$ ) mit der Operationalisierung durch globale Differenzwerte ( $TD_{DIFF}$ ) zeigt im ersten Schritt, dass beide Indikatoren bei allen Befragungsthemen in der Richtung der prognostizierten Anreize durch soziale Erwünschtheit übereinstimmen: es werden Antwortverzerrungen in Richtung auf ein höheres Lebensglück und häufigere Blutspenden sowie auf einen geringeren Alkohol- und Fernsehkonsum vorhergesagt. Es lässt sich weiterhin feststellen, dass auf der Grundlage der Differenzwerte signifikant stärkere Anreize zur Angabe häufigerer Blutspenden und einem höheren Lebensglück prognostiziert werden (Blutspende:  $t = 4.59$ ,  $df = 71$ ; Lebensglück:  $t = 5.23$ ,  $df = 71$ ; beide:  $p < 0.05$ ). Die weitere Analyse wird zeigen, dass dieser Unterschied auf die Verletzung der bei der Standardoperationalisierung zugrunde liegenden Neutralitätsannahme und der dadurch bei diesen Themen mehrheitlich resultierenden Unterschätzung der tatsächlich vorliegenden Anreize durch soziale Erwünschtheit zurückgeht (vgl. Abschnitt 3.4.2, unten).

**Tabelle 1** Durch verschiedene TD-Konstrukte prognostizierte Anreize durch soziale Erwünschtheit auf der Aggregatebene der Merkmale

<i>Thema</i>	$TD_{STARK}$ $\phi$ (STD)	$TD_{DIFF}$ $\phi$ (STD)	$TD_{ST-MI}$ $\phi$ (STD)	$TD_{MI-SCH}$ $\phi$ (STD)
Blutspende	2.83 (1.30)	3.38 (1.69)	1.01 (1.34)	2.37 (1.38)
Lebensglück	2.38 (1.75)	3.42 (2.78)	0.97 (1.78)	2.45 (2.24)
Fernsehkonsum	-1.31 (1.57)	-0.43 (2.47)	-1.35 (1.52)	0.91 (2.29)
Alkoholkonsum	-2.56 (1.32)	-2.72 (2.40)	-3.08 (1.75)	0.36 (2.63)

Fallzahl: Blutspende, Lebensglück, Fernsehkonsum:  $N=72$ ; Alkoholkonsum:  $N=71$ .

Beim Merkmal „Alkoholkonsum“ finden sich dagegen keine statistisch abgesicherten Unterschiede in der prognostizierten Anreizstärke zwischen der Standardoperationalisierung und den globalen Differenzwerten ( $t = 0.76$ ,  $df = 70$ ;  $p > 0.1$ ). Beim Thema „Fernsehkonsum“ werden durch die Standardoperationalisierung sogar signifikant stärkere Effekte sozialer Erwünschtheit vorhergesagt ( $t = 4.05$ ,  $df = 71$ ;  $p < 0.01$ ). Die aggregierten Kennzahlen der bereichsspezifischen Differenzwerte zeigen, dass bei den beiden letztgenannten Befragungsthemen die Monotonieannahme zumindest bei bestimmten Befragtengruppen verletzt ist: Die Richtung der aggregierten Anreize dreht sich im niedrigen Ausprägungsbereich der beiden Merkmalsdimensionen um. Da diese nicht-linearen Anreize durch die globalen Differenzwerte nicht berücksichtigt werden und sich (teilweise) gegenseitig aufheben, kann erklärt

werden, warum die Differenzwerte beim Fernseh- und Alkoholkonsum ein relativ geringes Verzerrungspotential durch den SD-Bias prognostizieren.

Ein weiteres interessantes Ergebnis ist, dass sich nach den bereichsspezifischen Differenzwerten die Anreize durch soziale Erwünschtheit bei den Themen „Blutspende“ und „Lebensglück“ stark zwischen dem Bereich einer starken und schwachen Merkmalsausprägung unterscheiden. Die Befragten nehmen bei beiden Themen mehr als doppelt so starke und damit statistisch signifikant stärkere Anreize zur Nennung einer mittleren statt einer schwachen Merkmalsausprägung wahr, als dies für die Nennung einer starken im Vergleich zu einer mittleren Ausprägung der Fall ist (Blutspende:  $t = 5.46$ ,  $df = 71$ ; Lebensglück:  $t = 4.28$ ,  $df = 71$ ; beide:  $p < 0.05$ ). Demnach wird vor allem das Antwortverhalten von Befragten mit sehr geringem Lebensglück und das von denen, die nie zum Blutspenden gehen, durch Anreize zu sozial erwünschtem Antwortverhalten beeinflusst.

### **3.4.2 Verletzung der Neutralitätsannahme und Konsequenzen für die Validität der Standardoperationalisierung**

Die Validität der empirisch unaufwendigen Standardoperationalisierung des TD-Konstruktes hängt unter anderem davon ab, dass die hierbei angenommene neutrale Bewertung einer schwachen Merkmalsausprägung erfüllt ist. Im Folgenden wird diese Annahme auf der Ebene der individuellen Erwünschtheitswahrnehmungen der Befragten überprüft und die Konsequenzen eventueller Verletzungen der Neutralitätsannahme analysiert. Zu diesem Zweck werden die Bewertungen einer starken und einer schwachen Ausprägung der unterschiedlichen Merkmalsdimensionen miteinander in Beziehung gesetzt (für die Vorhersagen vgl. Abschnitt 2.1).

Die in Tabelle 2 dargestellten Analyseergebnisse zeigen, dass im Durchschnitt über alle vier Befragungsthemen 62.4 Prozent der Befragten eine schwache Ausprägung der Merkmale gesellschaftlich nicht als neutral bewertet wahrnehmen. Dies ist mit 76.4 Prozent beim Thema „Fernsehkonsum“ am stärksten der Fall, wohingegen die Neutralitätsannahme beim Thema „Blutspenden“ mit einem Anteil von 33.3 Prozent noch am wenigsten verletzt ist. Die beiden anderen Themen weisen bei ungefähr 70 Prozent der Befragten Verletzungen der Neutralitätsannahme auf. Die Abweichung der sozialen Erwünschtheit der schwachen Merkmalsausprägungen von einer neutralen Bewertung führt beim Thema „Blutspende“ in 91.8 und beim „Lebensglück“ in 86.0 Prozent der Fälle zu einer Unterschätzung der vorliegenden Anreize durch

**Tabelle 2** Verletzung der Neutralitätsannahme und deren Konsequenzen für die prognostizierten Anreize durch soziale Erwünschtheit

Thema	Neutralitäts- annahme verletzt		Verletzung führt zu Unterschätzung der Anreizstärke		Verletzung führt zu Überschätzung der Anreizstärke		Verletzung führt zu falschen Schluss- folgerungen über die Richtung der Anreize	
	Anteil <sup>a)</sup>	Stärke <sup>b)</sup>	Anteil <sup>c)</sup>	Stärke <sup>b)</sup>	Anteil <sup>c)</sup>	Stärke <sup>b)</sup>	Anteil <sup>c)</sup>	Stärke <sup>b)</sup>
Blutspende	33.3	1.71	91.8	1.80	8.2	0.77	0.0	-
Lebensglück	69.4	2.18	86.0	2.16	8.0	2.00	6.0	2.68
Fernsehkonsument	76.4	2.09	47.2	1.88	34.6	1.89	18.2	3.00
Alkoholkonsum	70.4	1.78	48.0	2.26	52.0	1.35	0.0	-
Mittelwert	62.4	1.94	64.1	2.03	28.6	1.50	7.3	1.42

<sup>a)</sup> Prozentsatz aller Befragten mit einer bei der schwachen Merkmalsausprägung vom neutralen Nullpunkt abweichenden Erwünschtheitswahrnehmung.

<sup>b)</sup> Durchschnittlicher Betrag der bei einer schwachen Merkmalsausprägung beobachteten Abweichung vom neutralen Nullpunkt der Skala.

<sup>c)</sup> Anteil der Befragten mit einer Verletzung der Neutralitätsannahme, bei denen die betreffende Art von Messfehler auftritt. Fallzahl: Blutspende, Lebensglück, Fernsehkonsument: N=72; Alkoholkonsum: N=71.

soziale Erwünschtheit. Dabei ist die Unterschätzung beim Merkmal „Blutspende“ im Durchschnitt 1.8 und beim Lebensglück 2.2 Punkte der Anreizskala stark.

Bei den Befragungsthemen „Alkohol-“ und „Fernsehkonsument“ sind die Folgen der nicht erfüllten Neutralitätsannahme weniger homogen. Beim Thema „Alkoholkonsum“ führt die Verletzung bei 48.0 Prozent zu einer Unterschätzung und bei 52.0 Prozent zu einer Überschätzung der Anreize zum SD-Bias. Allerdings ist die durchschnittliche Unterschätzung mit 2.3 Skalenpunkten deutlich stärker als die Überschätzung, die im Mittel nur 1.4 Skalenpunkte beträgt. Beim Befragungsgegenstand „Fernsehkonsument“ ergeben sich bei einer Verletzung der Neutralitätsannahme in 47.2 Prozent der Fälle eine Unterschätzung und bei 34.6 Prozent eine Überschätzung der vorliegenden Anreize durch soziale Erwünschtheit. Bei den letztgenannten Themen heben sich somit die Messfehler in entgegengesetzter Richtung in gewissem Umfang auf der Aggregatebene gegenseitig auf. Die Verletzung der Neutralitätsannahme führt außerdem beim Merkmal „Fernsehkonsument“ bei 18.2 Prozent der Befragten zur Prognose von Anreizen durch soziale Erwünschtheit in die falsche Richtung.

### 3.4.3 Prognostizierte Richtung der Anreize durch soziale Erwünschtheit auf der Individualebene

Im folgenden Abschnitt wird die Heterogenität bezüglich der in unserer Stichprobe wahrgenommenen Richtung der Anreize durch soziale Erwünschtheit analysiert. Bei der Standardoperationalisierung ist die Richtung durch die Bewertung der starken Ausprägung festgelegt: wird diese positiv bewertet, so liegen Anreize zur Nennung einer starken und bei negativer Bewertung Anreize für die Angabe einer schwachen Ausprägung vor. Für die Operationalisierung mit globalen Differenzwerten ergibt sich die Richtung der Anreize aus dem Vorzeichen dieser Werte. Dies gilt auch für die bereichsspezifische Erfassung, wobei sich hier die Anreizrichtung zwischen den Bereichen unterscheiden kann und dann zu u- oder umgekehrt u-förmigen Erwünschtheitsverläufen führt. Es wird getestet, ob und gegebenenfalls bei welchen Befragungsthemen die verschiedenen Operationalisierungen der Erwünschtheitswahrnehmungen zu unterschiedlichen Vorhersagen über die Richtung und den Verlauf der Anreize zu Antwortverzerrungen durch soziale Erwünschtheit kommen.

In Tabelle 3 ist die von den unterschiedlichen Operationalisierungen des TD-Konstruktes prognostizierte Richtung der bestehenden Anreize durch soziale Erwünschtheit für die Themen „Blutspende“ und „Lebensglück“ dargestellt. Dabei lässt sich zuerst einmal feststellen, dass alle drei Operationalisierungen des TD-Konstruktes übereinstimmend vorhersagen, dass die Angaben über die Häufigkeit des Blutspendens potentiell in Richtung auf zu hohe Angaben verzerrt werden: Die Standardoperationalisierung sagt dies für 93.1 Prozent, die Differenzwerte für 94.4 Prozent und die bereichsspezifischen Differenzwerte für 88.9 Prozent der Befragten vorher. Weiterhin prognostizieren alle Operationalisierungen, dass nur um die 5 Prozent der Befragten bei diesem Thema keinerlei Anreize zu Antwortverzerrungen wahrnehmen. Andere Anreiztypen, wie die zugunsten einer schwachen Merkmalsausprägung und umgekehrt u-förmige Anreizprofile, sind wegen der sehr geringen Fallzahlen in diesen Kategorien nicht sinnvoll interpretierbar.

Beim Befragungsthema „Lebensglück“ zeigt sich prinzipiell ein sehr ähnliches Bild wie bei der Blutspende. Hier werden bei Verwendung der Standardoperationalisierung 84.7 Prozent, bei den globalen Differenzwerten 83.3 Prozent und bei den bereichsspezifischen Differenzwerten 76.3 Prozent der Befragten mit einer Anreizwahrnehmung zugunsten der Nennung einer starken Merkmalsausprägung vorhergesagt. Befragte, die keinerlei Anreize für die Nennung einer bestimmten Merkmalsausprägung wahrnehmen, bilden auch hier mit einem Anteil zwischen 5.6 und 9.7 Prozent nur eine kleine Minderheit. Bei der differenzierten, bereichsspezifischen

Betrachtung der Erwünschtheitsverläufe wird außerdem eine ebenfalls kleine, 9.7 Prozent starke Gruppe von Befragten mit einer umgekehrt u-förmigen Anreizwahrnehmung vorhergesagt. Diese Befragten gehen davon aus, dass ein mittleres, quasi normales Ausmaß an Lebensglück gesellschaftlich am wünschenswertesten angesehen wird. Da es sich hierbei allerdings absolut gesehen nur um 7 Befragte handelt, verzichten wir bis zu einer Replikation dieses Erwünschtheitsprofils mit einer größeren Stichprobe auf eine substantielle Interpretation dieses Ergebnisses. Insgesamt kann jedenfalls festgestellt werden, dass den verschiedenen Operationalisierungen des TD-Konstruktes bei den Befragungsthemen „Blutspende“ und „Lebensglück“ insgesamt ein hohes Ausmaß an konvergenter Validität hinsichtlich der prognostizierten Richtung des Verzerrungspotentials durch soziale Erwünschtheit zukommt.

**Tabelle 3** Durch die TD-Konstrukte prognostizierte Anteile von Befragten mit unterschiedlichen Anreizprofilen bei den Themen „Blutspende“ und „Lebensglück“

<i>Anreize zugunsten einer...</i>	TD <sub>STARK</sub>		TD <sub>DIFF</sub>		TD <sub>ST-MI</sub> /TD <sub>MI-SCH</sub>	
	%	N	%	N	%	N
Blutspende						
- starken Merkmalsausprägung	93.1	67	94.4	68	88.9	64
- schwachen Merkmalsausprägung	1.4	1	1.4	1	--	--
- keine Anreize	5.6	4	4.2	3	4.2	3
- u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	--	--
- umgekehrt u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	6.9	5
Gesamt	100.0	72	100.0	72	100.0	72
Lebensglück						
- starken Merkmalsausprägung	84.7	61	83.3	60	76.3	55
- schwachen Merkmalsausprägung	6.9	5	6.9	5	5.6	4
- keine Anreize	8.3	6	9.7	7	5.6	4
- u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	2.8	2
- umgekehrt u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	9.7	7
Gesamt	100.0	72	100.0	72	100.0	72

In Tabelle 4 wird die Richtung der nach den verschiedenen Operationalisierungen des TD-Konstruktes von den Befragten wahrgenommenen Anreize durch soziale Erwünschtheit für die Themen „Fernsehkonsum“ und „Alkoholkonsum“ dargestellt. Hierbei zeigt sich auf der Ebene der individuellen Befragten ein beträchtliches Ausmaß an Heterogenität. So lässt sich feststellen, dass die Standardoperationalisierung des TD-Konstruktes für eine Mehrheit von 72.2 Prozent der Befragten Anreize zugunsten der Angabe eines sehr geringen Ausmaßes an Fernsehkonsum prognostiziert. Für 9.7 Prozent wird eine umgekehrte Prognose gemacht, während 18.1 Prozent keinerlei Anreize durch soziale Erwünschtheit zugeschrieben wird. Bei der Verwendung der globalen Differenzwerte ändert sich die Prognose dergestalt, dass für deutlich mehr Befragte, nämlich 31.9 Prozent, Anreize für die Nennung eines

hohen Ausmaßes an Fernsehkonsum vorhergesagt wird. Dieser Unterschied geht auf die durch die Verletzung der Neutralitätsannahme bei diesem Thema im Abschnitt 3.4.2 festgestellten Fehlprognosen der Anreizrichtung zurück. Von besonderer Bedeutung ist jedoch das Ergebnis bei Verwendung der bereichsspezifischen Differenzwerte, wonach 43.1 Prozent der Befragten Anreize zur Nennung eines mittelstarken Ausmaßes an Fernsehkonsum wahrnehmen. In dieser Gruppe wird sowohl ein hoher Fernsehkonsum als auch ein völliger Verzicht als gesellschaftlich weniger anerkannt angesehen. Gleichzeitig nehmen aber auch 31.9 Prozent konsistent Anreize für die Angabe eines vollständigen Fernsehverzichtes wahr, und 18.1 gehen davon aus, dass ein intensiver Fernsehkonsum das höchste Maß an sozialer Anerkennung genießt.

**Tabelle 4** Durch die TD-Konstrukte prognostizierte Anteile von Befragten mit unterschiedlichen Anreizprofilen bei den Themen „Fernsehkonsum“ und „Alkoholkonsum“

<i>Anreize zugunsten einer...</i>	TD <sub>STARK</sub>		TD <sub>DIFF</sub>		TD <sub>ST-MI</sub> /TD <sub>MI-SCH</sub>	
	%	N	%	N	%	N
Fernsehkonsum						
- starken Merkmalsausprägung	9.7	7	31.9	23	18.1	13
- schwachen Merkmalsausprägung	72.2	52	48.6	35	31.9	23
- keine Anreize	18.1	13	19.4	14	4.1	3
- u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	2.8	2
- umgekehrt u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	43.1	31
Gesamt	100.0	72	100.0	72	100.0	72
Alkoholkonsum						
- starken Merkmalsausprägung	1.4	1	1.4	1	1.4	1
- schwachen Merkmalsausprägung	93.0	66	91.5	65	49.3	35
- keine Anreize	5.6	4	7.0	5	1.4	1
- u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	--	--
- umgekehrt u-förmige Anreize	n.v.	n.v.	n.v.	n.v.	47.9	34
Gesamt	100.0	71	100.0	71	100.0	71

Beim Thema „Alkoholkonsum“ prognostiziert die Standardoperationalisierung für 93.0 Prozent und die globalen Differenzwerte für 91.5 Prozent der Befragten Antwortverzerrungen in Richtung auf ein Antwortverhalten, in dem sich ein vollständiger Verzicht auf Alkoholgenuss ausdrückt. Bei beiden Anreizindikatoren werden außerdem für 5.6 bzw. 7.0 Prozent der Befragten keinerlei Anreize zu sozial erwünschtem Antwortverhalten vorhergesagt. Auf der Grundlage bereichsspezifischer Differenzwerte wird aber deutlich, dass mit einem Anteil von 47.9 Prozent fast die Hälfte der Befragten Anreize für die Nennung eines mittelstarken und damit moderaten Alkoholkonsums wahrnimmt. Demnach wird, ähnlich wie beim Fernsehkonsum, weder die vollständige Abstinenz noch ein hoher Konsum als die gesellschaftlich anerkannteste Art des Umgangs mit Alkohol angesehen.

### 3.4.4 Prognostizierte Gesamtstärke der Anreize durch soziale Erwünschtheit

Im folgenden und abschließenden Teil der Analyse wird nun die im Rahmen der drei unterschiedlichen Operationalisierungen des TD-Konstruktes prognostizierte Gesamtstärke der Anreize durch soziale Erwünschtheit verglichen. Hierzu werden die durchschnittlichen Anreizstärken, die als Mittelwerte der Beträge der bei den verschiedenen Operationalisierungen prognostizierten Verzerrungen definiert wurden (siehe Abschnitt 2.1 bzw. 3.3, oben), herangezogen. Die Richtung der Anreize wird somit vernachlässigt.

**Tabelle 5** Durch die verschiedenen Konstrukte prognostizierte Gesamtstärke der Anreize durch soziale Erwünschtheit auf der Individualebene

<i>Thema</i>	TD <sub>STARK</sub> <sup>a)</sup> φ (STD)	TD <sub>DIFF</sub> <sup>b)</sup> φ (STD)	TD <sub>ST-MI</sub>   +  TD <sub>MI-SCH</sub>    <sup>c)</sup> φ (STD)
Blutspende	2.89 (1.17)	3.42 (1.60)	3.62 (1.66)
Lebensglück	2.68 (1.22)	3.81 (2.20)	4.20 (2.10)
Fernsehkonsum	1.69 (1.13)	1.90 (1.61)	3.48 (2.14)
Alkoholkonsum	2.65 (1.14)	2.92 (2.16)	5.06 (2.60)
Gesamt	2.48 (0.65)	3.01 (1.04)	4.09 (1.24)

<sup>a)</sup> Wertebereich: 0, „keine Anreize“ bis +4, „sehr starke Anreize“;

<sup>b)</sup> Wertebereich: 0, „keine Anreize“ bis +8, „sehr starke Anreize“;

<sup>c)</sup> Wertebereich: 0, „keine Anreize“ bis +16, „sehr starke Anreize“. Fallzahl: Blutspende, Lebensglück, Fernsehkonsum: N=72; Alkoholkonsum: N=71.

In Tabelle 5 sind diese durchschnittlich vorhergesagten Anreizstärken durch soziale Erwünschtheit dargestellt. Dabei ist erkennbar, dass bei jedem der Befragungsthemen die Standardoperationalisierung das geringste und die bereichsspezifischen Differenzwerte das höchste Verzerrungspotential durch soziale Erwünschtheit vorhersagt. Die globalen Differenzwerte nehmen dabei in jedem Fall eine Mittelstellung ein. Es kann weiterhin festgestellt werden, dass sich bei den Themen „Blutspende“ und „Lebensglück“ der mit den globalen Differenzwerten verbundene Mehraufwand in der Identifikation eines signifikant höheren Verzerrungspotentials durch soziale Erwünschtheit niederschlägt (Blutspende:  $t = 4.43$ ,  $df = 71$ ;  $p < 0.05$ ; Lebensglück:  $t = 6.18$ ,  $df = 71$ ;  $p < 0.05$ ). Dieses Phänomen ist eine Folge der Verletzung der Neutralitätsannahme bei beiden Themen und führt dazu, dass die Standardoperationalisierung die Anreize durch soziale Erwünschtheit unterschätzt (vgl. Abschnitt 3.4.2, oben). Dagegen finden sich bei den Merkmalen „Fernsehkonsum“ und „Alkoholkonsum“ kein statistisch abgesicherter Anstieg der identifizierten Stärke der Anreize durch soziale Erwünschtheit, wenn statt der Standardoperationalisierung

globale Differenzwerte herangezogen werden (Fernsehkonsument:  $t = 1.05$ ,  $df = 71$ ;  $p > 0.1$ ; Alkoholkonsum:  $t = 1.32$ ,  $df = 70$ ;  $p > 0.1$ ). Bei diesen Themen hat sich in Abschnitt 3.4.2 herausgestellt, dass die Verletzung der Neutralitätsannahme teilweise zu einer Über- und teilweise zu einer Unterschätzung der tatsächlich vorliegenden Anreizstärken führen. Diese Messfehler gleichen sich daher in gewissem Umfang gegenseitig aus.

Ausgehend von den mithilfe der globalen Differenzwerte prognostizierten Anreize durch soziale Erwünschtheit, kann mit dem bei den bereichsspezifischen Differenzwerten betriebenen Mehraufwand bei jedem Befragungsthema ein signifikant höheres Ausmaß an Verzerrungspotential durch soziale Erwünschtheit identifiziert werden. So steigt beim Thema „Blutspende“ der Anreizwert von 3.42 um 0.2 Skalenpunkte auf 3.62 an ( $t = 2.13$ ,  $df = 71$ ,  $p < 0.05$ ). Beim Thema „Lebensglück“ beträgt dieser Anstieg von 3.81 auf 4.20 der Anreizskala 0.39 Skalenpunkte ( $t = 2.88$ ,  $df = 71$ ;  $p < 0.05$ ). Es ist weiterhin erkennbar, dass der bei der Erfassung von drei Messpunkten notwendige Mehraufwand insbesondere bei den beiden Themen mit einem substantiellen Ausmaß an nicht linearen Erwünschtheitsverläufen zu einer besseren Identifikation des Verzerrungspotentials durch soziale Erwünschtheit führt. Hier steigt das entdeckte Verzerrungspotential beim Befragungsthema „Fernsehkonsument“ von 1.90 auf 3.48 und damit um 1.58 Skalenpunkte an ( $t = 5.71$ ,  $df = 71$ ;  $p < 0.05$ ). Beim Thema „Blutspende“ beträgt dieser Anstieg, von 2.92 auf 5.06, sogar 2.14 Punkte der Anreizskala ( $t = 6.36$ ,  $df = 70$ ;  $p < 0.05$ ). Nimmt man die absolute Höhe der von den bereichsspezifischen Differenzwerten, als der differenziertesten und zuverlässigsten Operationalisierung des SD-Bias, prognostizierten Antwortverzerrungen als Referenzpunkt, so kommt bezüglich der erfassten Gesamtanreize keinem der beiden anderen Indikatoren konvergente Validität zu.

Ein Vergleich des mit bereichsspezifischen Differenzwerten identifizierten Verzerrungspotentials durch soziale Erwünschtheit zwischen den untersuchten Befragungsthemen zeigt, dass bei den Themen „Blutspende“ sowie „Fernsehkonsument“ das geringste und ein ähnliches Verzerrungspotential vorliegt: die Anreizstärken von 3.48 (Fernsehkonsument) und 3.62 (Blutspende) unterscheiden sich nicht signifikant ( $t = 0.46$ ,  $df = 71$ ;  $p > 0.1$ ). Beim Thema „Alkoholkonsum“ liegt mit einem mittleren Skalenwert von 5.06 das stärkste Verzerrungspotential durch soziale Erwünschtheit vor, das signifikant stärker ist, als das beim zweitsensibelsten Merkmal „Lebensglück“ mit einem Wert von 4.20 beobachtet wird ( $t = 2.3$ ,  $df = 70$ ;  $p < 0.05$ ).

In Tabelle 6 wird überprüft, in welchem Umfang die verschiedenen Indikatoren zumindest darin übereinstimmen, welche Befragten mehr oder weniger Anreize zu Antwortverzerrungen durch soziale Erwünschtheit wahrnehmen. Dieses weitere

Kriterium für die konvergente Validität wird durch die Korrelationen zwischen den drei Indikatoren für jedes Befragungsthema überprüft. Dabei zeigt sich, dass bei den Merkmalsdimensionen „Blutspende“ und „Lebensglück“, nun unabhängig von der Anreizhöhe, mittlere bis starke Zusammenhänge beobachtet werden. Dies trifft insbesondere für die Übereinstimmung zwischen den globalen und bereichsspezifischen Differenzwerten zu: hier finden sich Korrelationen von  $r = .87$  bzw.  $r = .86$ . Demnach kann die weniger aufwendige Operationalisierung mit zwei Messpunkten hier als recht gute Annäherung an die optimale bereichsspezifische Erfassung der Anreize angesehen werden. Für die Standardoperationalisierung ist dies mit Korrelationen zwischen  $.61$  und  $.65$  schon deutlich weniger der Fall.

**Tabelle 6** Korrelation zwischen der von den verschiedenen TD-Konstrukten prognostizierten Stärke der Gesamtanreize durch soziale Erwünschtheit (Pearsons  $r$ )

<i>Thema</i>	Korrelation von TD <sub>STARK</sub> mit TD <sub>DIFF</sub>	Korrelation von TD <sub>STARK</sub> mit TD <sub>ST-MI</sub> + TD <sub>MI-SCH</sub>	Korrelation von TD <sub>DIFF</sub> mit TD <sub>ST-MI</sub> + TD <sub>MI-SCH</sub>
	$r$	$r$	$r$
Blutspende	0.78**	0.61**	0.87**
Lebensglück	0.74**	0.65**	0.86**
Fernsehkonsum	0.27*	0.54**	0.24*
Alkoholkonsum	0.61**	0.58**	0.30*

Signifikanz: \*\*  $p < 0.01$ , \*  $p < 0.05$ ;

Fallzahl: Blutspende, Lebensglück, Fernsehkonsum:  $N=72$ ; Alkoholkonsum:  $N=71$ .

Bei den beiden Befragungsthemen „Fernsehkonsum“ und „Alkoholkonsum“ erweisen sich dagegen die globalen Differenzscores als sehr schlechte Annäherung an die Prognosen der bereichsspezifischen Messung: es finden sich nur schwache Zusammenhänge zwischen  $r = .24$  und  $r = .30$  zwischen den beiden Indikatoren. Dies kann auf die Vernachlässigung der bei diesen Befragungsthemen für substantielle Teile der Befragtenpopulation nicht-linear verlaufenden Anreizwahrnehmungen zurückgeführt werden. Dagegen erreicht die Standardoperationalisierung, mit Korrelationen mit der bereichsspezifischen Messung zwischen  $r = .54$  und  $r = .58$ , zumindest ein mittleres Korrelationsniveau. Dies trifft auch auf den Zusammenhang von  $r = .61$  zwischen den globalen Differenzwerten und der Standardoperationalisierung beim Befragungsthema „Alkoholkonsum“ zu. Dagegen zeigt sich beim Fernsehkonsum hier mit einer Korrelation von  $r = .27$  ein sehr geringes Ausmaß an konvergenter Validität. Dies geht darauf zurück, dass die Standardoperationalisierung bei diesem Thema wegen der verletzten Neutralitätsannahme für einen Teil der Befragten falsche Prognosen über die Richtung und damit auch über die Stärke der vorliegenden Anreize macht.

#### 4 Zusammenfassung und Diskussion

Die zentrale Fragestellung unserer Untersuchung besteht darin, wie differenziert die von den Befragten antizipierte gesellschaftliche Bewertung kontinuierlich variierender Merkmale erfasst werden muss, um eine zuverlässige Prognose des Verzerrungspotentials durch soziale Erwünschtheit zu ermöglichen. Damit untersucht der Beitrag eine von drei der im Rahmen der Rational-Choice-Theorie als relevant angesehenen Determinanten des SD-Bias. Es handelt sich hierbei daher um einen besonders hervorgehobenen Faktor, weil dieser nicht nur die Stärke sondern auch die Richtung der zu erwartenden Antwortverzerrungen vorhersagt. Die vorliegende Untersuchung hat als *erstes* und wichtigstes Ergebnis gezeigt, dass bei zwei der vier untersuchten Themen für einen substantiellen Teil der Befragten die in der bisherigen Forschung implizit gemachte Annahme monotoner Erwünschtheitsverläufe nicht erfüllt ist. So nehmen 43 Prozent der Befragten beim Thema „Fernsehkonsument“ und rund die Hälfte beim Thema „Alkoholkonsum“ einen umgekehrt u-förmigen Erwünschtheitsverlauf über das Kontinuum unterschiedlich starker Ausprägungen dieser Merkmale wahr. Demnach antizipieren die Befragten bei einer Nennung eines mittelstarken Fernseh- oder Alkoholkonsums positivere Bewertungsreaktionen als dies bei vollständiger Abstinenz oder intensivem Konsum der Fall ist. Der Vergleich der drei von uns berücksichtigten unterschiedlichen Indikatoren für das Verzerrungspotential durch soziale Erwünschtheit hat bei diesen Themen das Ergebnis erbracht, dass ausschließlich die differenzierteste und daher empirisch aufwendigste Erfassung durch bereichsspezifische Differenzwerte als angemessen angesehen werden kann. Weder die in der bisherigen Forschung annähernd ausschließlich verwendete Bewertung nur einer starken Merkmalsausprägung, noch die globale Erwünschtheitsdifferenz einer starken und schwachen Merkmalsausprägung erfassen unter der Bedingung nicht-monotoner Erwünschtheitsverläufe die wahre Stärke und Richtung der Anreize durch soziale Erwünschtheit.

Als *zweites* wichtiges Ergebnis unserer Untersuchung hat sich bei den Befragungsthemen „Blutspende“ und „Lebensglück“ gezeigt, dass alle drei Indikatoren der Anreize zu sozial erwünschtem Antwortverhalten übereinstimmend Anreize zur Angabe einer starken Merkmalsausprägung vorhersagen: häufigeres Blutspenden und ein größeres Lebensglück werden über die komplette Merkmalsdimension als erwünschter wahrgenommen. Damit kommt allen untersuchten Operationalisierungen konvergente Validität hinsichtlich der Richtung des prognostizierten Verzerrungspotentials durch soziale Erwünschtheit zu. Unsere Analyse hat aber auch gezeigt, dass die der Standardoperationalisierung des TD-Konstruktes implizit zugrunde liegende Annahme einer neutralen Bewertung einer schwachen Merkmalsausprägung im Durchschnitt bei allen Befragungsthemen für 62 Prozent der Befragten nicht er-

füllt ist. Dies führt bei den Themen „Blutspende“ und „Lebensglück“ zu einer massiven Unterschätzung des möglichen SD-Bias. Obwohl auch bei der Verwendung globaler Differenzwerte die vorliegenden Anreize durch soziale Erwünschtheit quantitativ nicht vollständig erfasst werden, ist dieser Indikator stark mit den bereichsspezifischen Differenzwerten als zuverlässigste Operationalisierung des TD-Konstruktes korreliert.

Zusammenfassend lassen unsere Ergebnisse die Schlussfolgerung zu, dass die differenzierteste Erfassung des möglichen SD-Bias mit bereichsspezifischen Differenzwerten bei allen untersuchten Befragungsthemen der überlegene Indikator ist. Bei Themen mit nicht-monotonen Erwünschtheitsverläufen ist diese aufwendigste Art der Operationalisierung des TD-Konstruktes unabdingbar. Wegen der bei monotonen Erwünschtheitsverläufen hohen konvergenten Validität globaler Differenzwerte kann unter diesen Umständen die Verwendung dieses sparsameren, weil nur auf zwei Messpunkten beruhenden Indikators als gerechtfertigt angesehen werden.

Es muss betont werden, dass es sich bei den in der vorliegenden Studie untersuchten Erwünschtheitswahrnehmungen nur um einen Indikator für die *möglichen* Antwortverzerrungen durch soziale Erwünschtheit handelt. Ob sich diese tatsächlich in einem SD-Bias niederschlagen, kann erst beurteilt werden, wenn deren Erklärungskraft für das Antwortverhalten untersucht wurde. Hierbei möglicherweise beobachtete Zusammenhänge können allerdings auch auf eine tatsächliche Übereinstimmung zwischen den wahren Merkmalen der Befragten mit gesellschaftlichen Normierungen oder darauf zurückgeführt werden, dass die Befragten ihre Eigenschaftsausprägungen auf ihre Erwünschtheitswahrnehmungen projektzieren. Eine abschließende Beurteilung der unterschiedlichen Operationalisierungen der Erwünschtheitswahrnehmungen macht daher den Einbezug weiterer Faktoren, wie etwa der Privatheit der Interviewsituation und des Anerkennungsbedürfnisses der Befragten, notwendig. Nach den Hypothesen der Rational-Choice-Theorie sagen Erwünschtheitswahrnehmungen einen SD-Bias in dem Ausmaß vorher, in dem sich deren Erklärungskraft nach dem Anerkennungsbedürfnis der Befragten und der Privatheit der Antwortsituation unterscheidet. Der Nachweis, dass sich die verschiedenen Operationalisierungen der Erwünschtheitswahrnehmungen im Rahmen einer solchen Gesamterklärung des Antwortverhaltens in ihrer Erklärungskraft unterscheiden, steht noch aus.

Ergebnisse einer Untersuchung von *Stocké* (2004) haben gezeigt, dass die mit globalen Differenzwerten gemessene Erwünschtheit unterschiedlicher Einstellungen zu Ausländern auch bei sichergestellter Privatheit der Antwortsituation, allerdings signifikant schwächer als bei der Wahrnehmbarkeit der Antwortinhalte, mit den Anga-

ben der Befragten über ihre Ausländereinstellung korreliert ist. Dies ist ein Beleg dafür, dass Erwünschtheitswahrnehmungen nicht nur die Anreize zu positiver Selbstdarstellung, sondern in gewissem Umfang die „wahren“ Werte der Befragten erfassen. Es sollte daher auch für die in der vorliegenden Untersuchung erfassten Erwünschtheitswahrnehmungen explizit getestet werden, in welchem Umfang diese tatsächlich einen SD-Bias bei den verschiedenen Befragungsthemen bewirken.

In unserer Studie wurde bei der Erfassung der sozialen Erwünschtheit die jeweils zu bewertende Intensität der analysierten Merkmale verbal beschrieben. Dies birgt eine gewisse Unsicherheit darüber, wie die Befragten diese relativ vagen Quantifizierungen, wie etwa „viel“ oder „mittelmäßig viel“ Zeit vor dem Fernseher zu verbringen, interpretieren (vgl. hierzu etwa *Bradburn* und *Miles* 1979). Eine alternative Vorgehensweise hätte sein können, den Informanten objektiv definierte Quantitäten vorzugeben, und etwa zu fragen, wie eine Person bewertet würde, die „4 Stunden“ oder „2 Stunden“ täglich vor dem Fernseher verbringt. Dabei hätte sich allerdings, neben einer gewissen Überforderung der Befragten durch eine zu genaue Charakterisierung der zu bewertenden Quantität, auch das Problem der angemessenen Auswahl der verschiedenen objektiven Quantitäten ergeben. Wir haben es daher vorgezogen, den Befragten selbst die Interpretation der verbalen Quantifizierungen zu überlassen. Es muss aber betont werden, dass unsere Daten streng genommen für solche Frageformulierungen Gültigkeit haben, die ähnliche Stimuli wie in unserer Studie zur Definition der Antwortmöglichkeiten verwenden. In jedem Fall aber definieren die Quantifizierungen in unserer Untersuchung eine ordinale Reihenfolge der Merkmalsausprägungen und die Ergebnisse sind jedenfalls in diesem Sinne gültig.

Wegen der Verwendung einer lokal definierten Zufallsstichprobe und der vergleichsweise geringen Ausschöpfung können unsere Ergebnisse nicht einfach auf die Verhältnisse einer bevölkerungsrepräsentativen Befragung generalisiert werden. Allerdings betrifft dieses Problem der Generalisierbarkeit vor allem die beobachteten Randverteilungen der Erwünschtheitswahrnehmungen und Anreizstärken. Die Schlussfolgerungen in Bezug auf unser Hauptziel, nämlich den Vergleich der Validität unterschiedlicher Operationalisierungen der wahrgenommenen Anreize zu sozial erwünschtem Antwortverhalten, sind hiervon nicht betroffen.

### Literatur

*Aguinis, Herman; Pierce, Charles A.; Quigley, Brian M.* 1995: Enhancing the Validity of Self-Reported Alcohol and Marijuana Consumption Using a Bogus Pipeline Procedure: A Meta Analytic Review. In: *Basic and Applied Social Psychology* 16: 515-527.

*Aquilino, William S.* 1998: Effects of Interview Mode on Measuring Depression in Younger Adults. In: *Journal of Official Statistics* 14: 15-29.

- Aquilino, William S.* 1994: Interview Mode Effects in Surveys of Drug and Alcohol Use. In: *Public Opinion Quarterly* 58: 210-240.
- Blair, Ed; Sudman, Seymour; Bradburn, Norman M.; Stocking, Carol* 1977: How to Ask Questions About Drinking and Sex: Response Effects in Measuring Consumer Behavior. In: *Journal of Marketing Research* 14: 316-321.
- Bochner, Stephen; van Zyl, Theresa* 1984: Desirability Ratings of 110 Personality-Trait Words. In: *The Journal of Social Psychology* 125: 459-465.
- Bradburn, Norman M.; Miles, Carrie* 1979: Vague Quantifiers. In: *Public Opinion Quarterly* 43: 92-101.
- Bradburn, Norman M.; Sudman, Seymour; Blair, Ed; Stocking, Carol* 1978: Question Threat and Response Bias. In: *Public Opinion Quarterly* 42: 221-234.
- Chen, Peter Y.; Dai, Tina; Spector, Paul E.; Jex, Steve M.* 1997: Relation Between Negative Affectivity and Positive Affectivity: Effects of Judged Desirability of Scale Items and Respondents' Social Desirability. In: *Journal of Personality Assessment* 69: 183-198.
- Cox, Brian J.; Swinson, Richard P.; Drenfeld, David M.; Bourdeau, Danielle* 1994: Brief Report. Social Desirability and Self-Reports of Alcohol Abuse in Anxiety Disorder Patients. In: *Behaviour Research and Therapy* 32: 175-178.
- Crott, Helmut W.; Roßbrucker, Karl* 1974: Erwünschtheit von Eigenschaften in Abhängigkeit von Alter, Geschlecht und Schichtzugehörigkeit. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 6: 241-261.
- Edwards, Allen L.* 1957: *The Social Desirability Variable in Personality Assessment and Research*. New York: The Dryden Press.
- Embree, Bryan G.; Whitehead, Paul C.* 1991: Validity and Reliability of Self-Reported Drinking Behavior: Dealing with the Problem of Response Bias. In: *Journal of Studies on Alcohol* 54: 334-344.
- Epstein, Joan F.; Barker, Peggy R.; Kroutil, Larry A.* 2001: Mode Effects in Self-Reported Mental Health Data. In: *Public Opinion Quarterly* 65: 529-549.
- Esser, Hartmut* 1991: Die Erklärung systematischer Fehler in Interviews: Befragtenverhalten als "rational choice". S. 59-78 in: *Wittenberg, Reinhard* (Hrsg.), *Person - Situation - Institution - Kultur*. Günter Büschges zum 65. Geburtstag. Berlin: Duncker & Humblot.
- Gove, Walter R.; Geerken, Michael R.* 1977: Response Bias in Surveys of Mental Health: An Empirical Investigation. In: *American Journal of Sociology* 82: 1289-1317.
- Hampson, Sarah E.; Goldberg, Lewis R.; John, Oliver P.* 1987: Category-Breadth and Social-Desirability Values for 573 Personality Terms. In: *European Journal of Personality* 1: 241-258.
- Huang, Chiou-yan; Liao, Hsin-ya; Chang, Sue-Hwang* 1998: Social Desirability and the Clinical Self-Report Inventory: Methodological Reconsideration. In: *Journal of Clinical Psychology* 54: 517-528.
- Jobe, Jared B.; Pratt, William F.; Tourangeau, Roger; Baldwin, Alison K.; Rasinski, Kenneth A.* 1997: Effects of Interview Mode on Sensitive Questions in a Fertility Survey. S. 311-329 in: *Lyberg, Lars; Biemer, Paul; Collins, Martin; De Leeuw, Edith; Dippo, Cathryn; Schwarz, Norbert; Trewin, Dennis* (Eds.), *Survey Measurement and Process Quality*. New York: John Wiley.
- Jones, Edward E.; Sigall, Harold* 1971: The bogus pipeline: A New Paradigm for Measuring Affect and Attitude. In: *Psychological Bulletin* 76: 349-364.
- Költringer, Richard; Gehmacher, Ernst; Klucsarits; Willibald* 1989: Response-Set-Messinstrumente: Zur Kontrolle systematischer Messfehler in der Umfrageforschung. In: *Österreichische Zeitschrift für Soziologie* 14: 34-45.

- Kozma, Albert; Stones, M. J.** 1988: Social Desirability in Measures of Subjective Well-Being: Age of Comparisons. In: *Social Indicators Research* 20: 1-14.
- Laicardi, Caterina; Baldassarri, Francesca; Artistic, Daniele** 2001: Unidimensionality of Life Satisfaction and its Relation to Social Desirability: A Confirmatory Study of a Short Form of the Life Satisfaction Scale. In: *Psychological Report* 88: 253-261.
- Locke, Shellie D.; Gilbert, Brenda O.** 1995: Method of Psychological Assessment, Self-Disclosure, and Experiential Differences: A Study of Computer, Questionnaire, and Interview Assessment Formats. In: *Journal of Social Behavior and Personality* 10: 255-263.
- Mangione, Thomas W.; Hingson, Ralph; Barrett, Jane** 1982: Collecting Sensitive Data. A Comparison of Three Survey Strategies. In: *Sociological Methods & Research* 10: 337-346.
- Mecklenbräuker, Silvia; Hager, Willi; Möller, Holger** 1994: Erwünschtheit von 908 Eigenschaften bei Frauen und bei Männern aus der Sicht von Männern und von Frauen. S. 310-328 in: **Hager, Willi; Hasselhorn, Marcus H.** *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.
- Meleddu, Mauro; Guicciardi, Marco** 1998: Self-Knowledge and Social Desirability of Personality Traits. In: *European Journal of Personality* 12: 151-168.
- Nicotera, Anne M.** 1996: An Assessment of the Argumentativeness Scale for Social Desirability Bias. In: *Communication Reports* 9: 23-36.
- Paulhus, Delroy L.; Reid, Douglas B.** 1991: Enhancement and Denial in Socially Desirable Responding. In: *Journal of Personality and Social Psychology* 60: 307-317.
- Phillips, Derek L.; Clancy, Kevin J.** 1970: Response Bias in Field Studies of Mental Illness. In: *American Sociological Review* 35: 503-514.
- Phillips, Derek L.; Clancy, Kevin J.** 1972: Some Effects of "Social Desirability" in Survey Studies. In: *American Journal of Sociology* 77: 921-940.
- Presser, Stanley; Stinson, Linda** 1998: Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance. In: *American Sociological Review* 63: 137-145.
- Singer, Eleanor** 1981: Reference Groups and Social Evaluations. S. 66-93 in: **Rosenberg, Morris; Turner, Ralph H.** (Eds.), *Sociological Perspectives*. New York: Basic Books.
- Smith, Seyda T.; Smith, Kyle D.** 1993: Social Desirability of Personality Items as a Predictor of Endorsement: A Cross-Cultural Analysis. In: *Journal of Social Psychology* 133: 43-53.
- Smith, Tom W.** 1992: Discrepancies Between Men and Women in Reporting Number of Sexual Partners: A Summary from Four Countries. In: *Social Biology* 39: 203-211.
- Stadt Mannheim** 1998: Elektronischer Datensatz über die Verteilung der Bevölkerungsmerkmale im Stadtgebiet, Stand 31.12.1998. Mannheim: Statistikstelle.
- Statistisches Bundesamt** 1998: Statistisches Jahrbuch für die Bundesrepublik Deutschland. Stand 31.12.1996. Wiesbaden: Online-Version:  
[http://infob1.uni-mannheim.de/IBplus/sessions/STATJB\\_DEUT98-72eca49b.html](http://infob1.uni-mannheim.de/IBplus/sessions/STATJB_DEUT98-72eca49b.html).
- Stocké, Volker** 2004: Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit. Ein Vergleich der Prognosen der Rational-Choice-Theorie und des Modells der Frame-Selektion. In: *Zeitschrift für Soziologie* (im Druck).
- Tourangeau, Roger; Smith, Tom W.; Rasinski, Kenneth A.** 1997: Motivation to Report Sensitive Behaviors on Surveys: Evidence From a Bogus Pipeline Experiment. In: *Journal of Applied Social Psychology* 27: 209-222.

*Turner, Charles F.; Ku, Leighton.; Rogers, Susan M.; Lindberg, Laura D.; Pleck, Joseph H.; Sonenstein, Freya L.* 1998: Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology. In: *Science* 280: 867-873.

*Turner, Charles F.; Lessler, Judith T.; Devore, James W.* 1992: Effects of Mode of Administration and Wording on Reporting of Drug Use. S. 177-219 in: *Turner, Charles F.; Lessler, Judith T.; Gfroerer, Joseph C.* (Eds.), *Survey Measurement of Drug Use. Methodological Studies.* Rockville: National Institute on Drug Abuse.

*Umesh, U. N.; Peterson, Robert A.* 1991: A Critical Evaluation of the Randomized Response Method: Applications, Validation, and Research Agenda. In: *Sociological Methods & Research* 20: 104-138.

*Waterton, Jennifer J.; Duffy, John C.* 1984: A Comparison of Computer Interviewing Techniques and Traditional Methods in the Collection of Self-Report Alcohol Consumption Data in a Field Survey. In: *International Statistical Review* 52: 173-182.

*Wyner, Gordon A.* 1980: Response Errors in Self-Reported Number of Arrests. In: *Sociological Methods & Research* 9: 161-177.

## Anhang

### Instrument zur Erfassung der Erwünschtheitswahrnehmungen der Befragten

Ich nenne Ihnen gleich eine Liste mit Verhaltensweisen oder Meinungen einer Person. Können Sie mir bitte jeweils sagen, ob diese Person deswegen von der Gesellschaft positiv, neutral oder negativ bewertet wird? Es geht also auch hier nicht darum, für wie wünschenswert Sie persönlich diese Eigenschaften halten, sondern was die Mehrheit der Bevölkerung in Deutschland darüber denkt. (Antwortskala wird vorgelegt). Sie können Ihre Meinung auf der vorliegenden Skala zwischen dem Wert minus 4, ‚gesellschaftlich sehr negativ bewertet‘, und plus vier, ‚gesellschaftlich sehr positiv bewertet‘, abstufen. Nennen Sie mir bitte immer den zutreffenden Skalenwert.

Wie wird in unserer Gesellschaft eine Person bewertet, die...

- |                    |  |
|--------------------|--|
|                    | ... sehr oft zum Blutspenden geht.                     |
| (Blutspenden)      | ... gelegentlich zum Blutspenden geht.                 |
|                    | ... nie zum Blutspenden geht.                          |
|                    | ... ein sehr glückliches Leben führt.                  |
| (Lebensglück)      | ... ein teilweise glückliches Leben führt.             |
|                    | ... überhaupt kein glückliches Leben führt.            |
|                    | ... sehr viel Zeit vor dem Fernseher verbringt.        |
| (Fernsehkonsument) | ... mittelmäßig viel Zeit vor dem Fernseher verbringt. |
|                    | ... nie Zeit vor dem Fernseher verbringt.              |
|                    | ... sehr häufig Alkohol trinkt.                        |
| (Alkoholkonsum)    | ... gelegentlich Alkohol trinkt.                       |
|                    | ... nie Alkohol trinkt.                                |

# Unterschiede zwischen Wertorientierungen nach Geschlecht, Alter, Berufsgruppe, Parteienpräferenz und Bundesland: Ergebnisse der Standardisierung des Österreichischen Wertefragebogens

von Walter Renner und Ingrid Salem<sup>1</sup>

## *Zusammenfassung*

*Ein neu entwickeltes österreichisches Werteinventar, welches die Skalen Intellektualität, Harmonie, Religiosität, Materialismus sowie Konservatismus und 13 Subskalen umfasst, wurde in einer explorativen Studie von einem Umfrageinstitut einer hinsichtlich der wichtigsten soziodemographischen Variablen repräsentativen, österreichischen Stichprobe (N = 421) vorgegeben. In nahezu allen Teilstichproben wurden Harmonie, Materialismus und Intellektualität deutlich stärker bevorzugt als religiöse und politische Werte. Frauen erzielten gegenüber Männern höhere Ausprägungen für Intellektualität, Harmonie und Religiosität und niedrigere Werte auf der Materialismusskala. Die Ergebnisse zu verschiedenen Altersgruppen widersprechen Ronald Ingleharts Theorie des Materialismus vs. Postmaterialismus, weil jüngere Personen im Vergleich zu älteren signifikant materialistischere Werte vertreten. Bei multivariater Signifikanzprüfung zeigten sich mit dem Alter zunehmende Religiositätswerte, während für Konservatismus keine signifikanten Unterschiede nach Altersgruppen gefunden wurden. Während Arbeiterinnen und Arbeiter gegenüber Angestellten Intellektualität niedriger bewerteten, unterschieden sich die beiden Gruppen nicht hinsichtlich Konservatismus, Beamtinnen und Beamte waren hingegen deutlich konservativer. Arbeiter, Angestellte und Selbständige orientierten sich im Vergleich zu Hausfrauen(-männern) und Pensionistinnen und Pensionisten stärker an materiellen Werten. Befürworter von ÖVP, SPÖ und FPÖ unterschieden*

---

<sup>1</sup> Dr. **Walter Renner** und Mag. **Ingrid Salem** sind Wissenschaftliche Mitarbeiter am Institut für Psychologie der Universität Klagenfurt. Universitätsstraße 65-67; A-9020 Klagenfurt; E-Mail: walter.renner@uni-klu.ac.at; Telefon: 0463 / 2700 / 1625; FAX: 0463 / 2700 / 1697

Die Untersuchung wurde vom Fonds zur Förderung der Wissenschaftlichen Forschung (FWF) im Rahmen des Forschungsprojekts P14966 finanziert.

*sich nicht hinsichtlich Konservatismus. Wählerinnen und Wähler der ÖVP waren aber durch hohe Ausprägungen von Religiosität, Intellektualität und Materialismus gekennzeichnet, jene der Grünen durch hohe Intellektualität neben geringem Konservatismus, Materialismus sowie niedriger Religiosität. Wien und Tirol sind durch ein hohes Maß an intellektueller Aufgeschlossenheit gekennzeichnet, auffallend sind für Tirol auch hohe Ausprägungen von Religiosität und niedrige von Materialismus. Aus den Ergebnissen wird gefolgert, dass soziodemographische Variablen einen ausgeprägt moderierenden Einfluss auf Wertorientierungen haben. Es wird vorgeschlagen, den neu entwickelten Fragebogen als zusätzliches Instrument in der Werteforschung weiter zu erproben.*

Schlüsselwörter: Wertorientierungen, Werthaltungen, Geschlecht, Alter, Beruf, Parteienpräferenz, Bundesland, Österreich.

### **Abstract**

*A newly developed Austrian questionnaire of human values which comprises the scales intellectualism, harmony, religiosity, materialism and conservatism including thirteen subscales was administered to a representative sample (N = 421) of the Austrian population. Mostly, harmony, materialism and intellectualism were rated higher than religious and political values. Women rated intellectualism, harmony and religiosity higher and materialism lower than men. Inglehart's theory of materialism vs. postmaterialism was not confirmed, as younger respondents valued materialism higher than older ones. Multivariate tests of significance showed that religiosity increased with age, whereas age groups did not differ significantly with respect to conservatism. Additional analyses yielded significant differences of value orientations for diverse professional groups and for voters of different political parties and in the various geographical parts of Austria. Therefore it is concluded that socio-demographic variables have a moderating influence on value orientations and it is suggested to introduce the new questionnaire as a research instrument.*

Werte sind persönliche und gesellschaftliche normative Standards, grundsätzliche Orientierungen oder Leitmotive (*Klages & Kmieciak*, 1981; *Kluckhohn*, 1951; *Rokeach*, 1973; *Scholl-Schaaf*, 1975; *Schwartz*, 1994). Wertorientierungen oder Werthaltungen sind persönliche, über die Zeit relativ stabile, Besonderheiten einer Person in der Bevorzugung von Werten; sie haben somit den Stellenwert von Persönlichkeitsmerkmalen (*Asendorpf*, 2004; *Kluckhohn*, 1951; *Scholl-Schaaf*, 1975). Sie werden durch Lernen bzw. im Zuge der Sozialisation erworben (*Scholl-Schaaf*, 1975; *Schwartz*, 1994) und sind daher kulturabhängig (*Hofstede*, 1984; *Triandis*, 1995).

Bisherige psychologische Messinstrumente zur Erfassung von Wertorientierungen wurden häufig in andere Sprachen übersetzt und gehen von der theoretischen Position aus, dass Werte kulturübergreifend (universell) gelten und erfragt werden können (z.B. *Allport, Vernon & Lindzey*, 1951, dt. *Roth*, 1972; *Rokeach*, 1973, dt. *Schneider*, 1983; *Schwartz*, 1992). Ohne die Berechtigung einer solchen kulturunabhängigen Werteforschung in Frage zu stellen, erscheint uns *zusätzlich* die Erstellung kulturspezifischer Erhebungsinstrumente gerechtfertigt.

Für Österreich legten *Denz, Friesl, Polak, Zuba* und *Zulehner* (2001) im Rahmen der österreichischen Wertestudie bedeutsame Ergebnisse einer repräsentativen Befragung vor, die für die letzten zehn Jahre Veränderungen in den Wertebereichen "Familie, Partnerschaft und Frauenrollen", "Arbeit", "Religion" sowie "Gesellschaft und Politik" belegen. Ein auf der Grundlage der Testtheorie entwickeltes, psychologisches Messinstrument (ein "Testverfahren"), das die kulturellen Gegebenheiten Österreichs berücksichtigt, existiert jedoch bislang nicht.

Ein solcher Test unterscheidet sich von anderen Erhebungsmethoden in erster Linie durch seine Normierung oder Standardisierung: bevor ein psychologischer Test in der Praxis eingesetzt wird, legt man ihn einer meist aus mehreren Hundert Personen bestehenden "Eichstichprobe" vor, welche für die Gesamtbevölkerung (oder andere Bezugsgruppen, etwa Jugendliche, Abiturientinnen und Abiturienten, Psychotherapiepatientinnen und -patienten) hinsichtlich der wichtigsten soziodemographischen Merkmale repräsentativ sein soll. Im nächsten Schritt werden die einzelnen Fragen ("Items"), aus denen der Test besteht, mittels statistischer Methoden auf ihre Eignung empirisch überprüft. So werden zum Beispiel Items, die von (nahezu) allen Personen gleich beantwortet wurden, die also schlecht zwischen den Personen differenzieren, ausgeschieden. Ebenso werden Testfragen eliminiert, die niedrige Korrelationen mit der Gesamtskala aufweisen und somit das gemessene Merkmal schlecht repräsentieren. Sobald nach Abschluss dieser "Itemanalyse" die endgültige Form des Tests vorliegt, erfolgt die Standardisierung oder Normierung. Das bedeu-

tet im einfachsten Fall, dass (wiederum anhand der Daten aus der Eichstichprobe) für die Testergebnisse Mittelwerte und Standardabweichungen berechnet werden. In der Praxis ermittelt man zusätzlich so genannte Prozenstränge (PR), welche angeben, wie viel Prozent der Einstichprobe von einer bestimmten Testperson mit ihrem Skalenwert übertroffen werden: somit bedeutet etwa ein  $PR = 80$ , dass 80% der vergleichbaren Gesamtbevölkerung (repräsentiert durch die Eichstichprobe) eine niedrigere und 20% eine höhere Ausprägung des betreffenden Leistungs- (z. B. Intelligenz, Merkfähigkeit) oder Persönlichkeitsmerkmals (z. B. Extraversion, Konservatismus) aufweisen. Sind die Merkmalsausprägungen in der Eichstichprobe normalverteilt, können zusätzlich zu den Prozensträngen noch weitere Normwerte bestimmt werden. In der Praxis bestehen Tests häufig aus mehreren "Skalen", die unterschiedliche Leistungs- bzw. Persönlichkeitsaspekte erfassen. Die genannten Analysen werden dann für jede Skala einzeln vorgenommen.

Psychologische Tests unterscheiden sich von anderen Erhebungsinstrumenten auch dadurch, dass mittels statistischer Kennwerte ihre Gütekriterien bestimmt werden können: die Objektivität eines Tests gibt an, in welchem Ausmaß seine Durchführung vom Anwender unabhängig ist. Sie wird mittels wiederholter Testvorgabe durch verschiedene Testleiter geprüft. Die Reliabilität ist mit der Messgenauigkeit des Tests gleichzusetzen und kann z. B. dadurch ermittelt werden, dass derselbe Test einer Personengruppe wiederholt vorgelegt wird. Die Validität gibt an, in welchem Maß ein Test tatsächlich misst, was er zu messen vorgibt. Sie wird unter anderem dadurch geprüft, dass der neue Test gemeinsam mit bereits etablierten Verfahren vorgegeben wird, wobei eine möglichst gute Übereinstimmung angestrebt wird. Weiterführende Informationen zur Vorgangsweise der "Klassischen Testtheorie" geben zum Beispiel *Bühner* (2004) sowie *Lienert* und *Raatz* (1998), eine sehr gute Darstellung der Grundlagen findet sich bei *Bortz* und *Döring* (1995).

Gegenüber den üblichen Befragungen im Rahmen der empirischen Sozialforschung bietet also ein psychometrisches Instrument – kurz auch als "Test" bezeichnet – die Möglichkeit, aufgrund seiner Normierung Aussagen über "den relativen Grad der individuellen Merkmalsausprägung" (*Lienert & Raatz*, 1998, S. 1) von Personen zu treffen, welche die Fragen nach Abschluss der Befragung der Eichstichprobe beantworten.

Die vorliegende Untersuchung berichtet von der Standardisierung des Österreichischen Wertfragebogens (ÖWF), wobei die Eichstichprobe durch ein Umfrageinstitut kontaktiert wurde. Dadurch war die Repräsentativität hinsichtlich Geschlecht, Alters- und Berufsgruppe, Bildungsniveau und geographischer Gegebenheiten

leicht erzielbar. Die einzelnen Konstruktionsschritte des ÖWF wurden in **Renner** (im Druck) dokumentiert und werden weiter unten zusammengefasst dargestellt.

Neben dem Ziel einer Standardisierung des ÖWF zielt die vorliegende Untersuchung auch darauf ab, einen Beitrag zu seiner Validierung zu leisten. Dabei wird davon ausgegangen, dass die mit dem neuen Instrument erzielten Resultate mit bisherigen empirischen Befunden konvergieren sollten. Nun könnte man in jedem derartigen Fall die Frage stellen, wozu ein neues Instrument entwickelt wurde, wenn seine Resultate mit bereits bestehenden – und womöglich angeblich "schlechteren" – Instrumenten übereinstimmen sollen. Gegen dieses Argument ist einzuwenden, dass der Österreichische Wertefragebogen keineswegs beabsichtigt, bisherige Forschungsansätze zur Seite zu drängen und für ungültig zu erklären; vielmehr war es uns ein Anliegen, diese durch ein exploratives Vorgehen zu *ergänzen*: In Abgrenzung zu der österreichischen Wertestudie von **Denz** et al. (2001) handelt es sich beim ÖWF um einen psychologischen "Test": die hier vorgestellte Repräsentativbefragung bildete die Grundlage für seine Standardisierung, so dass in weiterer Folge die Merkmalsausprägungen von Einzelpersonen oder Gruppen mit jenen der Eichstichprobe in Beziehung gesetzt werden können. Gegenüber kulturübergreifenden Testverfahren (z. B. **Schwartz**, 1992) betont der ÖWF kulturspezifische Gegebenheiten der österreichischen Gesellschaft, indem er zum Beispiel, im Gegensatz zum **Schwartz** Value Survey, über eine eigene Dimension für Religiosität verfügt. Dies bedeutet aber nicht, dass der ÖWF gänzlich "andere" Wertedimensionen abbildet als bisherige Erhebungsinstrumente. Vielmehr werden – neben kulturspezifischen Unterschieden – auch Gemeinsamkeiten erwartet. Im Zuge der Validierung des ÖWF versuchen wir somit zu überprüfen, inwieweit das Instrument *ähnliche* Aspekte von Wertorientierungen abdeckt wie sie durch bereits erfolgte Befragungen bzw. bereits bestehende psychometrische Instrumente erfasst wurden (zu bisherigen Validierungsschritten vgl. **Renner**, im Druck; **Renner**, **Salem** und **Alexandrowicz**, im Druck; **Salem** und **Renner**, im Druck).

Da Wertorientierungen im Zuge der Sozialisation erworben werden, liegt die Annahme nahe, dass diese mit soziodemographischen Variablen kovariieren und dass sich Männer und Frauen, verschiedene Alters- und Berufsgruppen und Personen unterschiedlicher politischer Orientierung von einander hinsichtlich des Ausmaßes der Bevorzugung oder Ablehnung bestimmter Werte unterscheiden. Auch regionale Unterschiede – in Österreich bezogen auf die einzelnen Bundesländer – sind plausibel.

Die Ergebnisse bisheriger Untersuchungen zu *Geschlechterunterschieden* in der Bevorzugung von Werten stimmen nur teilweise überein: Für Deutschland berichtet bereits **Nord-Rüdiger** (1985), dass es zu einer Nivellierung von Geschlechterdiffe-

renzen gekommen sei und sich Männer zunehmend an sozialen und gefühlsbezogenen, Frauen an materiellen und beruflichen Werten orientieren würden. **Prince-Gibson** und **Schwartz** (1998) fanden in einer israelischen Studie, dass sich Männer und Frauen in der Bevorzugung von Werten nicht unterschieden und auch die Bedeutungsstruktur von Werten für beide Geschlechter übereinstimmte. **Dietz, Kalof** und **Stern** (2002) fanden in den USA zwar keine Unterschiede in der faktoriellen Struktur der Wertorientierungen von Männern und Frauen, wohl aber, dass Frauen gegenüber Männern altruistische Werte stärker bevorzugen. In Übereinstimmung damit berichten **Badger, Craft** und **Jensen** (1998) ebenfalls in einer Studie aus den USA, dass weibliche Jugendliche beziehungs- und gefühlsorientierte und auf Kooperation bezogene Werte und männliche Jugendliche materielle und rationale Werte betonen. Auch **Di-Dio, Saragovi, Koestner** und **Aubé** (1996) legen dar, dass – in Übereinstimmung mit klassischen Geschlechterrollenstereotypen und sozialer Erwünschtheit – Frauen gemeinschaftsbezogene und Männer handlungsbezogene Werte bevorzugen. In einer weiteren U.S.-amerikanischen Untersuchung verglichen **Bennett, Stadt** und **Karmos** (1997) die von Männern und Frauen in den Jahren 1982 und 1992 bevorzugten Werte. Neben zahlreichen Übereinstimmungen zwischen den Geschlechtern fanden sie, dass Frauen 1992 Autonomie und Fortschritt bevorzugten, während sie 1982 altruistische Werte und Vielfalt für wichtiger gehalten hatten. Die von Männern bevorzugten Werte blieben weitgehend konstant, allerdings waren ihnen ästhetische Werte 1992 weniger wichtig als 1982. Für die USA berichten **McConatha** und **Schnell** (1997), dass sich Frauen, gegenüber Männern, stärker an Werten wie Gleichheit, Frieden und Selbstachtung orientieren.

**Inglehart** (1971, 1977) postulierte, dass sich die heute *älteren* Menschen an den für die Nachkriegsgeneration typischen "materialistischen" Werten (wirtschaftliche Absicherung, Verbrechensbekämpfung etc.) und die heute *jüngeren* Menschen an "postmaterialistischen" Werten (Mitbestimmung, Umweltschutz etc.) orientieren würden. Unterschiede in den Wertorientierungen verschiedener Altersgruppen können zunächst auf Kohorteneffekte zurückgeführt werden: wenn man mit **Ronald Inglehart** annimmt, dass sich Wertorientierungen früh im Leben etablieren und dann konstant bleiben, können Veränderungen in der Wertestruktur einer Gesellschaft nur durch nachrückende Generationen bewirkt werden. Hingegen behauptete unter anderem **Klages** (1987), dass sich individuelle Werte durchaus auch im Erwachsenenalter, etwa unter dem Einfluss gesellschaftlicher Veränderungen wandeln können. Schließlich gibt es Anhaltspunkte dafür, dass die mit dem Lebensalter wechselnde Bevorzugung bestimmter Werte durch Veränderungen im Lebenszyklus erklärt werden kann: so würden materielle Werte in der Mitte des Lebens am höchsten eingeschätzt, weil hier die ökonomische Verantwortung für die Familie im Vor-

dergrund steht, während Solidarität bei jungen und alten Menschen höher bewertet werde, weil diese im Gegensatz zu Menschen in der Lebensmitte in höherem Maß die Unterstützung anderer benötigen. Es gibt empirische Hinweise darauf, dass alle drei Erklärungsansätze berücksichtigt werden müssen und je nach untersuchtem Wertebereich in unterschiedlichem Maß zum Tragen kommen (*Klein*, 1991). – In Deutschland fand *Stiksrud* (1984) nur teilweise unterschiedliche Wertehierarchien für verschiedene Altersgruppen und konnte die theoretischen Annahmen *Ingleharts* nicht bestätigen, wonach die Nachkriegsgeneration "materialistische", die darauf folgende Generation "postmaterialistische" Werte bevorzugt. Auch *Hellevik* (2002) berichtet in einer norwegischen Untersuchung, dass junge Menschen heute Werte vertreten, die von materiellem Besitz und Konsum und nicht von postmaterialistischen Idealen gekennzeichnet sind. *Argue*, *Johnson* und *White* (1999) berichten, dass es zwischen dem 18. und dem 30. Lebensjahr zu einer Zunahme an Religiosität komme (zur Bedeutung von Jugendsekten siehe *Essmann*, 1988 und *Ottomeyer*, 1998). Alte Menschen weisen ein hohes Maß an Religiosität auf (*Stollberg*, 2002; vgl. *Utsch*, 1992), und sie sind auch deutlich konservativer als jüngere (*Truett*, 1993).

Über die von einzelnen *Berufsgruppen* bevorzugten Werte ist in der bisherigen Literatur wenig bekannt. *Kasser*, *Koestner* und *Lekes* (2002) fanden in einer Längsschnittstudie, dass niedriger sozioökonomischer Status in der Kindheit – häufig gepaart mit einem restriktiven Erziehungsstil – mit hoher Konformität im Erwachsenenalter einhergeht. In Übereinstimmung damit berichtet *Meulemann* (1997), dass in niedrigen sozialen Schichten Konformitätswerte, in höheren sozialen Schichten dagegen Autonomiewerte im Vordergrund stehen und die Unterschiede in der Bevorzugung unterschiedlicher Werte zwischen den Sozialschichten eher zu- als abnehmen.

Es ist plausibel anzunehmen, dass Wertorientierungen politische Einstellungen auf höherem Abstraktionsniveau repräsentieren. Wenn dies zutrifft, müssten sich Personen je nach ihrer *Präferenz für politische Parteien* auch hinsichtlich ihrer Wertorientierungen unterscheiden lassen. *Denz* (2001) fand in einer österreichischen Repräsentativbefragung, dass sich Wählerinnen und Wähler der Österreichischen Volkspartei (ÖVP), der Sozialdemokratischen Partei Österreichs (SPÖ) und der Freiheitlichen Partei Österreichs (FPÖ) von jenen der Grünen durch eine deutlich höhere Bevorzugung von "Ruhe und Ordnung" als Grundwert unterscheiden, während in der Bewertung von "Mitbestimmung", "Freiheit" und "Gleichheit" keine ausgeprägten Unterschiede bestehen. *Stromberg* und *Boehnke* (1997) untersuchten basierend auf der Wertetheorie von *Schwartz* (1992) den Zusammenhang von Werthaltungen und der Bevorzugung politischer Parteien in Deutschland. Sie leiteten daraus die Dimensionen "Ökonomischer Egalitarismus" und "Klassischer Libe-

ralismus" ab und entwickelten aufgrund soziologischer Expertenurteile Hypothesen über die einzelnen Parteien. So postulierten sie z.B., dass die Präferenz der Christlich Demokratischen Union (CDU) mit "'Bewahrung des Bestehenden' und 'Selbststärkung'" einhergehe und "in negativer Beziehung zu (...) 'Offenheit gegenüber Neuem' und 'Selbstüberwindung'" (S. 11) stehe. Sie fanden, dass Personen, welche die CDU bevorzugten, signifikant häufiger die Werte Konformität und Tradition sowie Macht befürworteten und Selbstbestimmung und Universalismus ablehnten. Wer Bündnis 90/Die Grünen bevorzugte, entschied sich signifikant häufiger für Selbstbestimmung und Universalismus und lehnte Macht, Leistung, Konformität und Sicherheit sowie Tradition ab. Für die übrigen Parteien ergaben sich keine signifikanten Unterschiede. – Auch *Barnea* und *Schwartz* (1998) fanden in Israel Zusammenhänge zwischen Wertorientierungen und Wählerverhalten. – *Braithwaite* (1998) geht von einem zweidimensionalen Wertemodell (Security vs. Harmony) unter Bezugnahme auf *Rokeach* (1973) und *Schwartz* (1992) aus und führt empirische Belege dafür an, dass Security mit politisch rechten und Harmony mit politischen linken Einstellungen korreliert.

Arbeiten, welche Unterschiede in den Wertorientierungen der Bevölkerung verschiedener österreichischer *Bundesländer* dokumentieren, sind uns nicht bekannt.

Die vorliegende explorative Untersuchung verfolgt die Zielsetzung, hinsichtlich der genannten soziodemographischen Variablen anhand des neu entwickelten Österreichischen Wertefragebogens Basisdaten zu erheben und gefundene Unterschiede zwischen den einzelnen Teilstichproben aufzuzeigen und zu interpretieren. Auf diese Weise sollte einerseits der Österreichische Wertefragebogen standardisiert werden, andererseits sollten Informationen darüber gewonnen werden, wie sich einzelne Bevölkerungsgruppen hinsichtlich ihrer Wertorientierungen unterscheiden. Dieses Wissen soll zu künftigen wissenschaftlichen Untersuchungen beitragen, aber auch die Praxis der psychologischen, psychotherapeutischen, pädagogischen, politischen und religiösen Arbeit bereichern.

## 1 Methode

**Untersuchungsteilnehmerinnen und -teilnehmer:** Die Stichprobe wurde von einem österreichischen Meinungsforschungsinstitut österreichweit rekrutiert und war hinsichtlich Geschlecht, Altersgruppe, Schulabschluss, Bundesland und Wohnortgröße repräsentativ entsprechend den jeweiligen Anteilen an der Gesamtbevölkerung. Tabelle 1 zeigt die soziodemographischen Statistiken der Stichprobe.

**Tabelle 1** Soziodemographische Statistiken der Stichprobe (N = 421)

	<i>Häufigkeit</i>	<i>Prozent</i>
<b><i>Geschlecht</i></b>		
männlich	200	47.5
weiblich	221	52.5
<b><i>Alter</i></b>		
18 - 30 Jahre	91	21.6
31 - 40 Jahre	96	22.8
41 - 50 Jahre	75	17.8
51 - 60 Jahre	64	15.2
61 - 70 Jahre	54	12.8
über 70 Jahre	41	9.7
<b><i>Schulabschluss</i></b>		
Pflichtschule	105	24.9
Berufsschule/Fachschule/Handelsschule	221	52.5
Matura/Universität/Fachhochschule	95	22.6
<b><i>Familienstand</i></b>		
ledig	128	30.4
verheiratet, Lebensgemeinschaft	215	51.1
geschieden, getrennt lebend	42	10.0
verwitwet	36	8.6
<b><i>Bundesland</i></b>		
Wien	87	20.7
Niederösterreich	80	19.0
Burgenland	9	2.1
Steiermark	63	15.0
Kärnten	38	9.0
Oberösterreich	66	15.7
Salzburg	32	7.6
Tirol	36	8.6
Vorarlberg	10	2.4
<b><i>Wohnortgröße</i></b>		
bis 5.000 Einwohner	168	39.9
bis 50.000 Einwohner	101	24.0
über 50.000 Einwohner	65	15.4
Wien	87	20.7

## 2 Vorgangsweise und Untersuchungsinstrument

Die Untersuchungsteilnehmerinnen und -teilnehmer waren vom Umfrageinstitut aufgrund von vorliegendem Adressmaterial ausgesucht worden und wurden für die aktuelle Befragung zu Hause aufgesucht. Der Österreichische Wertefragebogen wurde gemeinsam mit Untersuchungsinstrumenten zu religiösen, politischen und gesundheitsbezogenen Einstellungen (vgl. *Renner*, im Druck) vorgelegt und schriftlich beantwortet.

Der Österreichische Wertefragebogen wurde auf der theoretischen Basis des "Lexical Approach" konzipiert. Dieser ursprünglich auf *Francis Galton* zurückgehende persönlichkeitspsychologische Ansatz geht von der Annahme aus, dass interindivi-

duelle Unterschiede, die von zwischenmenschlicher und gesellschaftlicher Bedeutung sind, im Laufe der Zeit ihren Ausdruck in der Sprache gefunden haben, indem ein Wort zu ihrer Bezeichnung gebildet wurde. Durchsucht man ein Lexikon nach allen Begriffen, welche einen bestimmten Bereich interindividueller Unterschiede (etwa Persönlichkeitseigenschaften, körperliche Merkmale oder Wertorientierungen) beschreiben, erhält man eine nahezu erschöpfende Taxonomie dieser Merkmale. Durch nachfolgende Faktorenanalysen lassen sich die in der Taxonomie verzeichneten Begriffe auf wenige, im Allgemeinen orthogonale, Dimensionen reduzieren. Auf dieser theoretischen Basis wurde etwa das heute allgemein akzeptierte Fünf-Faktoren-Modell der Persönlichkeit konzipiert (vgl. *De Fruyt & Furnham*, 2000, *De Raad*, 2000, *John, Angleitner & Ostendorf*, 1988).

Die Zusammenstellung einer Taxonomie in Österreich relevanter Wertorientierungen wurde in *Renner* (2003) im Detail dargestellt. Anhand schriftlicher Richtlinien entnahmen zwei Beurteiler einem Lexikon der deutschen Sprache Listen von Substantiven und Adjektiven, die Wertorientierungen beschreiben. Sechs weitere Beurteiler reduzierten diese Listen zu 383 Hauptwörtern und 299 Eigenschaftswörtern, welche von 456 Österreicherinnen und Österreichern, überwiegend Studierenden, hinsichtlich ihrer Bedeutung als subjektive Leit motive beurteilt wurden. Die Faktorenanalyse dieser Ratings ergab fünf Faktoren, welche in einer zweiten Studie weitgehend repliziert werden konnten. Hier umfasste die Stichprobe 1160 Personen, welche nach Geschlecht, Altersgruppe und Bildungsniveau repräsentativ ausgewählt wurden (eine detaillierte Beschreibung der Stichprobe findet sich in *Renner*, im Druck). Anhand der in dieser Studie gewonnenen Daten wurden (ohne Berücksichtigung der Adjektive) Subdimensionen extrahiert und eine Itemanalyse vorgenommen, welche schließlich 54 "Factor Markers" ergab, die den Österreichischen Wertefragebogen konstituieren. Details zu dieser Untersuchung sind in *Renner* (im Druck) wiedergegeben. Wie dort ersichtlich, wurde die faktorielle Struktur des Fragebogens in einer weiteren Stichprobe bestätigt, welche mit jener identisch ist, die der vorliegenden Studie zugrunde liegt. Auch liegen erste Hinweise vor, welche für die Validität des neuen Instrumentes sprechen (*Renner*, im Druck; *Salem & Renner*, eingereicht).

Die fünf in beiden Studien gefundenen Faktoren sowie deren Subdimensionen lauten:

- *1. Intellektualität:* dieser Faktor umfasst Werte, welche mit geistiger Aufgeschlossenheit zu tun haben. Bei dem ersten Subfaktor (1.1 Weltoffenheit) liegt der Schwerpunkt bei einer offenen und toleranten Haltung gegenüber neuen Erkenntnissen und Einsichten, bei dem zweiten Subfaktor (1.2 Kultur) bei kulturellen Werten.

- *2. Harmonie:* Werte, die auf diesem Faktor laden, repräsentieren persönliche und zwischenmenschliche Ausgewogenheit. Der erste Subfaktor (2.1 Gemeinschaft) betont allgemein zwischenmenschliche, der zweite Subfaktor (2.2 Familie) familiäre Wärme. Subfaktor drei (2.3 Liebe zum Leben) hat persönliches Glückserleben, allerdings mit einer altruistischen Komponente, zum Gegenstand.
- *3. Religiosität:* Hier stehen spirituelle Werte im Vordergrund, wobei Subfaktor eins (3.1 Glaube) religiöse Verpflichtungen, Subfaktor zwei (3.2 Gnade) das Erleben von Zuwendung durch die Religion zum Gegenstand hat.
- *4. Materialismus:* hier stehen persönliche Vorteile im Vordergrund. Der erste Subfaktor (4.1 Eigentum) betont ökonomische, der zweite Subfaktor (4.2 Erfolg) auf berufliche Ziele bezogene Werte. Subfaktor drei (4.3 Genuss) umfasst hedonistische Werte, welche - im Gegensatz zu den auf Subfaktor 2.3 ladenden Begriffen - deutlich egozentrisch geprägt sind.
- *5. Konservatismus:* die auf diesem Faktor ladenden Begriffe sind durch gesellschaftliche Anpassung und Konformität gekennzeichnet. Der erste Subfaktor (5.1 Nationalismus) bezieht sich auf im politischen Sinn konservative und nationalistische Werte, Subfaktor zwei (5.2 Wehrhaftigkeit) hat Werte zum Gegenstand, welche die Verteidigung des Staates gegen mögliche Angreifer betonen. Der dritte Subfaktor (5.3 Pflichtbewusstsein) gilt schließlich der Beachtung und Befolgung gesellschaftlicher Verpflichtungen.

Wie alle faktorenanalytisch konzipierten dimensional Konzepte interindividueller Unterschiede umfasst der vorliegende Fragebogen nicht "sämtliche" relevanten Begriffe, weil nur "Factor Markers" mit ausreichend hohen Ladungen als Fragebogenitems in Betracht kommen. Aufgrund der Taxonomie<sup>2</sup> sind aber alle in der deutschen Sprache vorkommenden Wertebegriffe den Faktoren zuordenbar: so hat etwa "Freiheit" eine Ladung von  $a = 0,53$  auf Faktor II (Harmonie), wurde aber als Testitem nicht ausgewählt, weil andere Items höhere Ladungen aufwiesen. Das Item "Menschenrechte" hat eine substantielle Ladung von  $a = 0,47$  auf Faktor II (Harmonie), aber eine fast gleich hohe Ladung ( $a = 0,43$ ) auf Intellektualität. Aufgrund bedeutsamer Mehrfachladungen sind somit zahlreiche Begriffe als "Mischung" mehrerer Faktoren darstellbar. In Analogie dazu scheint etwa im Fünf-Faktoren-Modell der Persönlichkeit das Persönlichkeitsmerkmal "Aggressivität" zu "fehlen": es repräsentiert zwar keine selbständige Dimension, ist aber als Kombination von hohem Neurotizismus und geringer Verträglichkeit darstellbar.

---

2 Die Taxonomie ist auf Anfrage von den Autoren erhältlich.

**Tabelle 2** Skalen, Subskalen, Itemzuordnungen und Reliabilitäten (Cronbachs  $\alpha$ ) des Österreichischen Wertefragebogens

Skala (Subskala)	Items	Interne Konsistenz ( $\alpha$ )
<b>1. Intellektualität</b>		.86
1.1 Weltoffenheit	Erkenntnis, Erkenntnisfähigkeit, Sinn, Umsicht, Vielfalt, Völkerfreundschaft, Völkerverständigung, Weltoffenheit	.81
1.2 Kultur	Kultur, Kulturerbe, Kulturgut	.89
<b>2. Harmonie</b>		.85
2.1 Gemeinschaft	Gemeinsamkeit, Gemeinschaft, Gemeinschaftsgeist	.81
2.2 Familie	Elternliebe, Familiensinn, Friedensbereitschaft, Kindesliebe	.81
2.3 Liebe zum Leben	Lebensfreude, Lebenskraft, Liebe, Wohlbefinden	.77
<b>3. Religiosität</b>		.97
3.1 Glaube	Christlichkeit, Glaube, Glaubensfestigkeit, Glaubensstärke, Gottesglaube, Gottesgnade, Gottvertrauen, Religion, Religiosität	.97
3.2 Gnade	Gnade, Seelenheil, Vergebung	.83
<b>4. Materialismus</b>		.86
4.1 Eigentum	Vermögen, Wohlhabenheit, Wohlstand	.82
4.2 Erfolg	Aufstieg, Erfolg, Karriere	.78
4.3 Genuss	Genuss, Hochgenuss, Komfort, Stolz	.67
<b>5. Konservatismus</b>		.90
5.1 Nationalismus	Nationalbewusstsein, Nationalgefühl, Patriotismus, Tradition, Traditionsbewusstsein, Vaterlandsliebe	.88
5.2 Wehrhaftigkeit	Verteidigung, Verteidigungsbereitschaft	.91
5.3 Pflichtbewusstsein	Pflicht, Pflichterfüllung	.86

Die 54 Items des Österreichischen Wertefragebogens sind gemeinsam mit den Skalenzuordnungen und den Reliabilitäten (Cronbachs  $\alpha$ ) aus Tabelle 2 ersichtlich.

Die in der Tabelle angegebenen Reliabilitätskoeffizienten wurden in der aktuellen Untersuchung neu ermittelt. In der ursprünglichen Eichstichprobe waren Reliabilitäten in ähnlicher Höhe errechnet worden (*Renner*, im Druck). Jedes Item wurde mit einer kurzen Erklärung seiner Bedeutung und mit einer fünfstufigen Likertskala versehen, deren mögliche Ausprägungen wie folgt benannt und kodiert wurden: "Starke Ablehnung" (= 1), "Ablehnung" (= 2), "Neutral" (= 3), "Zustimmung" (= 4) sowie "Starke Zustimmung" (= 5).

Die Präferenz für politische Parteien wurde mit einem Stimmzettel erhoben, auf welchem anzugeben war, welche Partei gewählt würde, wenn am nächsten Sonntag Nationalratswahlen wären.

### 3 Statistische Methoden

Die vorliegende Untersuchung hat deskriptiven Charakter, weshalb komplexe statistische Verfahren nicht eingesetzt wurden. Es sei aber betont, dass die Verteilungen der Testpunktwerte aller Skalen signifikant von der Normalverteilung abweichen (Kolmogorov-Smirnov  $p < .01$  bzw.  $< .05$ ). Die Normierung des Fragebogens erfolgt daher ausschließlich anhand von Prozenträngen; soweit Gruppenvergleiche vorgenommen werden, wird zur Signifikanzprüfung auf parameterfreie Verfahren zurückgegriffen.

### 4 Ergebnisse

Tabelle 3 zeigt die Mittelwerte und Standardabweichungen für die Skalen und Subskalen, getrennt für Männer und Frauen.

**Tabelle 3** Deskriptive Statistiken nach Geschlecht

	Männer (N = 200)		Frauen (N = 221)		<i>p</i> (U-Test)
	$\bar{x}$	s	$\bar{x}$	s	
<b>1. Intellektualität</b>	3.73	0.48	3.85	0.48	< .01
1.1 Weltoffenheit	3.80	0.48	3.92	0.48	< .01
1.2 Kultur	3.56	0.73	3.67	0.66	n.s.
<b>2. Harmonie</b>	4.12	0.45	4.28	0.43	< .01
2.1 Gemeinschaft	3.89	0.64	4.08	0.63	< .01
2.2 Familie	4.22	0.56	4.41	0.51	< .01
2.3 Liebe zum Leben	4.20	0.51	4.30	0.52	< .05
<b>3. Religiosität</b>	3.08	0.82	3.36	0.82	< .01
3.1 Glaube	3.10	0.86	3.37	0.86	< .01
3.2 Gnade	3.04	0.77	3.34	0.79	< .01
<b>4. Materialismus</b>	3.89	0.46	3.77	0.49	< .01
4.1 Eigentum	3.99	0.53	3.91	0.58	n.s.
4.2 Erfolg	3.95	0.62	3.77	0.65	< .01
4.3 Genuss	3.77	0.54	3.66	0.53	< .05
<b>5. Konservatismus</b>	3.60	0.59	3.57	0.59	n.s.
5.1 Nationalismus	3.51	0.66	3.50	0.65	n.s.
5.2 Wehrhaftigkeit	3.51	0.86	3.40	0.81	n.s.
5.3 Pflichtbewusstsein	3.97	0.67	3.97	0.72	n.s.

Es zeigt sich, dass Frauen signifikant höhere Ausprägungen auf Skala 1 (Intellektualität) haben, was sich aber nur für die Subskala Weltoffenheit (1.1), nicht jedoch

für Kultur (1.2) bestätigt. Frauen weisen auch signifikant höhere Ausprägungen auf Skala 2 (Harmonie), sowie auf sämtlichen dazu gehörigen Subskalen auf. Dasselbe gilt für Skala 3 (Religiosität): Frauen sind signifikant religiöser als Männer. Jedoch haben Männer signifikant höhere Ausprägungen auf Skala 4 (Materialismus) und auf den dazu gehörigen Subskalen Erfolg (4.2) und Genuss (4.3). Subskala 4.1 (Eigentum) differenziert jedoch nicht zwischen Frauen und Männern. Auch die auf Skala 5 (Konservatismus) von Männern und Frauen erzielten Merkmalsausprägungen unterscheiden sich nicht signifikant.

In Tabelle 4 sind die von verschiedenen Altersgruppen erzielten Rohwerte wiedergegeben.

**Tabelle 4** Deskriptive Statistiken nach Altersgruppe

	18-30 Jahre		31-40 Jahre		41-50 Jahre		51-60 Jahre		61-70 Jahre		>70 Jahre		<i>p</i>
	(N = 91)		(N = 96)		(N = 75)		(N = 64)		(N = 54)		(N = 41)		
	$\bar{x}$	<i>s</i>	$\bar{x}$	<i>s</i>	Kruskal-Wallis								
<b>1. Intellektualität</b>	3.84	0.53	3.82	0.52	3.80	0.42	3.81	0.46	3.67	0.43	3.77	0.49	<i>n.s.</i>
1.1 Weltoffenheit	3.91	0.53	3.88	0.52	3.87	0.43	3.88	0.46	3.74	0.43	3.82	0.48	<i>n.s.</i>
1.2 Kultur	3.63	0.74	3.66	0.70	3.61	0.66	3.64	0.67	3.48	0.70	3.66	0.71	<i>n.s.</i>
<b>2. Harmonie</b>	4.23	0.51	4.21	0.41	4.26	0.44	4.15	0.52	4.15	0.36	4.20	0.40	<i>n.s.</i>
2.1 Gemeinschaft	4.02	0.71	3.98	0.59	4.09	0.62	3.86	0.68	3.90	0.63	4.05	0.59	<i>n.s.</i>
2.2 Familie	4.24	0.61	4.30	0.50	4.39	0.54	4.31	0.59	4.37	0.47	4.36	0.48	<i>n.s.</i>
2.3 Liebe zum Leben	4.36	0.52	4.29	0.49	4.27	0.51	4.21	0.59	4.11	0.43	4.15	0.56	< .05
<b>3. Religiosität</b>	2.99	0.75	3.17	0.82	3.18	0.78	3.35	0.86	3.25	0.83	3.76	0.84	< .05
3.1 Glaube	2.99	0.78	3.18	0.85	3.19	0.80	3.37	0.89	3.29	0.89	3.77	0.91	< .05
3.2 Gnade	3.01	0.77	3.14	0.80	3.16	0.77	3.28	0.80	3.14	0.79	3.71	0.70	<i>n.s.</i>
<b>4. Materialismus</b>	3.97	0.52	3.92	0.47	3.88	0.43	3.66	0.43	3.66	0.43	3.66	0.50	< .01
4.1 Eigentum	3.98	0.61	4.03	0.51	3.98	0.51	3.85	0.51	3.87	0.59	3.88	0.65	<i>n.s.</i>
4.2 Erfolg	4.08	0.67	3.97	0.65	3.94	0.57	3.69	0.54	3.59	0.58	3.58	0.65	< .01
4.3 Genuss	3.88	0.57	3.80	0.53	3.77	0.52	3.50	0.50	3.56	0.49	3.54	0.51	< .01
<b>5. Konservatismus</b>	3.46	0.66	3.53	0.58	3.62	0.55	3.52	0.55	3.66	0.51	3.96	0.56	<i>n.s.</i>
5.1 Nationalismus	3.40	0.71	3.44	0.63	3.50	0.62	3.41	0.67	3.60	0.54	3.91	0.62	<i>n.s.</i>
5.2 Wehrhaftigkeit	3.34	0.95	3.39	0.86	3.45	0.75	3.38	0.78	3.60	0.76	3.78	0.76	<i>n.s.</i>
5.3 Pflichtbewusstsein	3.75	0.74	3.96	0.69	4.16	0.59	3.97	0.75	3.88	0.58	4.28	0.67	<i>n.s.</i>

Die letzte Spalte der Tabelle zeigt die Ergebnisse der multivariaten Signifikanzprüfung<sup>3</sup>. Hier ist ersichtlich, dass die Ausprägungen der Skalen 1 (Intellektualität), 2 (Harmonie) und 5 (Konservatismus) nicht signifikant mit dem Alter variieren. Dies gilt auch für die zugehörigen Subskalen, mit Ausnahme von Skala 2.3 (Liebe zum Leben). Hier werden mit zunehmendem Alter niedrigere Werte erzielt. Mit zunehmendem Alter werden höhere Werte auf Skala 3 (Religiosität) erzielt, wobei sich die Altersgruppen nur hinsichtlich des Subfaktors 3.1 (Glaube), nicht aber bezüglich des Subfaktors 3.2 (Gnade) unterscheiden. Mit steigendem Alter werden auf Skala 4 (Materialismus) niedrigere Werte erzielt. Dies bestätigt sich auch für die Subskalen Erfolg (4.2) und Genuss (4.3).

Die Berufsgruppen (Arbeiterinnen/Arbeiter, Angestellte, Beamte, Selbständige, Landwirtinnen/Landwirte, Hausfrauen/männer, Pensionistinnen und Pensionisten) unterscheiden sich hinsichtlich des überwiegenden Teils der Skalen bei multivariater Prüfung signifikant voneinander<sup>4</sup>:

- Arbeiterinnen und Arbeiter erzielen mit  $\bar{x} = 3.63$  im Vergleich zu den anderen Berufsgruppen die niedrigsten Ausprägungen auf Skala 1 (Intellektualität) (z.B. Angestellte  $\bar{x} = 3.88$ , Beamtinnen und Beamte  $\bar{x} = 3.96$ ).
- Arbeiterinnen und Arbeiter ( $\bar{x} = 3.51$ ) unterscheiden sich von Angestellten ( $\bar{x} = 3.53$ ) nicht signifikant auf Skala 5 (Konservatismus).
- Beamtinnen und Beamte haben hohe Ausprägungen auf den Skalen Intellektualität ( $\bar{x} = 3.96$ , gegenüber z. B.  $\bar{x} = 3.63$  bei Arbeiterinnen und Arbeitern) und Konservatismus ( $\bar{x} = 3.74$ , gegenüber z. B.  $\bar{x} = 3.51$  bei Arbeiterinnen und Arbeitern).
- Hausfrauen (-männer) ( $\bar{x} = 3.69$ ) sowie Pensionistinnen und Pensionisten ( $\bar{x} = 3.65$ ) haben niedrige Ausprägungen auf Skala 4 (Materialismus) (gegenüber z. B. Arbeiterinnen und Arbeitern:  $\bar{x} = 3.88$ , Angestellten:  $\bar{x} = 3.93$  oder Selbständigen:  $\bar{x} = 3.97$ ).
- Arbeiterinnen und Arbeiter ( $\bar{x} = 2.94$ ) erzielen die niedrigsten, Pensionistinnen und Pensionisten ( $\bar{x} = 3.45$ ) die höchsten Ausprägungen auf Skala 3 (Religiosität).

---

3 Die Resultate der zusätzlich vorgenommenen Einzelvergleiche sind auf Anfrage von den Autoren erhältlich.

4 Aus Platzgründen werden die Ergebnisse zu den einzelnen Berufsgruppen, den Parteienpräferenzen und zu den österreichischen Bundesländern nur cursorisch dargestellt. Die Detailergebnisse sind, zusammen mit den Resultaten der multi- und univariaten Signifikanztests, auf Anfrage von den Autoren erhältlich.

Die Befürworterinnen und Befürworter der verschiedenen politischen Parteien unterscheiden sich hinsichtlich aller Skalen und nahezu aller Subskalen bei multivariater Prüfung signifikant voneinander:

- Auf Skala 1 (Intellektualität) weisen Wählerinnen und Wähler der Grünen ( $\bar{x} = 3.98$ ) und jene der ÖVP ( $\bar{x} = 3.90$ ) die höchsten Ausprägungen auf. SPÖ-Wählerinnen und -Wähler liegen mit  $\bar{x} = 3.71$  im mittleren Bereich, die niedrigsten Ausprägungen haben Personen, welche die FPÖ befürworteten ( $\bar{x} = 3.43$ ).
- Harmonie (Skala 2) wird am höchsten von Personen, die keine Angaben zur Parteienpräferenz machten ( $\bar{x} = 4.27$ ) und von Wählerinnen und Wählern der ÖVP ( $\bar{x} = 4.26$ ) bewertet, am niedrigsten von jenen der FPÖ ( $\bar{x} = 3.89$ ).
- Die höchsten Ausprägungen auf Skala 3 (Religiosität) haben Befürworterinnen und Befürworter der ÖVP ( $\bar{x} = 3.56$ ). Am niedrigsten liegen FPÖ ( $\bar{x} = 2.92$ ) und Grüne ( $\bar{x} = 2.90$ ). Wählerinnen und Wähler der SPÖ liegen dazwischen, aber mit  $\bar{x} = 3.29$  signifikant höher als jene der Grünen.
- Wählerinnen und Wähler der ÖVP ( $\bar{x} = 3.90$ ) liegen bei materialistischen Werten (Skala 4) gemeinsam mit Personen, die keine Angaben machten ( $\bar{x} = 3.94$ ) am höchsten. Befürworterinnen und Befürworter der Grünen liegen mit  $\bar{x} = 3.68$  am niedrigsten, unterscheiden sich aber nicht signifikant von SPÖ und ÖVP.
- Auf Skala 5 (Konservatismus) besteht kein signifikanter Unterschied zwischen Wählerinnen und Wählern von ÖVP ( $\bar{x} = 3.76$ ), FPÖ ( $\bar{x} = 3.83$ ) und SPÖ ( $\bar{x} = 3.60$ ). Die Grünen liegen mit  $\bar{x} = 3.10$  jedoch signifikant unter den anderen politischen Parteien.

Die neun österreichischen Bundesländer unterscheiden sich hinsichtlich der Ausprägungen der Skalen und Subskalen bei multivariater Prüfung voneinander signifikant:

- Auf Skala 1 (Intellektualität) weisen Wien und Tirol mit  $\bar{x} = 3.91$  die höchsten Ausprägungen auf, am niedrigsten liegt Kärnten ( $\bar{x} = 3.65$ ).
- Hinsichtlich Skala 2 (Harmonie) werden im Burgenland ( $\bar{x} = 4.37$ ) und in Niederösterreich ( $\bar{x} = 4.33$ ) die höchsten, in Oberösterreich ( $\bar{x} = 4.03$ ) die niedrigsten Ausprägungen erzielt.
- Die höchsten Ausprägungen auf Skala 3 (Religiosität) werden im Burgenland ( $\bar{x} = 3.53$ ) und in Tirol ( $\bar{x} = 3.51$ ) erzielt, am wenigsten religiös ist die Steiermark ( $\bar{x} = 2.92$ ). Wien liegt im mittleren Bereich ( $\bar{x} = 3.24$ ).
- Bezüglich Materialismus (Skala 4) erreicht das Burgenland mit  $\bar{x} = 4.07$  die höchsten, Tirol ( $\bar{x} = 3.69$ ) die niedrigsten Skalenwerte.
- Im Burgenland ( $\bar{x} = 3.88$ ) werden die höchsten, in Tirol ( $\bar{x} = 3.07$ ) und Wien ( $\bar{x} = 3.48$ ) die niedrigsten Ausprägungen auf der Dimension Konservatismus (Skala 5) erzielt.

Im Anhang sind Normen wiedergegeben, die es ermöglichen, von Einzelpersonen oder Gruppen erzielte Resultate im Österreichischen Wertefragebogen unter Berücksichtigung von Geschlecht und Alter zu den Daten der Eichstichprobe in Relation zu setzen.<sup>5</sup>

## 5 Zusammenfassung und Diskussion

Zunächst ist festzuhalten, dass Wertorientierungen, die auf innere und äußere Ausgewogenheit (Skala 2), materiellen Vorteil (Skala 4) oder Intellektualität (Skala 1) und somit im weitesten Sinn auf persönliches und gemeinschaftlich-familiäres Wohlergehen gerichtet sind, in fast allen untersuchten Teilstichproben stärker befürwortet werden als die ideologischen Werte Religiosität (Skala 3) und Konservatismus (Skala 5). Besonders bei religiösen Werten fällt eine Tendenz zur Indifferenz auf, wobei allerdings die hohe Standardabweichung darauf hindeutet, dass neben neutralen Bewertungen auch zahlreiche deutlich positive sowie viele ausgeprägt negative Stellungnahmen erfolgten. Religiöse Werte, wie sie mit diesem Fragebogen erfasst werden, dürften, von den Assoziationen her, die sie wecken, auf eine traditionell-kirchliche Form von Religiosität bezogen sein. *Friessl* und *Zuba* (2001) entwarfen für Österreich eine "sozioreligiöse Typologie" (S. 122), wonach je nach Gottesbild und Kirchengangshäufigkeit ein kirchlicher, ein kulturkirchlicher, religiöser, kulturreligiöser und unreligiöser Typ unterschieden werden; dabei sind "kulturreligiöse" Personen am häufigsten, das sind solche, die an "irgendein höheres Wesen" glauben, aber nur selten oder nie den Gottesdienst besuchen. Nur der "kirchliche" Typ, welcher den Glauben an einen personalen Gott mit regelmäßigem Kirchenbesuch verbindet, dürfte auf der vorliegenden Religiositätsskala und ihren Subskalen hohe Ausprägungen erzielen. Des weiteren muss bedacht werden, dass sich nach *Denz* et al. (2001) zahlreiche Personen an mehreren verschiedenen religiösen Ideologien orientieren ("Bricolagenbildung", S. 139). Auch diesen Fällen wird die vorliegende Religiositätsskala nur wenig gerecht.

Entsprechend unserer Vorannahme wurden hinsichtlich aller soziodemographischen Variablen und auf den meisten Skalen und Subskalen des Österreichischen Wertefragebogens signifikante Unterschiede gefunden.

Das Ergebnis, wonach Frauen auf Skala 2 (Harmonie) höhere und auf Skala 4 (Materialismus) niedrigere Ausprägungen als Männer erzielen, stimmt mit den Befun-

---

5 Weitere Normierungsdaten, etwa unter Berücksichtigung der Berufsgruppe, der Parteienpräferenz, des Bundeslandes, des Religionsbekenntnisses oder der Wohnortgröße sind von den Autoren auf Anfrage erhältlich.

den von *Badger* et al. (1998), *Dietz* et al. (2002) und *Di-Dio* et al. (1996) überein. Zusätzlich fanden wir bei Frauen aber höhere Weltoffenheit (Skala 1.1) und Religiosität (Skala 3) als bei Männern. Alle geschlechterspezifischen Unterschiede spiegeln die von *Di-Dio* et al. (1996) angesprochenen Geschlechterrollenstereotype wieder. Inwieweit es im Laufe der Zeit auch in Österreich zu der von *Nord-Rüdiger* (1985) und *Prince-Gibson* und *Schwartz* (1998) beobachteten Nivellierung der Geschlechterunterschiede kommt, müssen Längsschnittuntersuchungen klären.

Die vorliegenden Ergebnisse unterstützen die theoretischen Annahmen von *Inglehart* (1971, 1977) nicht. Bei der Interpretation des Ergebnisses muss berücksichtigt werden, dass die von *Ronald Inglehart* zur Erfassung von Materialismus und Postmaterialismus verwendeten Items eher politische Einstellungen bzw. Zielsetzungen als Wertorientierungen erfassen (z.B. Kampf gegen steigende Preise, Verschönerung der Städte...); mit dieser Einschränkung wäre aber zu erwarten, dass materialistische Einstellungen mit materialistischen (Skala 4) und konservativen (Skala 5) Wertorientierungen und postmaterialistische Einstellungen mit intellektuellen Wertorientierungen (Skala 1) einhergehen. Die vorliegende Studie zeigt aber, dass Werte der Skala 4 (Materialismus) von jüngeren Altersgruppen signifikant höher bewertet werden als von älteren und die Skalen 1 (Intellektualität) und 5 (Konservatismus) keine signifikanten Altersunterschiede aufweisen (lediglich der Einzelvergleich der Gruppen der 61-70jährigen und der >70jährigen ergibt für die ältere Gruppe signifikant höhere Ausprägungen der Skalen 5, 5.1 und 5.3). Die mit dem Alter fallenden Ausprägungen von Skala 4 (Materialismus) bestätigen die norwegischen Befunde von *Hellevik* (2002) ebenso wie jene von *Stiksrud* (1984) für Deutschland: Besitz und Konsum materieller Güter hat für junge Menschen im heutigen Österreich größere Bedeutung als für ältere. Erklärbar ist dies, ebenso wie die mit dem Alter abnehmende mittlere Skalenausprägung von Skala 2.3 (Liebe zum Leben) durch die im Vergleich zu älteren Menschen bessere ökonomische und gesellschaftliche Situation der Jüngeren. Freilich werden auch intellektuelle Werte, also solche, die *Ronald Ingleharts* Postmaterialismus nahe kommen, von jungen Menschen hoch bewertet. *Denz* (2001), der in der Österreichischen Wertestudie am häufigsten den Mischtyp des "Überwiegend Materialisten" fand, kommentiert dies so: "Am liebsten möchte man alles haben – hohes Einkommen, viel Luxus, ein schnelles Auto und viel Natur, reine Luft usw. Die These vom additiven Wertewandel dürfte tatsächlich zutreffen" (S. 218). Die alltagspsychologische Vorstellung von mit zunehmendem Alter wachsender Religiosität ist jedenfalls zu relativieren: für Subskala 3.2 (Gnade) wurden bei multivariater Prüfung keine signifikanten Altersunterschiede gefunden.

Beim Vergleich der Berufsgruppen ist zunächst festzuhalten, dass die für die Gruppe der Landwirte erzielten Resultate aufgrund des sehr kleinen Stichprobenumfangs nicht verlässlich interpretierbar sind. Auch die Ergebnisse jener Gruppe, die zum Beruf keine Angaben machte, sind schwer zu interpretieren. Obwohl anzunehmen ist, dass sich zahlreiche Studierende in dieser Teilstichprobe befanden, dürfte diese Gruppe doch heterogen sein. Im übrigen lassen sich die Befunde von *Kasser* et al. (2002) und *Meulemann* (1997), wonach niedrige soziale Schichten Konformitätswerte und höhere Schichten Autonomiewerte bevorzugen, nur teilweise bestätigen: zwar erzielten Arbeiterinnen und Arbeiter die niedrigsten Ausprägungen auf Skala 1 (Intellektualität), welche Autonomiewerte repräsentiert, sie unterscheiden sich von Angestellten aber nicht auf Skala 5 (Konservatismus). Es fällt auf, dass Personen-  
gruppen mit eher geringen ökonomischen Entfaltungsmöglichkeiten (Hausfrauen/-männer sowie Pensionistinnen und Pensionisten) im Vergleich zu anderen Bevölkerungsgruppen niedrigere Materialismuskwerte aufweisen. Die besonders geringe Ausprägung von Religiosität (Skala 3) in der Gruppe der Arbeiterinnen und Arbeiter entspricht der alltagspsychologischen Erwartung ebenso wie die höhere Religiosität der Pensionistinnen und Pensionisten, welche die deutlich religiösere Altersgruppe der >70jährigen mit umfasst.

Die Ergebnisse zum Wählerverhalten zeigen, dass sich die Wählerinnen und Wähler der SPÖ hinsichtlich der Bevorzugung konservativer Werte weder von denen der ÖVP noch von jenen der FPÖ signifikant unterscheiden; auch ÖVP und FPÖ differieren diesbezüglich nicht voneinander. Nur im Vergleich mit den Grünen erzielen die drei anderen Parlamentsparteien höhere Konservatismuskwerte. Dieses Ergebnis stimmt mit dem Bericht von *Denz* (2001) überein, wonach "Ruhe und Ordnung" von den Wählerinnen und Wählern der ÖVP, SPÖ und FPÖ ähnlich hoch und nur von jenen der Grünen niedriger bewertet werden. Daraus ist ableitbar, dass die Entscheidung, ÖVP, SPÖ oder FPÖ zu wählen – im Gegensatz zu einer Entscheidung für die Grünen – nicht mit mehr oder weniger konservativen Wertorientierungen zu tun hat.

Die Wählerinnen und Wähler von ÖVP, SPÖ und FPÖ lassen sich somit nur noch schlecht auf einem politischen Links vs. Rechts-Kontinuum positionieren; klarer ist die Differenzierung hinsichtlich religiöser Werte: diese werden von den Anhängern der ÖVP vertreten und von jenen der Grünen und der FPÖ im Durchschnitt eher abgelehnt. Nimmt man die deutsche CDU als Analogon zur ÖVP, die SPD als Analogon zur SPÖ, und Bündnis 90/Die Grünen zu den österreichischen Grünen, so bestätigen sich die Befunde von *Stromberg* und *Boehnke* (1997) insofern, als die Wählerinnen und Wähler von ÖVP und Grünen sehr markant durch die von ihnen bevorzugten Werte gekennzeichnet sind; die Wertestruktur der Anhänger der österreichischen Grünen ist jener der deutschen Grünen vergleichbar, und die Wähler

von CDU und ÖVP befürworten gleichermaßen konservative Werte; anders als die Anhänger der CDU befürworten österreichische ÖVP-Wähler aber im Durchschnitt universalistische und intellektuelle Werte (Skala 1). Erklärbar ist dies dadurch, dass konservative Personen, welche die zuletzt genannten Werte weniger hoch einschätzen, in Österreich tendenziell FPÖ wählen. Die Wählerinnen und Wähler der SPÖ sind am wenigsten klar durch eine typische, von den anderen Parteien abgehobene, Wertestruktur charakterisiert.

Die für die einzelnen Bundesländer ermittelten Ergebnisse sind mit Vorsicht zu interpretieren: insbesondere sind die Stichproben in Vorarlberg und im Burgenland zu klein, um Zufallseffekte ausschließen zu können. Das Ergebnis stimmt mit dem allgemein verbreiteten Bild vom religiösen Tirol überein, während die Resultate für Wien nicht bestätigen, dass die Bewohnerinnen und Bewohner hinsichtlich ihrer Religiosität "abgeschlagen am Ende der Tabelle" liegen, wie *Friessl* und *Zuba* (2001, S. 121) konstatieren. Diese Divergenz könnte mit den unterschiedlichen Erhebungsinstrumenten zusammenhängen: obwohl religiöse Wertorientierungen, wie sie mit dem Österreichischen Wertefragebogen erhoben werden, mit religiösen Einstellungen, gemessen mit der Religiositätsskala von *Kecskes* und *Wolf* (1993), hoch korrelieren ( $r = .78$ , *Renner*, im Druck), sollte weitere Forschung klären, inwieweit instrumentenspezifische Faktoren bei der Erhebung religiöser Wertorientierungen und Einstellungen eine Rolle spielen. Im vorliegenden Fall wurde der Fragebogen zwar vom Meinungsforschungsinstitut vorgelegt, dann aber schriftlich beantwortet, während bei herkömmlichen Umfragen die Items im allgemeinen mündlich beantwortet werden.

Zusammenfassend ist festzuhalten, dass in der vorliegenden Untersuchung zahlreiche in der Literatur berichtete Ergebnisse über unterschiedliche Wertepreferenzen in Abhängigkeit von soziodemographischen Variablen bestätigt werden konnten. Dies spricht für die konvergente Validität des Österreichischen Wertefragebogens. Derartige Übereinstimmungen eines neu entwickelten Instruments führen häufig zu der Frage nach seiner Notwendigkeit bzw. Existenzberechtigung. Im Falle des ÖWF besteht, wie eingangs erwähnt, zu den Untersuchungen von *Denz* et al. (2001) der grundsätzliche Unterschied, dass es sich bei dem Österreichischen Wertefragebogen um ein psychometrisches Instrument handelt, welches in der psychologischen Einzelfalldiagnostik eingesetzt werden kann. Gegenüber bestehenden psychometrischen Instrumenten bietet der ÖWF den Vorteil, speziell für Österreich entwickelt worden zu sein.

Man könnte kritisch einwenden, dass der ÖWF von vornherein in Kooperation mit der Arbeitsgruppe von *Denz* et al. (2001) hätte erstellt werden sollen. Im Gegensatz

dazu wurde er auf der Basis des lexikalischen Ansatzes mit Methoden der Persönlichkeitspsychologie unabhängig entwickelt. Die Tatsache, dass die so gefundenen Wertedimensionen mit jenen von **Denz** et al. (2001) gut übereinstimmen (Harmonie entspricht "Familie/Partnerschaft/Frauenrollen", Religiosität entspricht "Religion", Materialismus entspricht "Arbeit" und Konservatismus entspricht "Gesellschaft und Politik"), bestätigt im Nachhinein die Validität *beider* Forschungsansätze gerade deshalb, weil sie voneinander unabhängig vorgingen.

In Ergänzung zu den Resultaten von **Denz** et al. (2001) und zu bestehenden kulturübergreifenden Instrumenten will der ÖWF einen Beitrag zur Werteforschung in Österreich leisten. Wir empfehlen, die Resultate der vorliegenden explorativen Studie zum Ausgangspunkt weiterer Forschung zu machen.

## LITERATUR

- Allport, G. W., Vernon, P. E. & Lindzey, G.** (1951). *Study of values* (revised edition). Boston: Houghton Mifflin.
- Argue, A., Johnson, D. R. & White, L. K.** (1999). Age and religiosity: Evidence from a three-wave panel analysis. *Journal for the Scientific Study of Religion*, 38 (3), 423-435.
- Asendorpf, J.** (2004). *Psychologie der Persönlichkeit*. (3. Aufl.). Berlin u.a.: Springer.
- Badger, K., Craft, R. S. & Jensen, L.** (1998). Age and gender differences in value orientation among American adolescents. *Adolescence*, 33 (131), 591-596.
- Barnea, M. F. & Schwartz, S. H.** (1998). Values and Voting. *Political Psychology*, 19 (1), 17-40.
- Bennett, W., Stadt, R. & Karmos, J.** (1997). Values preferences by gender for nontraditional college students between 1992 and 1982. *Counseling and Values*, 41 (3), 246-252.
- Bortz, J. & Döring, U.** (1995). *Forschungsmethoden und Evaluation* (2. Aufl.). Berlin u.a.: Springer.
- Braithwaite, V.** (1998). The value balance model of political evaluations. *British Journal of Psychology*, 89, 223-247.
- Bühner, M.** (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- De Fruyt, F. & Furnham, A.** (2000). Advances in the assessment of the five factor model. *Psychologica Belgica*, 40, 51-75.
- Denz, H.** (2001). Staat und Zivilgesellschaft: Widersprüche, Verwerfungen, Bruchlinien. In **H. Denz, C. Friesl, R. Polak, R. Zuba & P. M. Zulehner** (Hrsg.), *Die Konfliktgesellschaft. Wertewandel in Österreich 1990-2000* (S. 169-243). Wien: Czernin.
- Denz, H., Friesl, C., Polak, R., Zuba, R. & Zulehner, P. M.** (2001). *Die Konfliktgesellschaft. Wertewandel in Österreich 1990-2000*. Wien: Czernin.
- De Raad, B.** (2000). *The Big Five personality factors*. Seattle: Hogrefe.
- Di-Dio, L., Saragovi, C., Koestner, R. & Aubé, J.** (1996). Linking personal values to gender. *Sex Roles*, 34 (9-10), 621-636.
- Dietz, T., Kalof, L. & Stern, P. C.** (2002). Gender, values and environmentalism. *Social Science Quarterly*, 83 (1), 353-364.

- Essmann, K. R.** (1988). "Sekten", Neue religiöse Bewegungen, New Age. Menschen auf der Suche nach Sinn. *Dialog*, 4 (3): Beilage: Dialog Spezial, 1-11.
- Friesl, C. & Zuba, R.** (2001). Die ÖsterreicherInnen und die Religion. In **H. Denz, C. Friesl, R. Polak, R. Zuba & P. M. Zulehner** (Hrsg.), *Die Konfliktgesellschaft. Wertewandel in Österreich 1990-2000* (S. 99-168). Wien: Czernin.
- Hellevik, O.** (2002). Age differences in value orientation: Life cycle or cohort effects? *International Journal of Public Opinion Research*, 14 (3), 286-302.
- Hofstede, G.** (1984). *Culture's consequences. International differences in work related values*. (2nd ed.). Beverly Hills, CA: Sage.
- Inglehart, R.** (1971). The silent revolution in Europe: Intergenerational change in postindustrial societies. *American Political Science Review*, 65, 991-1017.
- Inglehart, R.** (1977). *The silent revolution: Changing values and political styles among Western publics*. Princeton, NJ: Princeton University Press.
- John, O. P., Angleitner, A., & Ostendorf, F.** (1988). The lexical approach to personality: a historical review of trait taxonomic research. *European Journal of Personality*, 2, 171-203.
- Kasser, T., Koestner, R. & Lekes, N.** (2002). Early family experiences and adult values: A 26-year, prospective longitudinal study. *Personality and Social Psychology Bulletin*, 28 (6), 826-835.
- Kecskes, R. & Wolf, C.** (1993). Christliche Religiosität: Konzepte, Indikatoren, Meßinstrumente. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 45, 270-287.
- Klages, H.** (1987). Sozialpsychologie der Wohlfahrtsgesellschaft: Konturen eines Wissenschaftsprogramms. In **H. Klages & W. Herbert** (Hrsg.), *Sozialpsychologie der Wohlfahrtsgesellschaft. Zur Dynamik von Wertorientierungen, Einstellungen und Ansprüchen*. Frankfurt u.a.: Campus.
- Klages, H. & Kmieciak, P.** (1981). Einführung. In **H. Klages & P. Kmieciak** (Hrsg.) *Wertewandel und gesellschaftlicher Wandel* (2. Aufl.) (S. 11-19). Frankfurt, New York: Campus.
- Klein, T.** (1991). Zur Bedeutung von Alters-, Perioden- und Generationseinflüssen für den Wandel politischer Werte in der Bundesrepublik. *Zeitschrift für Soziologie*, 20 (2), 138-146.
- Kluckhohn, C.** (1951). Values and value-orientations in the theory of action. An exploration in definition and classification. In **T. Parsons & E. A. Shils** (Hrsg.) *Toward a general theory of action* (S. 388-433). New York: Harper & Row.
- Lienert, G. A. & Raatz, U.** (1998). *Testaufbau und Testanalyse* (6. Aufl.). München: Psychologie Verlags-Union.
- McConatha, J. T. & Schnell, F.** (1997). A cross-generational and gender comparison of value priorities of adults living in the U. S. *Psychology: A Journal of Human Behavior*, 34 (1), 27-35.
- Meulemann, H.** (1997). Werte wandeln sich - aber nicht ihre Unterschiede zwischen sozialen Schichten. Erziehungsziele in Bildungs- und Berufsgruppen der alten Bundesrepublik zwischen den fünfziger und neunziger Jahren. *Theorie und Praxis der sozialen Arbeit*, 48 (2), 12-22.
- Nord-Rüdiger, D.** (1985). Nivellierung der Geschlechtsunterschiede? Lebensziele und Wertorientierungen jugendlicher Berufstätiger. In **A. Stiksrud & F. Wobit** (Hrsg.), *Adoleszenz und Präadoleszenz. Beiträge zur angewandten Jugendpsychologie* (S. 199-210). Eschenborn: Fachbuchhandlung für Psychologie.
- Ottomeyer, K.** (1998). Psychotherapie, Religion und New Age. *Journal für Psychologie*, 6 (4), 16-29.
- Prince-Gibson, E. & Schwartz, S. H.** (1998). Value priorities and gender. *Social Psychology Quarterly*, 61 (1), 49-67.

- Renner, W.** (2003). Human values: a lexical perspective. *Personality and Individual Differences*, 34, 127-141.
- Renner, W.** (im Druck). A German value questionnaire developed on a lexical basis: construction and steps toward a validation. *Review of Psychology*.
- Renner, W., Salem, I. & Alexandrowicz, R.** (im Druck). Human values as predictors for political, religious and health related attitudes: a structural equation approach to validate the Austrian Value Questionnaire (AVQ). *Social Behavior and Personality*.
- Rokeach, M.** (1973). *The nature of human values*. New York: Free Press.
- Roth, E.** (1972). *Der Werteinstellungstest. Eine Skala zur Messung dominanter Interessen der Persönlichkeit*. Bern u. a.: Huber.
- Salem, I. & Renner, W.** (im Druck). Do human values reflect job decisions, pro-social and anti-social behavior? Findings from the Austrian Value Questionnaire. *Psychological Reports*.
- Schneider, M.** (1983). *Werte und Persönlichkeit*. Frankfurt/M.: Haag und Herchen.
- Scholl-Schaaf, M.** (1975). *Werthaltung und Wertsystem. Ein Plädoyer für die Verwendung des Wertkonzeptes in der Sozialpsychologie*. Bonn: Bouvier.
- Schwartz, S. H.** (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1-65.
- Schwartz, S. H.** (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, 50 (4), 19-45.
- Stiksrud, H. A.** (1984). Gibt es einen Generationen-Dissens? Empirische Untersuchungen zu Wertangstrepanzen bei Personen unterschiedlichen Alters. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 31 (1), 153-174.
- Stollberg, D.** (2002). Glaubensgewissheit und Lebenszufriedenheit im Alter. Pastoralpsychologische Beobachtungen, Deutungen und Vermutungen. In **H. D. Basler & S. Keil** (Hrsg.), *Lebenszufriedenheit und Lebensqualität im Alter* (S. 251-266). Graftschaft: Vektor.
- Stromberg, C. & Boehnke, K.** (1997). Parteienpräferenzen in Deutschland: Welchen Einfluß haben individuelle Werthaltungen? *Zeitschrift für Politische Psychologie*, 5, 5-20.
- Triandis, H. C.** (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.
- Truett, K. R.** (1993). Age differences in conservatism. *Personality and Individual Differences*, 14 (3), 405-411.
- Utsch, M.** (1992). Religiosität im Alter: Forschungsschwerpunkte und methodische Probleme. *Zeitschrift für Gerontologie*, 25 (1), 25-31.

## 6 Anhang

**Tabelle 5** Normierung des Österreichischen Wertefragebogens nach Geschlecht und Altersgruppe

	1.	1.1	1.2	2.	2.1	2.2	2.3	3.	3.1	3.2	4.	4.1	4.2	4.3	5.	5.1	5.2	5.3
<i>GESCHLECHT</i>																		
<i>Männer</i>																		
PR 25	3.36	3.50	3.00	3.91	3.33	4.00	4.00	2.75	2.67	2.67	3.60	3.67	3.67	3.50	3.20	3.00	3.00	3.50
PR 50	3.73	3.75	3.67	4.18	4.00	4.25	4.25	3.00	3.00	3.00	3.90	4.00	4.00	3.75	3.60	3.50	3.50	4.00
PR 75	4.09	4.00	4.00	4.45	4.33	4.75	4.50	3.58	3.78	3.33	4.20	4.33	4.33	4.00	4.00	4.00	4.00	4.50
<i>Frauen</i>																		
PR 25	3.55	3.63	3.00	4.00	3.67	4.00	4.00	3.00	3.00	3.00	3.40	3.67	3.33	3.25	3.20	3.00	3.00	3.50
PR 50	3.82	3.88	4.00	4.27	4.00	4.50	4.25	3.25	3.33	3.33	3.70	4.00	3.67	3.75	3.60	3.50	3.50	4.00
PR 75	4.18	4.25	4.00	4.64	4.67	4.75	4.75	3.92	4.00	4.00	4.10	4.33	4.00	4.00	4.00	4.00	4.00	4.50
<i>ALTERSGRUPPE</i>																		
<i>18-30 J.</i>																		
PR 25	3.45	3.63	3.00	3.91	3.67	3.75	4.00	2.75	2.67	2.67	3.60	3.67	3.67	3.50	3.00	3.00	3.00	3.00
PR 50	3.82	3.88	3.67	4.36	4.00	4.25	4.50	3.00	3.00	3.00	4.00	4.00	4.00	4.00	3.50	3.50	3.50	4.00
PR 75	4.18	4.25	4.00	4.64	4.33	4.75	4.75	3.33	3.33	3.33	4.30	4.33	4.67	4.25	3.90	3.83	3.40	4.00
<i>31-40 J.</i>																		
PR 25	3.40	3.38	3.00	4.00	3.67	4.00	4.00	2.75	2.78	2.67	3.60	3.67	3.67	3.25	3.10	3.00	3.00	3.50
PR 50	3.82	3.88	4.00	4.27	4.00	4.25	4.25	3.00	3.00	3.00	3.90	4.00	4.00	3.75	3.50	3.33	3.50	4.50
PR 75	4.18	4.25	4.00	4.45	4.33	4.75	4.75	3.73	3.78	3.67	4.30	4.33	4.33	4.25	3.89	3.83	4.00	4.50
<i>41-50 J.</i>																		
PR 25	3.55	3.63	3.00	4.00	3.67	4.00	4.00	2.83	2.78	2.67	3.60	3.67	3.67	3.50	3.30	3.17	3.00	4.00
PR 50	3.82	3.88	4.00	4.27	4.00	4.50	4.25	3.08	3.11	3.00	3.80	4.00	4.00	3.75	3.60	3.50	3.50	4.00
PR 75	4.09	4.13	4.00	4.64	4.67	4.75	4.75	3.67	3.67	3.67	4.11	4.33	4.33	4.00	3.90	3.83	4.00	4.50
<i>51-60 J.</i>																		
PR 25	3.45	3.50	3.00	3.82	3.33	4.00	4.00	3.00	3.00	3.00	3.40	3.33	3.33	3.25	3.20	3.00	3.00	3.50
PR 50	3.82	3.88	3.67	4.18	4.00	4.25	4.25	3.33	3.33	3.33	3.70	4.00	3.67	3.50	3.40	3.33	3.00	4.00
PR 75	4.18	4.13	4.00	4.55	4.33	4.75	4.50	3.92	4.00	3.67	3.90	4.00	4.00	3.75	4.00	3.83	4.00	4.50
<i>61-70 J.</i>																		
PR 25	3.36	3.38	3.00	4.00	3.33	4.00	3.75	2.92	3.00	2.67	3.30	3.33	3.33	3.25	3.40	3.33	3.00	3.50
PR 50	3.70	3.75	3.33	4.18	4.00	4.50	4.00	3.08	3.11	3.33	3.70	4.00	3.67	3.50	3.60	3.50	4.00	4.00
PR 75	4.00	4.00	4.00	4.36	4.33	4.75	4.25	3.83	4.00	3.67	4.00	4.00	4.00	3.75	4.00	4.00	4.00	4.00
<i>&gt;70 Jahre</i>																		
PR 25	3.36	3.50	3.00	3.91	3.67	4.00	4.00	3.08	3.11	3.33	3.30	3.33	3.00	3.25	3.60	3.50	3.00	4.00
PR 50	3.73	3.88	3.67	4.18	4.00	4.25	4.25	3.92	4.00	3.67	3.60	4.00	3.67	3.50	4.00	3.83	4.00	4.50
PR 75	4.18	4.13	4.00	4.55	4.33	4.75	4.50	4.42	4.44	4.00	3.90	4.00	4.00	4.00	4.40	4.33	4.00	5.00

## **Richtungseffekte bei Bildblattvorlagen im demoskopischen Face-to-face-Interview**

**von Erich Lamp<sup>1</sup>**

### ***Zusammenfassung***

*Anknüpfend an impressionistische und empirische Beobachtungen zur Links-Rechts-Asymmetrie in der Wahrnehmung optisch dargebotener Stimuli geht die vorliegende Untersuchung der Frage nach, ob sich eine Vorzugsrichtung in der Wahrnehmung auch bei Bildblattvorlagen im demoskopischen Interview auswirkt. Dazu wurden insgesamt 16 Fragebogenexperimente mit fünf verschiedenen Skalenmodellen durchgeführt, die in ihrer Richtung (horizontal/vertikal) und Achsenführung (links/rechts, rechts/links, unten/oben, oben/unten) variiert wurden. Die Ergebnisse zeigen, dass die Befragten beim Betrachten einer Bildblatt-Skala offenbar einer "natürlichen Blickbahn" folgen, die ihr Antwortverhalten unbewusst, jenseits aller rationalen Denk- und Entscheidungsvorgänge, wie sie von der kognitiven Methodenforschung in den Blick genommen werden, beeinflusst.*

### ***Abstract***

*Starting from impressionistic and empirical observations of the asymmetry between left and right in the perception of visual stimuli, this study tackles the question of whether a preferred orientation also affects the perception of illustrations in face-to-face interviews, and thus the use of such illustrations in survey research. Sixteen questionnaire experiments with five different scales were conducted, varying in orientation (horizontal vs. vertical) and direction (left to right vs. right to left; going down vs. going up) of the scales. Results show that respondents, when looking at visual scales, evidently follow a natural pattern of glaze, which unconsciously affects their answers beyond all rational processes of thinking and decision making that cognitive methodology usually looks at.*

---

1 Dr. **Erich Lamp** ist Akademischer Oberrat und Privatdozent am Institut für Publizistik der Universität Mainz, Colonel Kleinmann-Weg 2, 55099 Mainz., Tel. 06131/3925763. E-mail: erich.lamp@uni-mainz.de

## 1 Einleitung

In seinem Aufsatz "Über das Rechts und Links im Bilde" lenkt der Kunsthistoriker **Heinrich Wölfflin** die Aufmerksamkeit auf einen Sachverhalt, der über seine praktische Bedeutung in der darstellenden Kunst hinaus bis dahin kaum Beachtung gefunden hatte: die verschiedene Wertigkeit von rechter und linker Bildseite. **Wölfflin** selbst war auf dieses Phänomen eher durch Zufall gestoßen: "Man kann es im kunstgeschichtlichen Unterricht, wo mit Diapositiven gearbeitet wird, manchmal erleben", schreibt er dazu, "dass eine Platte, falsch eingestellt, das Bild im Gegen-sinn erscheinen lässt. Da gibt es dann gewöhnlich einen ärgerlichen Zwischenruf: 'Umdrehn! Die Platte ist verkehrt eingestellt!', aber man dürfte sich auch einmal besinnen, warum denn das Bild nicht umgekehrt werden kann und was sich in der Wirkung ändert" (**Wölfflin** 1940a). Zahlreiche Beobachtungen an seitenverkehrten Bildern führen ihn zu dem Schluss, dass die Stimmung eines Bildes offenbar maßgeblich von der rechten Bildseite determiniert wird. Kehrt man das Bild um, was am materiellen Bestand nichts ändert, verliere alles seinen Sinn und Zusammenhang und das Auge kämpfe mit Widerständen. Die Ursache für dieses "Wirkungsgesetz", das selbst auf scheinbar reine Symmetrien zutrefte, erkennt er in der "natürlichen Blickbahn" von links nach rechts: "Immer geben wir der rechten Seite den stärkeren Akzent" (**Wölfflin** 1940b).

Empirische Untersuchungen bestätigen die von **Wölfflin** beschriebene Vorzugsrichtung der Wahrnehmung von links nach rechts. In einer frühen Experimentalserie über Richtungsasymmetrien in der Wahrnehmung schrieben holländische und deutsche Versuchspersonen statischen Figuren in deutlich höherem Umfang eine Bewegungstendenz nach rechts als nach links zu. Die Richtungstendenz nach rechts war bei Versuchspersonen mit höherer Schulbildung (wenigstens Mittelschulabschluss) stärker ausgeprägt als bei Versuchspersonen mit niedrigerer Schulbildung. Des Weiteren wurde die Rechtsgerichtetheit in der Wahrnehmung erst ab einem Alter von 15 Jahren an beobachtet, während jüngere Versuchspersonen keine Bewegungsrichtung nach links oder nach rechts bevorzugten. Auch wurde bei kindlichen Versuchspersonen kein Zusammenhang zwischen Äugigkeit bzw. Links- oder Rechtshändigkeit und einer bevorzugten Richtungswahrnehmung gefunden. Insgesamt werden die Ergebnisse als Beleg dafür gesehen, dass die Rechtsgerichtetheit in der Wahrnehmung nicht die Folge angeborener körperlicher Dispositionen ist, sondern unter dem Einfluss kultureller Einwirkungen (dem Lesen und Schreiben von links nach rechts bzw. der größeren Lese- und Schreiberfahrung von links nach rechts) als *Gewohnheit* entsteht (**Van der Meer** 1958).

Je nach dargebotenem Stimulusmaterial oder Aufgabenstellung gelangten spätere Untersuchungen zu teilweise anderen Ergebnissen, die insbesondere einen Zusammenhang zwischen Rechts- bzw. Linkshändigkeit und bevorzugter Seitenorientierung zu belegen scheinen (*Banich* et al. 1989; *Beaumont* 1985; *Ellis, Miller* 1981; *Gross, Bornstein* 1978; *Levy* 1976; *McLaughlin* 1986; *McLaughlin* et al. 1983). Auch wurde eine Seitenorientierung schon bei jüngeren Versuchspersonen beobachtet (*Lieblich* et al. 1975) und sogar für Kindergartenkinder berichtet (*Dehaene* 1999). Zur Erklärung dieses Phänomens greifen manche Autoren auf Erkenntnisse der Gehirnforschung zurück und vermuten eine je nach Aufmerksamkeits- und Wahrnehmungsrichtung unterschiedliche Aktivierung der rechten bzw. linken Hemisphäre (*Beaumont* 1985; *Burt, Perett* 1997; *Levy* 1976; *Sattler* 1983). Dessen ungeachtet stimmen zahlreiche Untersuchungen darin überein, die Richtungsasymmetrie vor allem als kulturelle Variable zu betrachten und sie der Lese- und Schreibgewohnheit zuzuschreiben (*Cohen* 1973; *Dehaene* 1999; *Hatta, Kawakami* 1997; *Lieblich* et al. 1975; *Maass, Russo* 2003; *Nachson* 1981, 1985; *Nachson* et al. 1999; *Sakhuja* et al. 1996; *Vaid, Singh* 1989). So berichtet *Nachson* eine Rechtsgerichtetheit in der Wahrnehmung sowohl bei englischen (*Nachson* 1985) als auch bei russischen Versuchspersonen (*Nachson* et al. 1999), während bei arabischen und israelischen Versuchspersonen, die in entgegengesetzter Richtung lesen und schreiben, in beiden Experimenten eine Linksorientierung beobachtet wurde. Diese war je nach Vertrautheit mit der abendländischen Kultur stärker oder schwächer ausgeprägt. Experimente mit französischen im Vergleich zu iranischen Versuchspersonen, die das Lesen entgegen unserer Konvention zunächst von rechts nach links gelernt hatten, führten zu gleichen Ergebnissen (*Dehaene* 1999).

Lassen sich diese und ähnliche Beobachtungen, die ganz unterschiedlichen Forschungsinteressen und Wissenschaftsbereichen entstammen, auch für die Umfrageforschung fruchtbar machen? Vor allem das Face-to-face-Interview nutzt neben dem Fragestimulus vielfach die Möglichkeit der optischen Unterstützung mit Hilfe von Bildblattvorlagen, z.B. um Situationen lebhafter zu illustrieren, verschiedene Argumente gleichwertig nebeneinander darzustellen oder Meinungen gedanklich zu strukturieren und in eine zählbare Form zu bringen (*Noelle-Neumann, Petersen* 1996). Wirkt sich eine Vorzugsrichtung in der Wahrnehmung optisch dargebotener Stimuli auch bei Bildblattvorlagen im demoskopischen Interview aus, bleibt dies nicht ohne Konsequenzen für das Antwortverhalten der Befragten, die – neben der (bewussten) Reaktion auf den Fragestimulus – unbewusst auch auf die formale Gestaltung des Bildblatts reagieren. Dieser Problematik soll im Folgenden anhand verschiedener Skalen-Bildblätter nachgegangen werden.

## 2 Methodisches Vorgehen

Die berichteten Ergebnisse stammen sämtlich aus gegabelten Face-to-face-Befragungen der Mainzer Bevölkerung, die nach ihrer Logik dem kontrollierten Experiment folgen. Alle Umstände im Ablauf der Untersuchung sind gleich, bis auf einen, den experimentellen Faktor, der im Fragebogen variiert wird. Ziel einer solchen Anlage ist es, für Unterschiede, die nach Einwirkung des experimentellen Faktors gemessen werden, alle Erklärungen bis auf eine auszuschließen. Bei den mehrmals im Jahr durchgeführten Umfragen stehen insbesondere methodische Aspekte des Fragebogens im Vordergrund des Interesses, um das Entdeckungspotential der Umfrageforschung durch Anwendung der Technik der gegabelten Befragung systematisch zu nutzen.

Die Fragebögen enthalten in der Regel Fragen zu ganz unterschiedlichen Themen und Sachverhalten und decken in ihrer Konstruktion die ganze Bandbreite der verschiedenartigen Fragemodelle und Hilfsmittel einschließlich Listen, Kartenspielen und Bildblattvorlagen ab. Die besondere Aufmerksamkeit gilt hierbei vor allem drei Gruppen von Einflussgrößen, die sich erstens auf die Anordnung und Abfolge der Fragen im Fragebogen beziehen, zweitens auf den Einsatz verschiedener Fragemodelle und -instrumente sowie drittens auf die Art der Frageformulierung. Die Interviews wurden von Studierenden der Publizistikwissenschaft der Universität Mainz durchgeführt. Die Stichproben wurden nach dem Quotenverfahren gebildet, wobei die Merkmale Geschlecht und Alter (in der Verteilung 16 - 29, 30 - 44, 45 - 59, 60+ Jahre) vorgegeben waren.

Die Antworten der befragten Mainzer Bürger sind im statistischen Sinne nicht repräsentativ für die Gesamtbevölkerung der Bundesrepublik. Dies ist für die vorliegende Untersuchung unerheblich, da es nicht um die Meinungen aller Deutschen geht, sondern darum, die Antworten von Befragten, denen verschiedene Versionen von Skalen-Bildblättern vorgelegt wurden, systematisch miteinander zu vergleichen. Es gibt keinen Grund anzunehmen, dass Unterschiede, die sich in den Antworten der Mainzer Bürger zeigen, nicht auch für die Gesamtbevölkerung gelten (vgl. *Lamp* 2001).

### 3 Ergebnisse

#### 3.1 Die horizontale "Männchen-Skala"

Skalierte Abfragemodelle liefern einen empfindlichen Maßstab zur Messung von Meinungen, Einstellungen, Zusammenhängen usw., für die differenzierte Aussagen erwartet werden. Erstreckt sich die Skala über mehr als fünf Skalenpunkte, benötigt der Befragte in der Regel eine optische oder andere Hilfestellung, die über den nur sprachlich vermittelten Stimulus hinausgeht (*Sudman, Bradburn* 1982). Hierbei steht vor allem die Anschaulichkeit im Vordergrund, damit die Befragten mit der Skala auch umgehen können, selbst wenn die zu messende Dimension nur schwer zu fassen ist (*Ring* 1992). Ein gutes Beispiel hierfür liefert die von *Erp Ring* entwickelte so genannte "Männchen-Skala", die seit vielen Jahren in deutschen und internationalen Umfragen eingesetzt wird, um den Faktor "Tempo" zu messen, etwa wie schnell bei bestimmten Tätigkeiten oder in bestimmten Situationen die Zeit vergeht (vgl. *Noelle-Neumann, Köcher* 1993; 1997; 2002). Das Skalen-Bildblatt zeigt sieben stilisierte Männchen in unterschiedlichen Bewegungspositionen, angefangen von einem stehenden Männchen am linken Bildrand bis hin zu einem mit ausladendem Schritt laufenden Männchen am rechten Bildrand (Abb. 1a). Durch diese Übersetzung der ansonsten kaum fassbaren Frage, wie schnell die Zeit vergeht, die von den Befragten in aller Welt verstanden wird, wird das Tempo gegliedert und sowohl dem Befragten für seine Antwort als auch dem Forscher für seine Analyse zugänglich gemacht. Anders als bei einer verbalen Abstufung, die aufgrund ihrer Formulierung das Antwortverhalten der Befragten beeinflussen kann (*Worcester, Burns* 1975), werden die Urteile und Antworten nicht durch verschiedene Begriffsverwendungen beeinflusst. Ist es möglich, dass statt von der sprachlichen Darstellungsweise hier von der optischen Darstellung ein Einfluss ausgeübt wird? Macht es – im Sinne der angesprochenen Links-Rechts-Asymmetrie in der Wahrnehmung – einen Unterschied, ob die Laufrichtung der Männchen auf dem Bildblatt von links nach rechts oder von rechts nach links dargestellt ist (Abb. 1b)?

Zur Überprüfung dieser Frage wurde folgende Testfrage gestellt: *"Was würden Sie sagen, wie schnell der technische Fortschritt voranschreitet, wie schnell sich alles entwickelt? So kann man das natürlich nur schwer sagen, aber wenn Sie sich bitte einmal dieses Bild hier ansehen ..."* An dieser Stelle wurde den Befragten das – im Split-ballot-Verfahren variierte – Skalen-Bildblatt übergeben, anhand dessen sie ihre Einschätzung vornehmen sollten: Eine Befragtengruppe erhielt das herkömmliche Bildblatt mit den nach rechts laufenden Männchen, die andere Befragtengruppe

erhielt ein modifiziertes Bildblatt mit nach links laufenden Männchen.<sup>2</sup> In der für alle Befragten gleich formulierten Anweisung für die Auswahl hieß es dazu weiter: "... das stehende Männchen auf der einen Seite würde bedeuten, die Technik macht überhaupt keine Fortschritte, und das schnell laufende Männchen auf der anderen Seite würde bedeuten, die Technik macht rasende Fortschritte. Welches von den sieben Männchen würden Sie für den technischen Fortschritt heute typisch finden, welches Männchen würden Sie sagen?"

### Abbildung 1 Männchen-Skala



Abb. 1a Laufrichtung von links nach rechts



Abb. 1b Laufrichtung von rechts nach links

Um einen möglichen Kontexteffekt auszuschließen, wurde die Testfrage unverändert in vier Befragungen mit unterschiedlichem Frageumfeld aufgenommen. Die Ergebnisse zeigen für alle vier Befragungen ein einheitliches Antwortmuster: In beiden Versionen der Bildblattvorlage häufen sich die Antworten mit über 90 Prozent der Nennungen auf den drei oberen Skalenstufen, die man als "schnell", "sehr schnell" und "rasend schnell" umschreiben kann, und auch das Verteilungsbild der Antworten ist *innerhalb* der beiden Versuchsgruppen in allen vier Befragungen identisch. *Zwischen* den beiden Gruppen zeigt sich jedoch ein deutlicher Unterschied: wie aus den zusammengefassten Daten hervorgeht, schätzten Befragte, de-

2 Um sicherzustellen, dass die auf **Rings** Originalbildblatt den Männchen zugeordneten Zahlen für die Stufen von 1 bis 7 sich nicht verzerrend auf das Ergebnis auswirken, erhielten die Befragten ein Bildblatt ohne Ordinalzahlen. In der spiegelbildlichen Version hätte sonst entweder die Bezifferung geändert werden müssen oder der Zahlenstrahl wäre von 7 nach 1 verlaufen. In beiden Versionen wurden die Interviewer angewiesen, im Zweifelsfall nachzufragen, das wievielte Männchen von links oder rechts gewählt wurde.

nen die Männchen-Skala mit Laufrichtung von links nach rechts vorgelegt wurde, das Tempo des technischen Fortschritts mehrheitlich als "sehr schnell" (Skalenstufe 6) ein, wohingegen Befragte, denen das Bildblatt mit umgekehrter Laufrichtung von rechts nach links vorgelegt wurde, das Tempo des technischen Fortschritts mehrheitlich als "rasend schnell" (Skalenstufe 7) einschätzten (Tabelle 1). Obwohl die spiegelbildliche Anordnung nichts am materiellen Bildbestand ändert, verschiebt sich das Schwergewicht der Angaben in der Bildblattumkehrung um eine Skalenstufe. Die maximale Differenz in der Tempo-Einschätzung des technischen Fortschritts zwischen den beiden Versuchsgruppen beträgt 18 Prozentpunkte.

**Tabelle 1** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der "Männchen-Skala"

Frage: "Was würden Sie sagen, wie schnell der technische Fortschritt voranschreitet, wie schnell sich alles entwickelt? So kann man das natürlich nur schwer sagen, aber wenn Sie sich bitte einmal dieses Bild hier ansehen: das stehende Männchen auf der einen Seite würde bedeuten, die Technik macht überhaupt keine Fortschritte, und das schnell laufende Männchen auf der anderen Seite würde bedeuten, die Technik macht rasende Fortschritte. Welches von den sieben Männchen würden Sie für den technischen Fortschritt heute typisch finden, welches Männchen würden Sie sagen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit von links nach rechts laufenden Männchen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit von rechts nach links laufenden Männchen.

		Laufrichtung der Männchen auf dem Bildblatt		
		von links nach rechts %	von rechts nach links %	Differenz
Der technische Fortschritt kommt...				
... überhaupt nicht voran	(Stufe 1)	--	--	--
	Stufe 2	--	1	+1
	Stufe 3	2	1	-1
	Stufe 4	6	3	-3
... rasend schnell voran	Stufe 5	21	18	-3
	Stufe 6	46	34	-12
	(Stufe 7)	25	43	+18
		100	100	
n =		367	340	

IfP-MP, Umfragen SS 1999, WS 99/00, SS 2000, WS 00/01 (zusammengefasste Daten)

Informelle Ermittlungen des Verfassers lassen vermuten, dass die Befragten das Tempo der laufenden Männchen selbst als schneller wahrnehmen, wenn diese kon-

form zur Lese- und Schreibgewohnheit von links nach rechts laufen anstatt "gegen den Strich" von rechts nach links. Das Zurückbleiben um einen Skalenpunkt in der Einschätzung der Mehrheit, wie schnell der technische Fortschritt voranschreitet, würde sich so – der Logik des kontrollierten Experiments entsprechend – aus der "natürlichen Blickbahn" der Befragten erklären: nicht das Tempo des technischen Fortschritts wird unterschiedlich eingeschätzt, vielmehr wird die Geschwindigkeit der in verschiedene Richtungen laufenden Männchen unterschiedlich wahrgenommen.

Handelt es sich bei diesem – in mehreren Befragungen empirisch erhärteten – Ergebnis um einen Einzelbefund, der nur auf den Sonderfall der "Männchen-Skala" zutrifft, oder gibt es systematische Unterschiede in der Wahrnehmung von Bildblatt-Skalen, je nachdem, ob sie den Befragten als "Original" oder in der spiegelbildlichen Umkehrung vorgelegt werden?

### 3.2 Die horizontale "Kreise-Skala"

Ersten Aufschluss hierüber liefert ein Skalen-Bildblatt mit einer gegenüber der "Männchen-Skala" deutlich zurückgenommenen Dynamik. Zur abgestuften Messung von Unterschieden und Gemeinsamkeiten (vgl. *Noelle-Neumann, Köcher* 1993; 1997; 2002) zeigt das Bildblatt mehrere Kreispaae, die zunächst ohne Überschneidung über kontinuierlich größer werdende Schnittmengen bis zur vollständigen Deckungsgleichheit von links nach rechts zusammenlaufen (Abb. 2a).<sup>3</sup> Anders als die "Männchen-Skala" mit ihren immer ausladender sich bewegenden Figuren bezieht die "Kreise-Skala" ihre Dynamik lediglich aus dem allmählichen Zusammenlaufen der Kreise. Macht es beim Betrachten des Bildblatts einen Unterschied, ob die Kreise von links nach rechts zusammenlaufen oder von rechts nach links (Abb. 2b)?

---

3 Die vom Institut für Demoskopie Allensbach verwendete Skala ist in der Originalbildblattvorlage senkrecht organisiert. Zu Testzwecken wurde sie hier in die Horizontale gedreht.

## Abbildung 2 Kreise-Skala

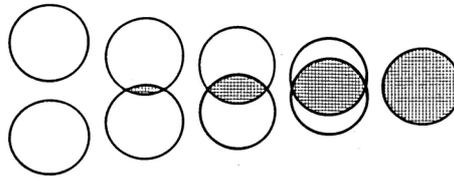


Abb. 2a Von links nach rechts zusammenlaufende Kreise

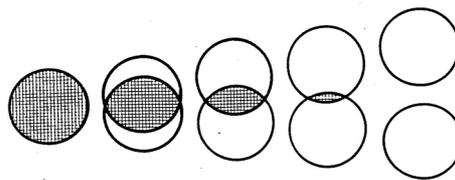


Abb. 2b Von rechts nach links zusammenlaufende Kreise

In insgesamt drei Befragungen wurde dazu die folgende Testfrage gestellt: *"Was meinen Sie, wie sehr Männer und Frauen sich in ihrem Fühlen, ihrem Gefühl unterscheiden? Wenn Sie es mir bitte nach diesem Bildblatt mit den Kreisen hier sagen: Die beiden weißen Kreise würden bedeuten, Männer und Frauen fühlen ganz verschieden, der dunkle Kreis würde bedeuten, sie fühlen ganz gleich. Was meinen Sie, wie sehr sich Männer und Frauen in ihrem Fühlen gleichen oder unterscheiden, welche Stufe würden Sie wählen?"* Das vorgelegte Bildblatt zeigte die "Kreise-Skala" einmal mit den von links nach rechts zusammenlaufenden Kreisen, eine andere Befragengruppe erhielt das Bildblatt mit den in der umgekehrten Richtung von rechts nach links zusammenlaufenden Kreisen.<sup>4</sup>

In beiden Untersuchungsgruppen bevorzugten die Befragten mehrheitlich die neutrale Mittelposition: Männer und Frauen wurden in ihrem Fühlen weder als ganz verschieden noch als ganz gleich gesehen. Befragte, die sich für eine der beiden Alternativen entschieden, vertraten in beiden Gruppen überwiegend die Ansicht, Männer und Frauen fühlten eher unterschiedlich, ein geringerer Prozentsatz äußerte die Meinung, sie fühlten eher gleich. Wie die zusammengefassten Daten zeigen, gibt es *zwischen* den Gruppen dennoch einen deutlichen Unterschied: Befragte, denen das Bildblatt mit den von links nach rechts zusammenlaufenden Kreisen vorgelegt wurde, entschieden sich in höherem Maße für die Mittelposition und in geringerem

4 Die Befragten erhielten ein Bildblatt ohne den Kreisen zugeordnete Ordinalzahlen (Vorgehen wie in Anm. 1 beschrieben).

**Tabelle 2** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der "Kreise-Skala"

Frage: "Was meinen Sie, wie sehr Männer und Frauen sich in ihrem Fühlen, ihrem Gefühl unterscheiden? Wenn Sie es mir bitte nach diesem Bildblatt mit den Kreisen hier sagen: Die beiden weißen Kreise würden bedeuten, Männer und Frauen fühlen ganz verschieden, der dunkle Kreis würde bedeuten, sie fühlen ganz gleich. Was meinen Sie, wie sehr sich Männer und Frauen in ihrem Fühlen gleichen oder unterscheiden, welche Stufe würden Sie wählen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit von links nach rechts zusammenlaufenden Kreisen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit von rechts nach links zusammenlaufenden Kreisen.

	Zusammenlaufen der Kreise auf dem Bildblatt		
	von links nach rechts %	von rechts nach links %	Differenz
Männer und Frauen sind in ihrem Fühlen			
... verschieden (Stufen 1 und 2)	30	34	+4
... teils verschieden, teils ähnlich (Stufe 3)	51	43	-8
... gleich (Stufen 4 und 5)	19	23	+4
	100	100	
n =	407	392	

IfP-MP, Umfragen SS 2000, WS 00/01, SS 2001 (zusammengefasste Daten)

Umfang für die beiden Alternativen als Befragte mit der Bildblatturnkehrung. Diese neigten deutlich seltener zur neutralen Mitte und bezogen häufiger für die eine oder die andere Position Stellung (Tabelle 2). Insgesamt fallen die Ergebnisse in der Belegung der einzelnen Skalenstufen um bis zu 8 Prozentpunkte auseinander.

Die gegenüber der "Männchen-Skala" in ihrer Dynamik stark zurückgenommene "Kreise-Skala" führt je nach Bildblattvorlage zwar ebenfalls zu unterschiedlichen Ergebnissen, die Unterschiede fallen jedoch deutlich geringer aus als bei der "Männchen-Skala". Könnte es sein, dass mit abnehmender Dynamik einer Skala auch die in der Bildblatturnkehrung beobachtete Wahrnehmungsverschiebung geringer wird?

### 3.3 Die horizontale "Baum-Skala"

Die Vermutung wurde anhand der "Baum-Skala" überprüft. Das Skalen-Bildblatt zeigt fünf in gleichen Abständen in der Horizontalen angeordnete Äpfel, die sich lediglich durch ihre Entfernung zum Stamm eines seitwärts stehenden Apfelbaums unterscheiden. In Anlehnung an das Sprichwort "Der Apfel fällt nicht weit vom Stamm" misst die "Baum-Skala" die Nähe zu den Eltern (vgl. *Noelle-Neumann, Köcher* 1993; 1997; 2002). Der dem Stamm am nächsten liegende Apfel drückt danach eine große Nähe zu den Eltern aus, der am weitesten weg liegende Apfel entsprechend eine große Entfernung (Abb. 3a). Im Unterschied zu den immer ausladender sich bewegenden Figuren der "Männchen-Skala" und den wie eine Pfeilspitze zusammenlaufenden Kreisen der "Kreise-Skala" geht von den regelmäßig in gerader Folge angeordneten Äpfeln der "Baum-Skala", einer Perlschnur vergleichbar, keinerlei Dynamik aus. Wie nehmen die Befragten die "Baum-Skala" in der Bildblattumkehrung wahr, in der sich an der Skala selbst, also der Anordnung der Äpfel auf der Horizontalen, nichts ändert, und einzig der Apfelbaum auf die andere Seite zu stehen kommt (Abb. 3b)?

#### Abbildung 3 Baum-Skala

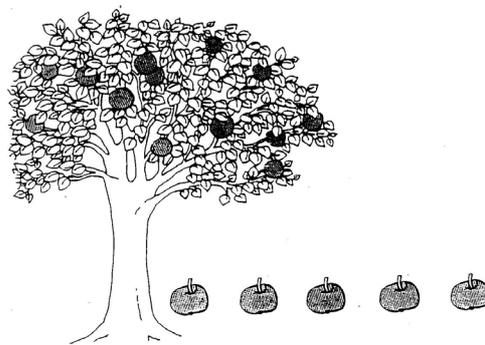


Abb. 3a Apfelbaum am linken Bildrand

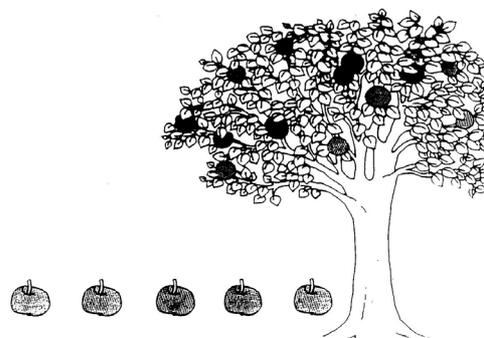


Abb. 3b Apfelbaum am rechten Bildrand

Die in drei Befragungen gestellte Testfrage lautete: *"Es gibt ja das alte Sprichwort 'Der Apfel fällt nicht weit vom Stamm', und man meint damit, daß einem die Kinder ähnlich sind. Wenn man sich das jetzt einmal als Bild vorstellt, könnte das so aussehen. Wie ist das bei Ihnen, wie weit sind Sie von Ihren Eltern entfernt? Welchen Apfel würden Sie da wählen?"* Im Split-ballot-Verfahren erhielt eine Befragten-Gruppe jeweils das Bildblatt mit dem Apfelbaum auf der linken Seite, der anderen Befragten-Gruppe wurde das Bildblatt in der spiegelbildlichen Version mit dem Apfelbaum auf der rechten Seite vorgelegt.<sup>5</sup> Die zusammengefassten Daten sprechen eine eindeutige Sprache: Die Ergebnisse in beiden Gruppen sind nahezu identisch (Tabelle 3).

**Tabelle 3** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der "Baum-Skala"

Frage: *"Es gibt ja das alte Sprichwort 'Der Apfel fällt nicht weit vom Stamm', und man meint damit, daß einem die Kinder ähnlich sind. Wenn man sich das jetzt einmal als Bild vorstellt, könnte das so aussehen. Wie ist das bei Ihnen, wie weit sind Sie von Ihren Eltern entfernt? Welchen Apfel würden Sie da wählen?"*

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragten-Gruppe ein Bildblatt mit am linken Bildrand abgebildetem Apfelbaum vorgelegt, die andere Befragten-Gruppe erhielt ein Bildblatt, auf dem der Apfelbaum auf der rechten Bildseite abgebildet war.

	Abbildung des Apfelbaums auf dem Bildblatt		
	linke Bildseite %	rechte Bildseite %	Differenz
Einstufung der Entfernung zu den Eltern als...			
... nah (Stufen 1 und 2)	44	43	-1
... weder nah noch fern (Stufe 3)	33	34	+1
... fern (Stufen 4 und 5)	23	23	--
	100	100	
n =	552	545	

IfP-MP, Umfragen SS 2000, WS 00/01, SS 2001 (zusammengefasste Daten)

<sup>5</sup> Die Befragten erhielten ein Bildblatt ohne den Äpfeln zugeordnete Ordinalzahlen (Vorgehen wie in Anm. 1 beschrieben).

Die bei der "Männchen-Skala" und der "Kreise-Skala" in der Umkehrung der Bildblattvorlage aufgetretenen Ergebnisverzerrungen könnten demnach, wie vermutet, eine Funktion der Skalendynamik sein. Zumindest konnte eine Wahrnehmungsverzerrung bei der Bildblattumkehrung im Fall der statischen "Baum-Skala" nicht nachgewiesen werden. Die empirische Evidenz scheint somit für die Annahme zu sprechen, dass von statischen Skalen keine Richtungstendenz in der Wahrnehmung ausgeht, die das Antwortverhalten der Befragten unabhängig vom Frageinhalt im Sinne einer Links-Rechts-Asymmetrie beeinflusst.

Dennoch hinterlässt das Ergebnis, möglicherweise aufgrund seiner Eindeutigkeit, ein eigenartiges Gefühl. Gerade an einem der "Baum-Skala" so ähnlichen Bild: der "Landschaft mit den drei Bäumen" von **Rembrandt**, hatte der eingangs zitierte Kunsthistoriker **Wölfflin** seine Beobachtungen der unterschiedlichen Wirkung seitenverkehrter Bilder illustriert. Danach überwog in der Radierung ein stimmungs-mäßiger Akzent von Energie, wenn sich die Vertikale der Baumgruppe auf der rechten Bildseite befand, wohingegen im umgekehrten Fall ein Akzent von Ruhe vorherrschte (**Wölfflin** 1940a). Könnte es sein, dass die Bildblattumkehrung der "Baum-Skala" im Split-ballot-Experiment nur deswegen zu gleichen Ergebnissen führte, weil eine mögliche Richtungsorientierung der Befragten durch den rechts bzw. links stehenden Apfelbaum "abgelenkt" wurde?

### 3.4 Die horizontale "Stapel-Skala"

Die Frage wurde anhand einer anderen statischen Skala überprüft. Dazu wurde die Versuchsanordnung der Bildblattumkehrung horizontal organisierter Skalen abgewandelt und als Prüf-Skala eine vertikal ausgerichtete Skala verwendet, die in der experimentellen Variation statt in der spiegelbildlichen Darbietung lediglich um 90 Grad gedreht wurde. Sollte die Annahme zutreffen, dass von statischen Skalen im Unterschied zu dynamischen Skalen keine Richtungstendenz in der Wahrnehmung ausgeht, dürften sich die Antworten der Befragten in den Untersuchungsgruppen auch in diesem Fall nicht voneinander unterscheiden. Die Überprüfung erfolgte anhand der – nach ihrem Erfinder **Jan Stapel** benannten – "Stapel-Skala", die zur Messung bipolarer Reaktionsmöglichkeiten verwendet wird (vgl. **Noelle-Neumann, Köcher** 1993; 1997; 2002). Das statisch-seitensymmetrische Skalen-Bildblatt zeigt mehrere in gleich großen Abständen senkrecht übereinander angeordnete schwarze und weiße Kästchen, die von einem Nullpunkt in der Mitte in die positive (+5) und negative (-5) Richtung reichen (Abb. 4a). Die dazu gestellte Testfrage lautete: "*Eine Frage zu Bundespräsident **Rau**: Wie sympathisch ist Ihnen **Johannes Rau**? Vielleicht sagen Sie es mir nach diesem Bildblatt mit den weißen und schwarzen Käst-*

chen hier: -5 würde bedeuten, er ist Ihnen ganz und gar unsympathisch, und +5, er ist Ihnen äußerst sympathisch. Welches Kästchen würden Sie sagen?" Das vorgelegte Bildblatt zeigte die "Stapel-Skala" einmal in der Vertikalen, eine andere Befragten-Gruppe erhielt das Bildblatt mit der "Stapel-Skala" in der horizontalen Anordnung (Abb. 4b).<sup>6</sup>

#### Abbildung 4 Stapel-Skala



Abb. 4a Vertikale Skalenausrichtung



Abb. 4b Horizontale Skalenausrichtung

Unabhängig von der Art der Bildblattvorlage überwogen in beiden Untersuchungsgruppen die positiven Werte: Hier wie da bekundete die Hälfte der Befragten ihre Sympathie für *Johannes Rau* (Skalenstufen +2 bis +5). Das Votum der übrigen Befragten fiel dagegen je nach Skalenversion sehr unterschiedlich aus. Befragte, denen das Bildblatt in der senkrechten Skalenanordnung vorgelegt wurde, entschieden sich mit 23 Prozent der Stimmen fast doppelt so häufig für die negativen Werte "unsympathisch" (Skalenstufen -2 bis -5) wie Befragte, die das Skalen-Bildblatt in der horizontalen Anordnung erhielten (12 Prozent). Entsprechend deutlich fielen die Unterschiede in der Belegung der neutralen Mittelposition aus (Tabelle 4). Das bei Verwendung der "Baum-Skala" gefundene Ergebnis, wonach bei der Bildblattem-

6 Da die Skala in der vorliegenden Versuchsanordnung nicht in der Horizontalen gespiegelt, sondern aus der Vertikalen um 90 Grad in die Horizontale gedreht wurde, konnte die Bezifferung der Kästchen beibehalten werden.

**Tabelle 4** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der "Stapel-Skala"

Frage: "Eine Frage zu Bundespräsident Rau: Wie sympathisch ist Ihnen Johannes Rau? Vielleicht sagen Sie es mir nach diesem Bildblatt mit den weißen und schwarzen Kästchen hier: -5 würde bedeuten, er ist Ihnen ganz und gar unsympathisch, und +5, er ist Ihnen äußerst sympathisch. Welches Kästchen würden Sie sagen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit der Skala in der Vertikalen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit der Skala in der Horizontalen.

	Anordnung der Skala auf dem Bildblatt		
	in der Vertikalen %	in der Horizontalen %	Differenz
<b>Sympathieempfinden für <i>Johannes Rau</i></b>			
unsympathisch (Stufen -5 bis -2)	23	12	-11
neutral (Stufen -1 bis +1)	27	39	+12
sympathisch (Stufen +2 bis +5)	50	49	-1
%	100	100	
n =	77	90	

IfP-MP, Umfrage WS 2000/01

kehrung keine Ergebnisverzerrung beobachtet wurde, mag somit in der Tat mit einer "Ablenkung" der Richtungsorientierung der Befragten durch den rechts bzw. links stehenden Apfelbaum zusammenhängen.

### 3.5 Die horizontale "Leiter-Skala"

Weiteren Aufschluss in dieser Frage liefert ein anderes Skalenmodell, das eine Mischform aus "Stapel-Skala" und – vom Nullpunkt aus nur in eine Richtung weisende – "Leiter-Skala" darstellt. Das Skalen-Bildblatt zeigt ebenfalls elf in einer Linie in gleich großen Abständen angeordnete Kästchen, allerdings in einer kontinuierlich von schwarz nach weiß heller werdenden Stufenfolge, deren Zählung nach dem Modell der "Leiter-Skala" von 0 bis 10 reicht (Abb. 5a). Als Testfrage wurde die Sympathiefrage zu **Johannes Rau** beibehalten und lediglich die Stufenbenennung zu Null (statt -5) für ganz und gar unsympathisch, und 10 (statt +5) für äußerst sympathisch abgeändert. Die experimentelle Variation der Bildblattvorlagen in der

### Abbildung 5 Modifizierte Leiter-Skala



Abb. 5a Vertikale Skalenausrichtung



Abb. 5b Horizontale Skalenausrichtung

vertikalen und horizontalen Skalenanordnung (Abb. 5b)<sup>7</sup> erfolgte in derselben Befragung wie die Ermittlung anhand der "Stapel-Skala".<sup>8</sup>

Befragte, denen das vertikale Skalen-Bildblatt vorgelegt wurde, entschieden sich fast doppelt so häufig für die negativen Werte "unsympathisch" (Skalenstufen 0 bis 3)

7 Auch in diesem Fall wurde die Bezifferung der Kästchen beibehalten (vgl. Anm. 5).

8 Aus den Befunden eines anderen Skalen-Experiments (IfP-MP SS 1999) geht hervor, dass sich die kontinuierlich von schwarz nach weiß heller werdende Stufenfolge der elfstufigen "Leiter-Skala" nicht auf das Antwortverhalten der Befragten auswirkt. Im Split-ballot-Verfahren wurde einer Befragtengruppe die "Leiter-Skala" in dieser Version vorgelegt, eine andere Befragtengruppe erhielt das Skalen-Bildblatt in der herkömmlichen Version als Abbildung einer schwarzen Leiter auf weißem Grund. Abgefragt wurde das Sympathieempfinden für **Johannes Rau** nach dem Wortlaut der vorliegenden Testfrage, wobei zum damaligen Messzeitpunkt nur noch der Einschub "dem Nachfolger von **Roman Herzog** als Bundespräsident" hinzugefügt wurde. Der Ergebnisvergleich beider Skalen zeigt, dass die unterschiedliche graphische Gestaltung (weiß vs. grundiert) sich nicht in den Antworten der Befragten niederschlägt: Unsympathisch (Stufen 0 bis 3): 15 Prozent (weiß) vs. 12 Prozent (grundiert); neutral (Stufen 4 bis 6): 48 Prozent vs. 50 Prozent; sympathisch (Stufen 7 bis 10): 37 Prozent vs. 38 Prozent.

**Tabelle 5** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der modifizierten "Leiter-Skala"

Frage: "Eine Frage zu Bundespräsident **Rau**: Wie sympathisch ist Ihnen Johannes Rau? Vielleicht sagen Sie es mir nach diesem Bildblatt mit den weißen und schwarzen Kästchen hier: 0 würde bedeuten, er ist Ihnen ganz und gar unsympathisch, und 10, er ist Ihnen äußerst sympathisch. Welches Kästchen würden Sie sagen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit der Skala in der Vertikalen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit der Skala in der Horizontalen.

	Anordnung der Skala auf dem Bildblatt		
	in der Vertikalen %	in der Horizontalen %	Differenz
<b>Sympathieempfinden für <i>Johannes Rau</i></b>			
unsympathisch (Stufen 0 bis 3)	24	13	-11
neutral (Stufen 4 bis 6)	45	42	-3
sympathisch (Stufen 7 bis 10)	31	45	+14
%	100	100	
n =	79	79	

IfP-MP, Umfrage WS 2000/01

wie Befragte, die das horizontale Skalen-Bildblatt erhielten (24 Prozent vs. 13 Prozent). Umgekehrt blieben sie in ihrem Votum "sympathisch" (Skalenstufen 7 bis 10) deutlich hinter diesen zurück (31 Prozent vs. 45 Prozent). Lediglich in der Belegung der neutralen Mittelposition (Skalenstufen 4 bis 6) hielten die Befragten beider Versuchsgruppen sich mit gut zwei Fünfteln der Stimmen die Waage (Tabelle 5).

Der Befund untermauert das bei Verwendung der "Stapel-Skala" gefundene Ergebnis, wonach auch von statischen Skalen eine Richtungstendenz in der Wahrnehmung ausgeht, die das Antwortverhalten der Befragten unabhängig vom Frageinhalt im Sinne einer Links-Rechts-Asymmetrie beeinflusst. Damit erhält zugleich die Annahme Auftrieb, dass im Fall der statischen "Baum-Skala" trotz der in gerader Linie und regelmäßigen Abständen angeordneten Äpfel von der Baumstellung am linken bzw. rechten Bildrand ein zusätzlicher Impuls ausgeht, der die "natürliche Blickbahn" der Befragten ablenkt.

### 3.6 Bipolare "Stapel-Skala" und unipolare "Leiter-Skala" im Vergleich

Über die Einzelbefunde hinaus sind die bei der "Stapel-Skala" und ihrer abgeänderten Version (nach der Stufenfolge der "Leiter-Skala") gefundenen Ergebnisse insbesondere in der Zusammenschau aufschlussreich. Die experimentelle Variation der Skalenmodelle wurde nach dem Split-ballot-Verfahren in vier Untersuchungsgruppen derselben Befragung durchgeführt. Die Konstanz aller Umstände im Ablauf der Untersuchung mit Ausnahme des experimentellen Faktors, der im Fragebogen variiert wurde, ist somit nicht nur für die vertikale bzw. horizontale *Skalenanordnung* der Bildblattvorlagen gesichert, sondern auch für die Variation der *Skalenmodelle*. Für die zwischen den beiden Skalenmodellen gefundenen Unterschiede können mithin ebenfalls alle Erklärungen bis auf eine – die verschiedene Polarität der Skalen – ausgeschlossen werden.

Bei Vorlage der *vertikal* organisierten Bildblatt-Skalen ist der Anteil der negativen Werte "unsympathisch" (Skalenstufen -5 bis -2 bzw. 0 bis 3) mit einem knappen Viertel der Stimmen in beiden Untersuchungsgruppen gleich. Deutlich verschieden fällt das Votum der Befragten hingegen für die neutrale Mittelposition (Skalenstufen -1 bis +1 bzw. 4 bis 6) und bei den positiven Werten "sympathisch" (Skalenstufen +2 bis +5 bzw. 7 bis 10) aus. Befragte, die ihre Sympathieeinstufung anhand der bipolaren "Stapel-Skala" vornahmen, entschieden sich in erheblich höherem Umfang für die positiven Werte als Befragte, die ihre Sympathieeinstufung anhand der unipolaren Skala vornahmen (50 Prozent vs. 31 Prozent), wohingegen diese sich deutlich häufiger für die neutrale Mittelposition entschieden als jene (45 Prozent vs. 27 Prozent).

Der Befund deckt sich mit dem in der Methodenliteratur beschriebenen Sachverhalt, "that bipolar scales need not yield the same results as ratings on the constituent unipolar scales" (*Ostrom* 1987), wie ebenfalls von *Noelle-Neumann* und *Petersen* (1996) am Beispiel von mit "Stapel-Skala" und "Leiter-Skala" erhobenen Daten gezeigt wird. Da es sich hierbei um die gleiche Skalenanordnung handelt wie in der vorliegenden Untersuchung, bestätigt das hier berichtete Ergebnis die dort vorgelegten Befunde. Vor diesem Forschungshintergrund ist es umso bemerkenswerter, dass die empirisch gesicherten Unterschiede zwischen uni- und bipolaren Skalen (vgl. *Hippler* u.a. 1991; *Schwarz* u.a. 1991) nicht auftreten, wenn die Skalen-Bildblätter statt in der Vertikalen in der Horizontalen dargeboten werden. In diesem Fall sind die Ergebnisse trotz der unterschiedlichen Skalenabfrage in beiden Untersuchungsgruppen nahezu identisch (Tabelle 6).

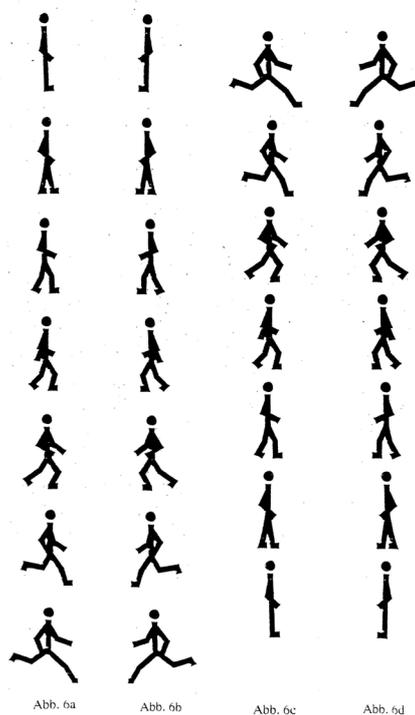
**Tabelle 6** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung bipolarer und unipolarer Skalen (Tabellen 4 und 5)

	bipolare Skala (-5 bis +5) in der Vertikalen (n = 77) %	bipolare Skala (-5 bis +5) in der Horizontalen (n = 90) %	Differenz
unsympathisch	23	12	-11
neutral	27	39	+12
sympathisch	50	49	-1
<hr/>			
	unipolare Skala (0 bis 10) in der Vertikalen (n = 79) %	unipolare Skala (0 bis 10) in der Horizontalen (n = 79) %	Differenz
unsympathisch	24	13	-11
neutral	45	42	-3
sympathisch	31	45	+14
<hr/>			
Differenz			
unsympathisch	+1	+1	
neutral	+18	+3	
sympathisch	-19	-4	

### 3.7 Die vertikale "Männchen-Skala"

Die vorgestellten Befunde sprechen dafür, dass die – durch zahlreiche empirische Untersuchungen belegte – Links-Rechts-Asymmetrie in der Wahrnehmung optisch dargebotener Stimuli offenbar auch auf Bildblattvorlagen im demoskopischen Interview zutrifft. Dies ist dann der Fall, wenn der Blick des Befragten beim Betrachten des Bildblatts entlang der Links-Rechts-Achse wandert. Entfällt diese Möglichkeit der "Blickwanderung", sollte folgerichtig auch die beobachtete Verzerrung im Antwortverhalten der Befragten entfallen. Zur Überprüfung dieser Annahme wurde die horizontale "Männchen-Skala" aus der ersten Testreihe um 90 Grad in die Vertikale gedreht, so dass das stehende Männchen vom linken Bildrand nach oben zu stehen kam und das mit ausladendem Schritt laufende Männchen vom rechten Bildrand nach unten (Abb. 6a). In der experimentellen Variation wurde die Skala statt auf der x-Achse auf der y-Achse gespiegelt. Eine Befragtengruppe erhielt ein Bildblatt mit nach rechts gewendeten Männchen, einer anderen Befragtengruppe wurde

Abbildung 6 Vertikale Männchen-Skala



ein Bildblatt mit nach links gewendeten Männchen vorgelegt (Abb. 6b). Die *vertikal* organisierte Skala unterschied sich in den Bildblattvorlagen somit ebenfalls nur durch die Laufrichtung der Männchen.<sup>9</sup>

Wie schon bei der horizontalen Skala liegt auch bei der vertikal organisierten Skala das Schwergewicht der Antworten (auf die Frage nach dem Tempo des technischen Fortschritts) in beiden Versionen der Bildblattvorlage auf den drei Skalenstufen "schnell", "sehr schnell" und "rasend schnell". Befragte, denen das Bildblatt mit den nach rechts gewendeten Männchen vorgelegt wurde, schätzten das Tempo des technischen Fortschritts mit je einem Drittel der Stimmen gleichermaßen als "sehr schnell" und "rasend schnell" ein, wohingegen Befragte, denen die Männchen-Skala mit den nach links gewendeten Männchen vorgelegt wurde, das Tempo des technischen Fortschritts mehrheitlich als "sehr schnell" einschätzten (40 Prozent gegenüber 23 Prozent "rasend schnell"). Die maximale Differenz zwischen den beiden Versuchsgruppen ist mit 8 Prozentpunkten zwar deutlich geringer als bei Verwendung der horizontalen "Männchen-Skala" (Tabelle 7, vgl. Tabelle 1), doch übt die Laufrichtung der Figuren offenbar auch bei der vertikal organisierten Skala einen Einfluss auf das Ergebnis aus.

<sup>9</sup> Die Befragten erhielten ein Bildblatt ohne den Männchen zugeordnete Ordinalzahlen (Vorgehen wie in Anm. 1 beschrieben), um die Vergleichbarkeit zu dem nachfolgend beschriebenen Experiment zu gewährleisten.

**Tabelle 7** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der vertikalen "Männchen-Skala"  
 – stehendes Männchen oben, schnell laufendes Männchen unten –

Frage: "Was würden Sie sagen, wie schnell der technische Fortschritt voranschreitet, wie schnell sich alles entwickelt? So kann man das natürlich nur schwer sagen, aber wenn Sie sich bitte einmal dieses Bild hier ansehen: das stehende Männchen oben würde bedeuten, die Technik macht überhaupt keine Fortschritte, und das schnell laufende Männchen unten würde bedeuten, die Technik macht rasende Fortschritte. Welches von den sieben Männchen würden Sie für den technischen Fortschritt heute typisch finden, welches Männchen würden Sie sagen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit nach rechts gewendeten Männchen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit nach links gewendeten Männchen.

		Laufrichtung der Männchen auf dem Bildblatt		
			nach rechts %	nach links %
Der technische Fortschritt kommt...				
... überhaupt nicht voran	(Stufe 1)	-	-	-
	Stufe 2	1	-	-1
	Stufe 3	3	3	-
	Stufe 4	9	11	+2
	Stufe 5	22	23	+1
... rasend schnell voran	Stufe 6	34	40	+6
	(Stufe 7)	31	23	-8
	%	100	100	
	n =	149	145	

IfP-MP, Umfrage WS 2001/02

In einem zweiten Experiment wurde die vertikale Anordnung der Figuren der "Männchen-Skala" umgekehrt. Das stehende Männchen kam nun nach unten zu stehen und das mit ausladendem Schritt laufende Männchen nach oben (Abb. 6c). Im Übrigen wurde die Links-Rechts-Versuchsordnung beibehalten: eine Befragtengruppe erhielt ein Bildblatt mit nach rechts gewendeten Männchen, eine andere Befragtengruppe erhielt ein Bildblatt mit nach links gewendeten Männchen (Abb. 6d).<sup>10</sup> Das schon bekannte Muster der Antwoorthäufung auf den drei letzten Skalenstufen wiederholt sich auch in dieser Bildblatt-Variation, doch gleichen sich die

<sup>10</sup> Wie bei dem zuvor beschriebenen Experiment enthielt das Bildblatt keine den Männchen zugeordneten Ordinalzahlen. Beide Experimente wurden nach dem Split-ballot-Verfahren in vier Untersuchungsgruppen derselben Befragung durchgeführt.

**Tabelle 8** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der vertikalen "Männchen-Skala"  
– stehendes Männchen unten, schnell laufendes Männchen oben –

Frage: "Was würden Sie sagen, wie schnell der technische Fortschritt voranschreitet, wie schnell sich alles entwickelt? So kann man das natürlich nur schwer sagen, aber wenn Sie sich bitte einmal dieses Bild hier ansehen: das stehende Männchen unten würde bedeuten, die Technik macht überhaupt keine Fortschritte, und das schnell laufende Männchen oben würde bedeuten, die Technik macht rasende Fortschritte. Welches von den sieben Männchen würden Sie für den technischen Fortschritt heute typisch finden, welches Männchen würden Sie sagen?"

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit nach rechts gewendeten Männchen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit nach links gewendeten Männchen.

	Laufrichtung der Männchen auf dem Bildblatt			
		nach rechts %	nach links %	Differenz
Der technische Fortschritt kommt...				
... überhaupt nicht voran	(Stufe 1)	-	-	-
	Stufe 2	1	-	-1
	Stufe 3	2	-	-2
	Stufe 4	4	5	+1
	Stufe 5	20	18	-2
... rasend schnell voran	Stufe 6	39	39	-
	(Stufe 7)	34	38	+4
	%	100	100	
	n =	150	120	

IfP-MP, Umfrage WS 2001/02

Ergebnisse in beiden Versuchsgruppen nun deutlich einander an. Befragte, denen das Bildblatt mit den nach rechts gewendeten Männchen vorgelegt wurde, unterscheiden sich in ihrer Tempo-Einschätzung des technischen Fortschritts kaum von den Befragten, die ihr Urteil anhand des Bildblatts mit den nach links gewendeten Männchen abgaben. Die maximale Differenz zwischen den beiden Versuchsgruppen beträgt 4 Prozentpunkte (Tabelle 8).

Die bei der *horizontal* organisierten "Männchen-Skala" beobachtete Verzerrung der Ergebnisse je nach Laufrichtung der Figuren auf der Bildblattvorlage, die vermutlich auf die "natürliche Blickbahn" der Befragten entlang der Links-Rechts-Achse zurückgeht, ist bei der *vertikalen* Ausrichtung der Skala zwar erkennbar schwächer ausgeprägt bzw. nahezu aufgehoben. Statt der Links-Rechts-Asymmetrie in der Wahrnehmung wirkt sich bei der vertikalen Organisation der Skala offenbar jedoch eine entlang der Senkrechten verlaufende Richtungstendenz in der Wahrnehmung im Sinne einer Unten-Oben-Asymmetrie aus.

Unabhängig von der "Bedeutung" der Figuren im Fragekontext scheint die oben gelegene Position gegenüber der unten gelegenen Position eher bevorzugt zu werden. Befragte, denen das Bildblatt mit den stehenden Männchen oben vorgelegt wurde, schätzten das Tempo des technischen Fortschritts häufiger als "langsam" ein als Befragte, die das Bildblatt mit den stehenden Figuren unten erhielten. Umgekehrt schätzten Befragte, die das Bildblatt mit dem rasenden Männchen oben erhielten, das Tempo des technischen Fortschritts häufiger als "rasend schnell" ein wie Befragte, denen das Bildblatt mit dem rasenden Männchen unten vorgelegt wurde. Mit einer Differenz von 15 Prozentpunkten in der Einschätzung des technischen Fortschritts als "rasend schnell" ist diese Richtungstendenz nach oben bei dem Skalen-Bildblatt mit den nach links gewendeten Männchen deutlich stärker ausgeprägt als bei dem Skalen-Bildblatt mit den nach rechts gewendeten Männchen (Tabelle 9). Gibt es neben dem Links-Rechts-Effekt mithin auch einen Unten-Oben-Effekt, der sich auf das Antwortverhalten der Befragten auswirkt?

**Tabelle 9** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der vertikalen "Männchen-Skala" (Tabellen 7 und 8)

Skalenorganisation	Voranschreiten des technischen Fortschritts		
	Langsam (Stufen 1-4) %	Schnell (Stufen 5 und 6) %	Rasend (Stufe 7) %
<i>Laufriichtung nach rechts</i>			
stehendes Männchen oben/ rasendes Männchen unten (n=149)	13	56	31
vs. stehendes Männchen unten/ rasendes Männchen oben (n=150)	7	59	34
<i>Differenz</i>	-6	+3	+3
<i>Laufriichtung nach links</i>			
stehendes Männchen oben/ rasendes Männchen unten (n=145)	14	63	23
vs. stehendes Männchen unten/ rasendes Männchen oben (n=120)	5	57	38
<i>Differenz</i>	-9	-6	+15

Anmerkung: Die Zusammenfassung der Skalenstufen folgt dem optischen Eindruck der Bewegungsdynamik der Figuren auf dem Skalenbildblatt.

### 3.8 Die vertikale "Kreise-Skala"

Weiteren Aufschluss in dieser Frage liefert die in der Testreihe horizontal organisierter Skalen ebenfalls bereits verwendete "Kreise-Skala". Macht es beim Betrachten des Bildblatts einen Unterschied, ob die – nun in der Vertikalen angeordneten Kreise – von unten nach oben zusammenlaufen (Abb. 7a) oder von oben nach unten (Abb. 7b)? Anhand der so variierten Bildblattvorlagen<sup>11</sup> in den beiden Versuchs-

<sup>11</sup> Wie beim Test der horizontalen "Kreise-Skala" erhielten die Befragten auch im vorliegenden Experiment ein Bildblatt ohne den Kreisen zugeordnete Ordinalzahlen (Vorgehen wie in Anm. 1 beschrieben).

## Abbildung 7 Vertikale Kreise-Skala

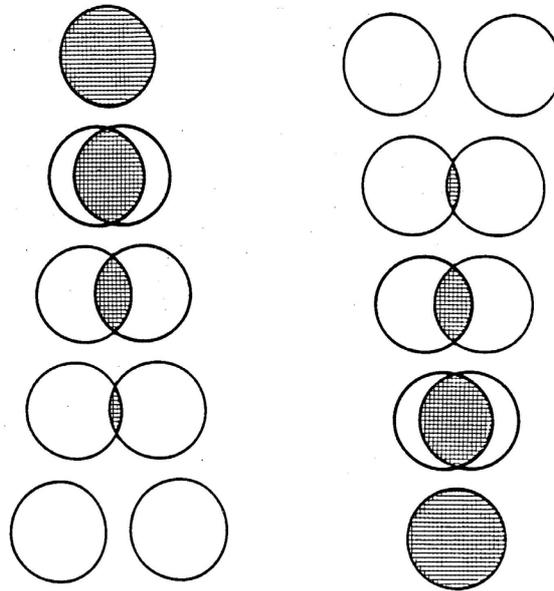


Abb. 7a

Abb. 7b

gruppen gaben die Befragten auf die (beibehaltene) Testfrage Auskunft, wie sehr sich Männer und Frauen ihrer Meinung nach in ihrem Fühlen gleichen oder unterscheiden.

Das Ergebnis untermauert die Vermutung eines Unten-Oben-Effekts. Befragte, denen das Bildblatt der "Kreise-Skala" mit den unten auseinander liegenden Kreisen (Männer und Frauen fühlen verschieden) und den oben deckungsgleichen Kreisen (Männer und Frauen fühlen gleich) vorgelegt wurde, äußerten mit einem Drittel der Stimmen die Ansicht, Männer und Frauen fühlten gleich. Erhielten die Befragten das Bildblatt mit den deckungsgleichen Kreisen unten und den auseinander liegenden Kreisen oben, halbierte sich der Stimmenanteil "gleich" auf 16 Prozent. Entsprechend verschoben sich, wenn auch in geringerem Umfang, die Aussagen, Männer und Frauen fühlten verschieden (Tabelle 10).

**Tabelle 10** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der vertikalen "Kreise-Skala"

*Frage: "Was meinen Sie, wie sehr Männer und Frauen sich in ihrem Fühlen, ihrem Gefühl unterscheiden? Wenn Sie es mir bitte nach diesem Bildblatt mit den Kreisen hier sagen: Die beiden weißen Kreise würden bedeuten, Männer und Frauen fühlen ganz verschieden, der dunkle Kreis würde bedeuten, sie fühlen ganz gleich. Was meinen Sie, wie sehr sich Männer und Frauen in ihrem Fühlen gleichen oder unterscheiden, welche Stufe würden Sie wählen?"*

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit den von unten nach oben zusammenlaufenden Kreisen vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit den von oben nach unten zusammenlaufenden Kreisen.

	Zusammenlaufen der Kreise auf dem Bildblatt		
	von unten nach oben %	von oben nach unten %	Differenz
Männer und Frauen sind in ihrem Fühlen...			
... verschieden (Stufen 1 und 2)	29	33	+4
... teils verschieden, teils ähnlich (Stufe 3)	39	51	+12
... gleich (Stufen 4 und 5)	32	16	-16
%	100	100	
n =	167	144	

IfP-MP, Umfrage SS 2001

Unabhängig von der Art der Bildblattvorlage und der Art der Fragestellung scheint es eine – mit der Links-Rechts-Blickbahn korrespondierende – Vorzugsrichtung in der Wahrnehmung offenbar auch entlang der Senkrechten zu geben. Im Unterschied zu der Links-Rechts-Asymmetrie ist dieser Zusammenhang empirisch bisher nicht gesichert. So wurden in einem von *Van der Meer* durchgeführten Experiment in einer anders gearteten Wahrnehmungssituation 322 Versuchspersonen nach ihrem Eindruck gefragt, ob ein Punkt im Zentrum eines Vierecks nach oben gehe oder nach unten. Für 119 Versuchspersonen zeigte der Punkt *keine* Bewegungstendenz, für 107 Versuchspersonen hatte der Punkt eine Bewegungstendenz *nach oben* und für 96 Versuchspersonen eine Bewegungstendenz *nach unten*. In einer anderen Versuchsanordnung, in der die Richtungstendenz von Strichen innerhalb eines Vierecks beurteilt wurde, zeigte sich über die variierten Anordnungen der Striche hinweg ein

leichtes Übergewicht der Einschätzungen zugunsten einer Richtungstendenz nach oben (*Van der Meer* 1958). Wenn es sich bei den hier vorgelegten Befunden um mehr als bloße Zufallsergebnisse handelt, sollte sich die Unten-Oben-Asymmetrie anhand einer weiteren Prüf-Skala erneut beobachten lassen.

### 3.9 Die vertikale "Stapel-Skala"

Die Überprüfung erfolgte anhand der streng seitensymmetrisch aufgebauten bipolaren "Stapel-Skala", nach der die Befragten ihre Sympathie für *Johannes Rau* einstuften. In der experimentellen Variation erhielt eine Befragtengruppe ein Skalen-Bildblatt mit dem positiven Pol (+5) oben und dem negativen Pol (-5) unten (Abb. 8a), einer anderen Befragtengruppe wurde ein Bildblatt mit umgekehrter Skalenanordnung vorgelegt, auf dem der positive Pol unten und der negative Pol oben zu liegen kam (Abb. 8b).<sup>12</sup>

Abbildung 8 Vertikale Stapel-Skala

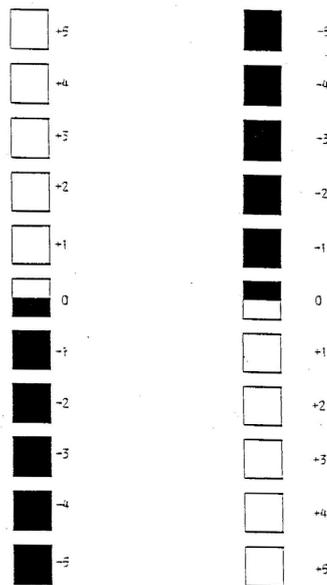


Abb. 8a

Abb. 8b

Zur besseren Anschaulichkeit werden die elf Skalenstufen in der Auswertung zu den vier Bewertungskategorien "sehr sympathisch" (Stufen +5, +4), "sympathisch"

<sup>12</sup> Wie bei dem oben beschriebenen Experiment zur Links-Rechts-Asymmetrie wurde auch in diesem Experiment die Bezifferung der Kästchen beibehalten.

(Stufen +3 bis +1), "unsympathisch" (Stufen -1 bis -3), "sehr unsympathisch" (Stufen -4, -5) sowie der neutralen Mittelposition (Stufe 0) zusammengefasst. Im Mittelfeld der gemäßigt positiven, gemäßigt negativen und ambivalenten Stimmen fällt das Urteil der Befragten in beiden Untersuchungsgruppen identisch aus. Jeweils die Hälfte der Befragten bekundete ihre Sympathie für **Johannes Rau** (Stufen +1 bis +3), 16 Prozent fanden ihn unsympathisch (Stufen -1 bis -3), jeder fünfte entschied sich für die neutrale Mitte (Stufe 0). Deutliche Unterschiede zwischen den Untersuchungsgruppen zeigen sich jedoch in den Extremwerten. Befragte, denen die Bildblatt-Skala mit dem Pluspol oben und dem Minuspol unten vorgelegt wurde, entschieden sich mit 15 Prozent der Stimmen für die Werte "sehr sympathisch" (Stufen +4, +5) gegenüber 1 Prozent "sehr unsympathisch" (Stufen -4, -5). Befragte, die das umgekehrte Bildblatt mit dem Pluspol unten und dem Minuspol oben erhielten, urteilten hingegen zu 8 Prozent "sehr sympathisch" gegenüber 6 Prozent "sehr unsympathisch" (Tabelle 11).

**Tabelle 11** Der Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei Verwendung der vertikalen "Stapel-Skala"

*Frage: "Eine Frage zu Bundespräsident Rau: Wie sympathisch ist Ihnen Johannes Rau? Vielleicht sagen Sie es mir nach diesem Bildblatt mit den weißen und schwarzen Kästchen hier: -5 würde bedeuten, er ist Ihnen ganz und gar unsympathisch, und +5, er ist Ihnen äußerst sympathisch. Welches Kästchen würden Sie sagen?"*

(Vorlage eines Bildblatts)

Im Split-ballot-Verfahren wurde einer Befragtengruppe ein Bildblatt mit den positiven Werten oben und den negativen Werten unten vorgelegt, die andere Befragtengruppe erhielt ein Bildblatt mit den positiven Werten unten und den negativen Werten oben.

	Anordnung der Skala auf dem Bildblatt		
	Pluspol oben Minuspol unten %	Minuspol oben Pluspol unten %	Differenz
Sympathieempfinden für Johannes Rau			
sehr unsympathisch (Stufen -5 und -4)	1	6	+5
Unsympathisch (Stufen -3 bis -1)	16	16	-
neutral (Stufe 0)	20	20	-
Sympathisch (Stufen +1 bis +3)	48	50	+2
sehr sympathisch (Stufen +4 und +5)	15	8	-7
%	100	100	
n =	150	120	

Obwohl die Differenz zwischen den Untersuchungsgruppen im Fall der streng seiten-symmetrischen "Stapel-Skala" geringer ausfällt als bei der vertikalen "Männchen-Skala" und der vertikalen "Kreise-Skala", weist das Ergebnis doch in dieselbe Richtung. Es spricht mithin einiges dafür, dass es neben dem beobachteten Links-Rechts-Effekt bei horizontal organisierten Skalen auch einen Unten-Oben-Effekt bei vertikal organisierten Skalen gibt.<sup>13</sup>

#### 4 Diskussion

Die vorgestellte Experimentalserie ging der Frage der Richtungsorientierung in der Wahrnehmung von Bildblattvorlagen im demoskopischen Interview nach. Anhand von fünf verschiedenen Skalenmodellen wurde geprüft, wie sich eine – auf die Richtungstendenz in der Wahrnehmung bezogene – Variation in der Präsentation horizontaler und vertikaler Skalen auf das Antwortverhalten der Befragten auswirkt (Tabelle 12).

Die Ergebnisse belegen einmal mehr die in der Methodenliteratur vielfältig beschriebene "große Wirkung scheinbarer Kleinigkeiten in der Umfrageforschung" (*Lamp* 2000), wie sie beispielsweise im Allensbacher Archiv durch mehrere Tausend Feldexperimente dokumentiert wird (vgl. *Petersen* 2002). So wenig die Problematik der Richtungsorientierung in der Wahrnehmung optisch dargebotener Stimuli hierbei in den Blick kommt, so bedeutsam erweist sie sich für die Fragebogenarbeit. Zwar stellen die in der vorliegenden Experimentalserie eingeschlossenen Bildblatt-Skalen nur einen kleinen Ausschnitt des in der Umfrageforschung zur Verfügung stehenden Skalenarsenals dar. Sie vermitteln aber bereits einen hinreichenden Eindruck sowohl von der Vielfalt und Anschaulichkeit der Skalen als auch von der damit verbundenen Problematik. Da rein formale Aspekte der Skalenorganisation offenbar eine größere Rolle spielen als bisher angenommen, stellt sich die Frage, ob der Entwicklung und Verwendung von Bildblatt-Skalen nicht doch engere Grenzen gezogen sind als "only by the imagination of the investigator and the necessity that they be commonly known to respondents", wie *Sudman* und *Bradburn*

---

<sup>13</sup> Das standardisierte Interview, in das die verschiedenen Skalenexperimente eingebunden waren, liefert keine Erkenntnisse über das "Verständnis" der Skalen. Insbesondere im Fall der vertikalen "Stapel-Skala" muss daher offen bleiben, wie die Befragten mit der gedrehten Skalenversion, in der die negativen Werte nach oben zu liegen kamen, umgegangen sind, da diese der Alltagserfahrung, bspw. bei der Einteilung des Thermometers, zuwiderläuft. *Dehaene* (1999) vermutet eine mentale Repräsentation numerischer Größen in Form eines Zahlenstrahls, der in der horizontalen Dimension von links (kleine Zahlen) nach rechts (große Zahlen) gerichtet ist und in der vertikalen Dimension von unten (kleine Zahlen) nach oben (große Zahlen) führt. Im Rahmen der kognitiven Methodenforschung entwickelte Verfahren zur Evaluation von Survey-Fragen (vgl. *Prüfer, Rexroth* 1996) könnten hierüber Auskunft geben.

**Tabelle 12** Experimentalserie zur Richtungsorientierung in der Wahrnehmung

Skala	Tabelle	Dynamik/ Statik	experimentelle Variation	maximale Differenz
"Männchen-Skala"	1	dynamisch	Links-Rechts- Umkehrung in der Horizontalen	18
"Kreise-Skala"	2	dynamisch	Links-Rechts- Umkehrung in der Horizontalen	8
"Baum-Skala"	3	statisch	Links-Rechts- Umkehrung in der Horizontalen	1
"Stapel-Skala"	4	statisch	Drehung aus der Vertikalen in die Horizontale	12
"Leiter-Skala" (modifiziert)	5	statisch	Drehung aus der Vertikalen in die Horizontale	14
"Männchen-Skala"	9	dynamisch	Unten-Oben- Umkehrung in der Vertikalen	15
"Kreise-Skala"	10	dynamisch	Unten-Oben- Umkehrung in der Vertikalen	16
"Stapel-Skala"	11	statisch	Unten-Oben- Umkehrung in der Vertikalen	7

(1982) darlegen. Aus den vorgestellten Befunden lassen sich dazu mindestens zwei Feststellungen ableiten, die sich zum einen auf die *Skalendynamik* beziehen und zum anderen auf die *Skalenrichtung*.

Der in der Testserie zunächst vermutete Einfluss der *Skalendynamik* auf das Antwortverhalten der Befragten konnte nicht erhärtet werden. Zwar haben sich mit abnehmender Dynamik der Skala die Ergebnisunterschiede zwischen den Versuchsgruppen verringert. Das bei der statischen "Baum-Skala" gefundene Null-Ergebnis scheint indessen eher auf andere Ursachen – die Konterkarierung der "natürlichen Blickbahn" der Befragten durch den seitlich stehenden Apfelbaum – zurückzugehen als auf die Aufhebung der *Skalendynamik*. Bei der Testung weiterer statischer Skalen hat sich das Null-Ergebnis nicht bestätigt. Da statische Skalen sich in der experimentellen Variation als ebenso „anfällig“ für Ergebnisverzerrungen erwiesen wie dynamische Skalen, kann die Frage nach dem Einfluss der *Skalendynamik* auf das

Antwortverhalten der Befragten vernachlässigt werden. Eine hierauf bezogene Forderung für die Skalengestaltung lässt sich aus den vorliegenden Ergebnissen nicht ableiten.

Anders stellt sich der Sachverhalt in der Frage der Skalen*richtung* dar. Unabhängig von der Skalendynamik oder -statik hat sich gezeigt, dass Bildblatt-Skalen offenbar einer Vorzugsrichtung in der Wahrnehmung Vorschub leisten. *Dieselbe* Skala führt zu unterschiedlichen Ergebnissen, je nachdem, ob sie in der Horizontalen oder in der Vertikalen organisiert ist bzw. eine Achsenführung von links nach rechts oder von rechts nach links bzw. von oben nach unten oder von unten nach oben besitzt. Diese Richtungsasymmetrie in der Wahrnehmung hängt vermutlich mit der Lese- und Schreibgewohnheit zusammen, die empirischen Untersuchungen zufolge als kulturelle Variable anzusehen ist. Experimente mit arabischen und israelischen Versuchspersonen, die entgegen unserer Konvention von rechts nach links lesen und schreiben, haben gezeigt, dass diese auch eine gegenläufige Richtungsorientierung in der Wahrnehmung besitzen. Für die Praxis der Umfrageforschung bleibt dies nicht ohne Konsequenzen, wie bereits das erste Beispiel der Experimentalserie, die horizontale "Männchen-Skala", deutlich macht.

Würde in einer deutsch-israelischen Untersuchung zum technischen Fortschritt den Befragten in Israel das gleiche (in der Laufrichtung der Männchen identische) Bildblatt vorgelegt wie den Befragten in Deutschland, liefen die Männchen in einem Land konform zur Lese- und Schreibrichtung, in dem anderen Land jedoch "gegen den Strich". Aufgrund der unterschiedlichen "natürlichen Blickbahn" würde die Skala von den Befragten beider Länder unterschiedlich wahrgenommen und besäße mithin eine verschiedene Bedeutung. Die Einschätzung des technischen Fortschritts durch die Befragten beider Länder wäre wegen der mangelnden Reliabilität des Messinstruments mithin nicht zu vergleichen. Kaum anders dürfte sich der Sachverhalt darstellen bei einer Untersuchung in Ländern mit horizontaler bzw. vertikaler Lese- und Schreibsozialisation. Die Ergebnisse der experimentellen Skalenvariation lassen vermuten, dass die gleiche (in der horizontalen oder vertikalen Dimension identische) Bildblatt-Skala für die Befragten beider Länder ebenfalls eine unterschiedliche Bedeutung besitzt.

Eine aus den vorliegenden Befunden abgeleitete Forderung für international vergleichende Umfragen, die Länder mit unterschiedlichen Lese- und Schreibrichtungen einschließen, könnte mithin darauf abzielen, Bildblattvorlagen der jeweiligen Lese- und Schreibkonvention anzupassen und verschiedenen Befragtengruppen ggf. richtungsvertauschte Bildblätter vorzulegen. Ein *nicht* umgekehrt vorgelegtes Bildblatt könnte das (Teil-)Ergebnis einer solchen Umfrage derart verzerren, dass es für

die vergleichende Analyse unbrauchbar wird. Ähnlichkeit oder Unähnlichkeit in den Einstellungen und Werthaltungen der Bevölkerung verschiedener Länder, die in einer vergleichenden Untersuchung gefunden werden, wären unter Umständen dann bloße Artefakte.

Ähnlich, wenn auch mit umgekehrten "Vorzeichen", stellt sich der Sachverhalt bei Zeitreihenuntersuchungen dar. Für den Ergebnisvergleich im Zeittrend schreibt der Methodenkanon der Umfrageforschung vor, dass sich bei der Wiederholungsbefragung nichts am Wortlaut einer Frageformulierung ändert, das Frageumfeld im Fragebogen gleich ist, die Erhebungstechnik beibehalten wird und anderes mehr. Nicht minder gerechtfertigt scheint die aus den vorliegenden Ergebnissen ableitbare Forderung, bei Verwendung von Fragemodellen mit Bildblattmaterial auch auf die Links-Rechts bzw. Unten-Oben-Anordnung der Vorlagen zu achten. Ein aus Unachtsamkeit (weil vordergründig ohne Bedeutung) richtungsverkehrt vorgelegtes Bildblatt kann das Ergebnis einer Umfrage derart verzerren, dass es für die Trendanalyse unbrauchbar wird. Dies gilt ebenfalls für die Adaption von Fragemodellen aus veröffentlichten Umfragen, bei denen die abgebildete Bildblattvorlage – sei es aus Versehen, sei es aus drucktechnischen Gründen – in der horizontalen oder vertikalen Dimension verkehrt wiedergegeben ist. Solange man sich über den Einfluss optischer Darstellungs- und Wahrnehmungsmuster bei der Verwendung von Bildblattvorlagen hinwegtäuscht, wird man sich auch über die Tragweite solcher vermeintlichen Nebensächlichkeiten, wie sie sich selbst in Standardwerken zur Umfrageforschung finden, keine Rechenschaft ablegen. Stabilität oder Wandel der Bevölkerungsmeinung, die in einer Trenduntersuchung gefunden werden, sind unter Umständen dann ebenfalls bloße Artefakte.

Es wäre indessen ein Missverständnis, wollte man die aufgezeigte Richtungsorientierung in der Wahrnehmung von Bildblattvorlagen und die hierdurch bewirkte Ergebnisverzerrung im demoskopischen Face-to-face-Interview als Argument gegen das persönlich-mündliche Interview und Plädoyer für die telefonische Befragung verstehen, bei denen *diese* Effekte naturgemäß keine Rolle spielen. Die unbestreitbaren Vorzüge von Telefonumfragen wiegen die eingeschränkten Erkenntnismöglichkeiten nicht auf, die aus dem (notwendigen) Verzicht auf komplexere Fragemodelle resultieren, die gerade die Stärke des persönlich-mündlichen Interviews ausmachen. Bildblattvorlagen sind häufig fester Bestandteil persönlich-mündlicher Befragungen, und diese gehören zum festen Bestandteil der Umfrageforschung. Die Strategie kann daher nur lauten, Schwachstellen zu identifizieren und Wege zu ihrer Behebung aufzuzeigen, um die Umfragemethode als Ganzes zu stärken. Die Frage, was ein Bild "im Auge des Betrachters" bewirkt und welchen Einfluss visuelle Stimuli auf die Meinungsbildung und das Gefühlsleben ausüben, ist noch lange nicht

geklärt (vgl. *Frey* 1999). Dass in dieser Hinsicht auch im Bereich der Umfrageforschung Klärungsbedarf besteht, hat die experimentelle Variation der verschiedenen Bildblatt-Skalen deutlich vor Augen geführt.

### Literatur

*Banich, M.T., Heller, W., Levy, J.:* Aesthetic preference and picture asymmetries. In: *Cortex* 25 (1989), S. 187-195.

*Beaumont, J.G.:* Lateral organization and aesthetic preference: The importance of peripheral visual asymmetries. In: *Neuropsychologia* 23 (1985), S. 103-113.

*Burt, D.M., Perrett, D.I.:* Perceptual asymmetries in judgments of facial attractiveness, age, gender, speech and expression. In: *Neuropsychologia* 35 (1997), S. 685-693.

*Cohen, A.S.:* Die Links-Rechts-Asymmetrie in der optischen Wahrnehmung mit besonderer Berücksichtigung von Schreib- und Lesegewohnheiten. Diss. Innsbruck 1973.

*Dehaene, S.:* Der Zahlensinn oder Warum wir rechnen können. Basel u.a. 1999.

*Ellis, A.W.; Miller, D.:* Left and wrong in adverts: Neuropsychological correlates of aesthetic preference. In: *British Journal of Psychology* 72 (1981), S. 225-229.

*Frey, S.:* Die Macht des Bildes. Der Einfluß der nonverbalen Kommunikation auf Kultur und Politik. Bern u.a. 1999.

*Gross, C.G., Bornstein, M.H.:* Left and right in science and art. In: *Leonardo* 11 (1978), S. 29-38.

*Hatta, T., Kawakami, A.:* Image generation and handedness: Is the hemi-imagery method valid for studying the hemisphere imagery generation process? In: *Neuropsychologia* 35 (1997), S. 1499-1502.

*Hippler, H.-J., Schwarz, N., Noelle-Neumann, E., Knäuper, B., Clark, L.:* Der Einfluss numerischer Werte auf die Bedeutung verbaler Skalenendpunkte. In: *ZUMA-Nachrichten* 28 (1991), S. 54-64.

*Lamp, E.:* Die Kunst zu fragen als Wissenschaft. Über die große Wirkung scheinbarer Kleinigkeiten in der Umfrageforschung. Öffentliche Antrittsvorlesung an der Universität Mainz am 10. Mai 2000.

*Lamp, E.:* Ist einer von drei gleich jedem Dritten? Der Einfluss numerischer Äquivalente auf die Wahrnehmung und Bewertung identischer Sachverhalte. In: *ZA-Information* 49 (2001), S. 49-68.

*Levy, J.:* Lateral dominance and aesthetic preference. In: *Neuropsychologia* 14 (1976), S. 431-445.

*Lieblich, A., Ninio, A., Kugelmass, S.:* Developmental trends in directionality of drawing in Jewish and Arab Israeli children. In: *Journal of Cross-Cultural Psychology* 6 (1975), S. 504-511.

*Maass, A.; Russo, A.:* Directional Bias in the Mental Representation of Spatial Events: Nature or Culture? In: *Psychological Science* 14 (2003), S. 296-301.

*McLaughlin, J.P.:* Aesthetic preference and lateral preferences. In: *Neuropsychologia* 24 (1986), S. 587-590.

*McLaughlin, J.P.; Dean, P.; Stanley, P.:* Aesthetic preference in dextrals and sinistrals. In: *Neuropsychologia* 21 (1983), S. 147-153.

*Nachson, I.:* Cross-cultural differences in directionality. In: *International Journal of Psychology* 16 (1981), S. 199-211.

*Nachson, I.:* Directional preferences in perception of visual stimuli. In: *International Journal of Neuroscience* 25 (1985), S. 161-174.

*Nachson, I., Argaman, E., Luria, A.:* Effects of directional habits and handedness on aesthetic preference for left and right profiles. In: *Journal of Cross-Cultural Psychology* 30 (1999), S. 106-114.

*Noelle-Neumann, E., Köcher, R. (Hrsg.):* Allensbacher Jahrbuch der Demoskopie 1984-1992. Band 9. München u.a. 1993. 1993-1997. Band 10. München u.a. 1997. 1998-2002. Band 11. München u.a. 2002.

*Noelle-Neumann, E., Petersen, T.:* Alle, nicht jeder. Einführung in die Methoden der Demoskopie. München 1996.

*Ostrom, T.M.:* Bipolar Survey Items: An Information Processing Perspective. In: *Hippler, H.-J., Schwarz, N., Sudman, S. (Hrsg.):* Social Information Processing and Survey Methodology. New York u.a. 1987, S. 71-85.

*Petersen, T.:* Das Feldexperiment in der Umfrageforschung. Frankfurt u.a. 2002.

*Prüfer, P., Rexroth, M.:* Verfahren zur Evaluation von Survey-Fragen: Ein Überblick. In: ZUMA-Nachrichten 39 (1996), S. 95-115.

*Ring, E.:* Signale der Gesellschaft. Psychologische Diagnostik in der Umfrageforschung. Göttingen u.a. 1992.

*Sakhuja, T., Gupta, G.C., Singh, M., Vaid, J.:* Reading habits affect asymmetries in facial affect of judgments: A replication. In: Brain and Cognition 32 (1996), S. 162-164.

*Sattler, J.B.:* Ikonographische und psychologische Aspekte der 'Seitigkeit' in der Kunst. Diss. München 1983.

*Schuman, H., Presser, S.:* Questions and Answers in Attitude Surveys. Experiments on Question Form, Wording, and Context. New York u.a. 1981.

*Sudman, S., Bradburn, N.M.:* Asking Questions. A Practical Guide to Questionnaire Design. San Francisco u.a. 1982.

*Sudman, S., Bradburn, N.M., Schwarz, N.:* Thinking about Answers. The Application of Cognitive Processes to Survey Methodology. San Francisco 1996.

*Schwarz, N., Knäuper, B., Hippler, H.-J., Noelle-Neumann, E., Clark, L.:* Rating Scales. Numeric values may change the meaning of scale labels. In: Public Opinion Quarterly 55 (1991), S. 570-582.

*Vaid, J., Singh, M.:* Asymmetries in the perception of facial affect: Is there an influence of reading habits? In: Neuropsychologia 27 (1989), S. 1277-1287.

*Van der Meer, H.C.:* Die Links-Rechts-Polarisation des phänomenalen Raumes. Eine experimentelle Untersuchung. Groningen 1958.

*Wölfflin, H.:* Über das Rechts und Links im Bilde. Paul Wolters zum 70. Geburtstag. In: Ders.: Gedanken zur Kunstgeschichte. Gedrucktes und Ungedrucktes. Basel 1940a, S. 82-90.

*Wölfflin, H.:* Das Problem der Umkehrung in Raffaels Teppichkartons. In: Ders.: Gedanken zur Kunstgeschichte. Gedrucktes und Ungedrucktes. Basel 1940b, S. 90-96.

*Worcester, R.M., Burns, T.:* A Statistical Examination of the Relative Precision of Verbal Scales. In: Journal of the Market Research Society 17 (1975), S. 181-197.

# **Labour Migration, Market Competition and Ethnocentrism: Guest workers in Israel and Germany**

**Ergebnisse eines Forschungsprojektes und Beschreibung des Datensatzes (ZA-Studien-Nr. 3974)**

**by Julia Iser and Silvia Keil**

## **1 Abstract of the research project**

The major objective of the research was to examine attitudes and values toward guest workers in Israel and Germany in a comparative perspective. The main objects of examinations are the social conditions which shape the normative system of beliefs, attitudes, stereotypes, prejudice and behaviour toward foreign workers and ethnic minorities.

The research was cast within theoretical frameworks that focus on the emergence of prejudice, discrimination and ethnic antagonism in modern societies. Two alternative, but by no means contradictory, models were used. The first is social psychological and centres on such concepts as social distance, prejudice, authoritarianism, and ethnocentrism. The second approach focuses on group competition over resources and rewards in national markets. Data to test the hypotheses were taken from national representative samples of adult citizens in Israel and Germany, respectively. The data was analysed using a variety of multivariate statistical models. It was our aim to contribute to a better understanding of the Israeli and German societies and to the theoretical knowledge regarding the social conditions that affect the system of values, attitudes, and behaviour toward migrant workers and ethnic minorities in modern societies. The comparison between Germany and Israel was one special point of interest, because Germany has been employing guest workers for more than a quarter of a century whereas Israel only recently began to rely on foreign workers.

The influence of guest worker communities on contemporary societies is one significant social and economic problem of modern states. The study tried to increase the understanding of these mechanisms and phenomena.

## 2 The data

The data sources are two national population surveys, one in Israel and one in Germany. The German data (sample size 3306) was collected in 1996 and the Israeli survey (sample size 1105), was done in 1999. The whole sample consists of  $N = 4411$  cases.

The German data (General Social Survey; ALLBUS) is a national representative survey, which has been conducted every two years since 1980. It is a random sample of the adult German population, supervised by ZUMA (centre for survey research and methodology). In the German data set the foreigners living in Germany have been excluded because they were not asked all the relevant questions.

Beside the demographic topics, the ALLBUS questionnaire in 1996 concentrates on the subject: "Attitudes towards ethnic groups in Germany". Therefore it is especially suited for analysing attitudes towards immigrants of different ethnic origins. We took a selection of the ALLBUS 1996.

The database consists of all people who were born before 1.1.1978 and who are living in private households in Germany. In a first step, a random sample of 162 communities had been taken from all communities in Germany. In a second step, a random selection of 40 addresses from the inhabitant lists of the chosen communities had been sampled.

The Israeli data is based on a survey of the adult Israeli population based on a stratified sample of 1105 respondents conducted in Israel during the second half of 1999. A questionnaire containing attitudes and beliefs towards immigrants and labour migrants was modelled after instruments developed in following studies: Eurobarometer Survey (Western Europe), the General Social Survey ALLBUS (Germany), and the survey organized by the Canadian Institute for Public Opinion. In addition, the questionnaire contained data on demographic, socio-economic and labour force characteristics. The interview was conducted face to face and lasted about 40 minutes. The rate of response was 50 percent, which is typical in the Israeli society.

There had been two different treatments for the missing values in the two data sets. The German data contained the value labels: does not apply, answer refused and no information as missing values. There were exclusively system missings in the Israeli data. In the merging file procedure all the missing values were transformed into system missings.

The two data files were merged in order to create a joint dataset. The variables were matched and coded in identical manner in order to conduct a comparative analysis.

In the German ALLBUS the East German Population was over-represented. So was the Arabic population in Israel. To correct this, the following analyses were done by using a weighting variable (variable name: weight).

### 3 Results and Publications

The two research teams from Israel and Germany published a number of articles in the last three years. They focus the common topic “Labour Migration, Market Competition and Ethnocentrism: Guest Workers in Israel and Germany” from different perspectives. Due to the different theoretical frameworks mentioned above some articles highlight the *explanatory* power of the split labour market approach whereas others focus on ethnocentrism and authoritarianism. Especially those articles that were written together did the innovate step to combine the two approaches.

A short overview over the output of two years of cooperation: *Peter Schmidt* and *Aribert Heyder* focused in their first two articles on the causal relationship between authoritarianism and ethnocentric attitudes in Germany. Their first article “Authoritarianism and Ethnocentrism in East and West Germany: Does the system matter?” deals with the differences between East and West Germany concerning the different types of political systems of the former GDR and West Germany, the second article “Autoritarismus und Ethnozentrismus in Deutschland: Ein Phänomen der Jugend oder der Alten?” focuses on the impact of age on hostile attitudes toward foreigners and guest workers.

The article “Konkurrenz, Bedrohung oder Bereicherung? Einstellungen gegenüber Ausländern in Deutschland und Israel” by *Julia Iser*, *Silvia Keil* and *Peter Schmidt* gives an descriptive insight in the German and Israeli societies concerning the level of Xenophobia in both countries.

Also descriptive but with the centre of attention on attitudes towards labour migration is the article “Labour migration in the public eye: Attitudes towards foreign workers in Israel” by *Moshe Semyonov* and *Rebeca Raijman*.

Next, these two authors and *Anat Yom Tov* published a paper, which takes into account the split labour market approach. “Labour Market Competition, Perceived Threat, and Propensity for Economic Discrimination” examines the impact of la-

bour market competition and fear of economic competition on propensity for economic discrimination against foreign workers in Israel.

The impact of perceived threat to social and economic well-being of individuals as well as threat to national identity on granting equal social rights to foreign workers is examined in the article “Perceived Threat and Exclusionary Attitudes towards Foreign Workers in Israel” by *Rebeca Raijman* and *Moshe Semyonov*.

“Labor Market Competition, Perceived Threat, and Endorsement of Economic Discrimination Against Foreign Workers in Israel” is the title and the topic of another article by *Moshe Semyonov*, *Rebeca Raijman* and *Anat Yom Tov*. Here the three researchers found support to socio-psychological explanations of discrimination, suggesting that fear of economic competition is more pronounced among the disadvantaged, subordinate, and vulnerable populations; and that support of economic discrimination against foreign workers is affected first and foremost by perceived threat.

“Do Foreigners Deserve Rights?” is the leading question of the likewise titled article by *Rebeca Raijman*, *Moshe Semyonov* and *Peter Schmidt*. They compare public views regarding equality of rights foreigners deserve. The comparison is interesting especially for two reasons. First, the incorporation of foreign workers is of special significance in both countries since their presence challenges the ethno-national definition of the two societies. Second, Germany has long been employing foreign workers while Israel has only recently begun relying- on labour migrants.

In “Population Size, Perceived Threat and Exclusion: A Multiple Indicators Analysis of Attitudes Toward Foreigners in Germany” *Moshe Semyonov*, *Rebeca Raijman*, *Peter Schmidt* and *Anat Yom Tov* challenge the long-standing theoretical proposition that hostility and discriminatory attitudes toward foreigners are likely to rise with the relative size of foreign population. By use of a sophisticated analytical model they combine the two theoretical approaches mentioned above. The results show that perceived size is associated with perceived threat, and perceived threat is found to mediate the relations between perceived size and support for exclusionary practices against foreigners.

#### 4 Publications

**Heyder, Aribert** and **Schmidt, Peter**, 2001: Autoritarismus und Ethnozentrismus in Deutschland: Ein Phänomen der Jugend oder der Alten? in: **U. Boehnke, D. Fuss** and **W. Hagan** (eds.), Jugendliche und Rechtsextremismus. Weinheim: Juventa Verlag, S. 119-142

**Iser, Julia; Keil, Silvia** and **Schmidt, Peter**, 2002: Konkurrenz, Bedrohung oder Bereicherung? Einstellungen gegenüber Ausländern in Deutschland und Israel. in: **C. Gansel** und **A.-P. Enslin** (Hrsg.), Literatur, Kultur und Medien. Facetten der Informationsgesellschaft. Berlin: Weidler Verlag, S. 185-203.

**Raijman, Rebeca** and **Semyonov, Moshe**, 2004: Perceived Threat and Exclusionary Attitudes Towards Foreign Workers in Israel, *Ethnic and Racial Studies*, forthcoming.

**Schmidt, Peter** and **Heyder, Aribert**, in press: Authoritarianism and Ethnocentrism in East and West Germany: Does the system matter? in: **R. Alba, P. Schmidt** and **M. Wasmer** (eds.), Germans and Foreigners. Theoretical Models and Empirical Analyses. New York: Palgrave Macmillan, S. 187-210

**Raijman, Rebeca; Semyonov, Moshe,** and **Schmidt, Peter;** 2003: Do Foreigners Deserve Rights? Comparing Public Views towards Foreigners in Germany and Israel. in: *European Sociological Review* 19, 379-392.

**Semyonov, Moshe; Raijman, Rebeca; Anat Yom Tov, Schmidt, Peter,** 2004: Population Size, Perceived Threat and Exclusion: A Multiple Indicators Analysis of Attitudes Toward Foreigners in Germany, accepted for publication in *Social Science Research*.

**Semyonov, Moshe; Raijman, Rebeca** and **Anat Yom Tov**, 2002: Labor Market Competition, Perceived Threat, and Endorsement of Economic Discrimination Against Foreign Workers in Israel. in: *Social Problems* Vol. 49, no 3, 416-431.

**Semyonov, Moshe; Raijman, Rebeca** and **Anat Yom Tov**: Labor Market Competition, Perceived Threat, and Propensity for Economic Discrimination." The paper was submitted to the Annual Meeting of the American Sociological Association, August, 2001. It was presented in Oxford by **Moshe Semyonov** in May 2001.

**Semyonov, Moshe; Raijman, Rebeca,** 2000: Labor Migration in the Public Eye: Attitudes Towards Foreign Workers in Israel. in: *ZA-Information* 47, 6-28.

## Archivierung und Sekundärnutzung qualitativer Interviewdaten – eine Machbarkeitsstudie

von Reiner Mauer

Die Akquisition und Bereitstellung *quantitativer* Forschungsdaten sowie deren kontinuierliche Aufbereitung und Dokumentation sind Kernaufgaben des Zentralarchivs. Eine vergleichbare Institution, die *qualitativen* Datenmaterial systematisch sammelt, archiviert, dokumentiert und einer wissenschaftlichen Sekundärnutzung zuführt, gibt es in Deutschland derzeit nicht. Die Archivierung des Materials liegt allein in der Verantwortung des einzelnen Forschers. Häufig lagern die Daten im Büro oder zu Hause, wo sie im Regelfall für andere nicht zugänglich sind und der dauerhafte Verbleib ungewiss ist.

Vor diesem Hintergrund fördert die DFG ein Gemeinschaftsprojekt des Zentralarchivs und des Archivs für Lebenslaufforschung der Graduate School of Social Sciences, GSSS, (Universität Bremen)<sup>1</sup>, das die Machbarkeit einer Serviceinfrastruktur für qualitative Forschung untersucht und überprüft, welche Anforderungen die scientific community an den Aufbau einer solchen Infrastruktur stellt.

In einem ersten Arbeitsschritt werden Leiter empirisch-qualitativ ausgerichteter, sozialwissenschaftlicher Forschungsprojekte nach Art und Umfang sowie den Nutzungsmöglichkeiten des von ihnen erhobenen qualitativen Datenmaterials befragt.

Das Sample der aktuell laufenden, schriftlichen Befragung wurde mit Unterstützung des Informationszentrums Sozialwissenschaften (IZ) aus der FORIS-Datenbank gewonnen. Dabei konnten für den Zeitraum von 1984 bis 2003 rund 18.000 mit qualitativen Methoden arbeitende Projekte identifiziert werden. Für die schriftliche Befragung wurden daraus alle Projekte der letzten zehn Jahre ausgewählt, die als ‚empirisch-qualitativ‘ gekennzeichnet sind, die als Methode zur Datengewinnung ‚qualitative Interviews‘ oder ‚Experteninterviews‘ verwendet haben und eine Angabe

---

1 Projektleiter: Prof. Dr. *K. Schumann* und Dr. *A. Witzel* (GSSS); Prof. Dr. *W. Jagodzinski* und Dr. h.c. *E. Mochmann* (ZA). Projektbearbeiter: *I. Medjedovic*, *D. Opitz* (GSSS), *R. Mauer* (ZA)

zum Projektleiter enthielten. Nach Überprüfung aller Adressen verblieben 1.750 Projekte mit 1.105 Projektleitern.

Die Befragung zielt zunächst darauf ab, den potentiell archivierbaren Datenbestand zu erfassen. Dazu gehören Fragen nach dem Verbleib der Daten, vorhandenen Datenformaten, dem Stand der Anonymisierung und der Datendokumentation sowie eine Einschätzung notwendiger Arbeitsschritte der Datenaufbereitung, um die Daten Dritten zur Verfügung stellen zu können. Darüber hinaus wird die Bereitschaft ermittelt, Daten aus beendeten bzw. zukünftigen Forschungsprojekten für eine Sekundäranalyse an andere weiterzugeben. Ein weiterer Fragenkomplex befasst sich mit der Sekundärnutzung von qualitativen Daten: Wer nutzte bzw. nutzt die eigenen Daten und für welche Zwecke? Hat der Befragte selbst schon einmal Daten sekundär genutzt, und welche Erfahrungen hat er dabei gemacht? Gibt es das Interesse, qualitative Daten zukünftig sekundär zu nutzen? Abschließend werden noch Fragen zum Interesse am Aufbau eines Archivs für qualitative Daten und zu den Leistungserwartungen an ein solches Archiv gestellt. Die Ergebnisse der Befragung werden in der nächsten Ausgabe der ZA-Information vorgestellt.

Im Anschluss an die schriftliche Befragung werden noch ergänzende, qualitative Interviews durchgeführt, um dann in einem weiteren Schritt Arbeitsroutinen, Kriterien und Konzeptionen für die Archivierung qualitativer Daten zu entwickeln. Die exemplarische digitalisierte Archivierung einer qualitativen Studie und ein Gutachten zum Datenschutz und der Anonymisierungsproblematik runden das Arbeitsprogramm der auf zwei Jahre angelegten Machbarkeitsstudie ab.

## **Dimensional Analysis**

### **34<sup>th</sup> Spring Seminar at the Zentralarchiv 14 Feb. - 4 March, 2005**

‘Dimensional Analysis’ was already part of the very first Spring Seminar in 1972. At that time, also ‘Causal Models’ were included. Over the years, both fields of data reduction have substantially grown. Each of the two is valid for a whole course of three weeks’ intensive teaching. In the coming year, we will focus on types of analysis that visualize data in terms of components, axes, dimensions, scales.

The Spring Seminar is a training course for social scientists interested in advanced techniques of data analysis and in the application of these techniques to data. Participants must have a sound basic knowledge of statistics as well as experience in the handling of PCs and of working with SPSS.

The Spring Seminar comprises lectures, exercises and practical work using personal computers. While in the lectures the logic of models and the corresponding analysis strategies will be explained, during the exercises and in the practical work the participants are given the opportunity to apply these methods to data. As in the past Spring Seminars, the focus will be on teaching multivariate analysis techniques. In 2005, the general topic will be Dimensional Analysis. In addition to the lectures, the participants will be provided with information about functions and services of the Zentralarchiv which is the German data archive for survey data.

The seminar covers three modules of one week each, to some extent based upon one another. The courses can be booked either separately or as a block: Dimensions in data: Are all data flat? Principal Component Analysis and Procrustes Analysis; Visualization of Multidimensional Data; Correspondence Analysis and Related Methods.

A more detailed programme will be published on ([www.gesis.org/za](http://www.gesis.org/za)) and will appear in the November issue of ZA-Information. Please make sure to refer to our web pages occasionally in order to follow new developments in the announcement.

The Zentralarchiv web server will provide up-to-date information on the Spring Seminar and other forthcoming events ([www.gesis.org.za](http://www.gesis.org.za)).

The lectures and workshops will be given in English by:

Dr. ***Garnt Dijksterhuis***

University of Groningen, The Netherlands

*Dimensions in data: Are all data flat?*

*Principal Component Analysis and Procrustes Analysis.*

14 - 18 February 2005

Professor Dr. ***Patrick J.F. Groenen***

Erasmus University Rotterdam, The Netherlands

*Visualization of Multidimensional Data*

21 - 25 February 2005

Professor Dr. ***Jörg Blasius***

University of Bonn, Germany

*Correspondence Analysis and Related Methods*

28 February - 4 March 2005

### **Fees and accommodation**

The participation fee is 50. -- € per week; no reduction nor refunds possible. For participants from non-university institutions, the fee is 200. -- € per week. Participants from Germany might be interested to learn that the Spring Seminar is recognized as „Bildungsurlaub“. Application forms to the employer can be requested from the Zentralarchiv. Travelling and accommodation have to be organized and paid by the participants.

### **Registration**

For your registration please use the attached registration form. You can also register via e-mail ([spring-seminar@za.uni-koeln.de](mailto:spring-seminar@za.uni-koeln.de)). For more details please check our web pages (address: see below). Participants will be accepted by order of application date, the number of participants is limited to 40 persons per week.

### **Address**

Zentralarchiv für Empirische Sozialforschung, Universität zu Köln, Bachemer Str. 40, D-50931 Köln, Germany.

***Maria Rohlinger***, scientific coordinator, e-mail: [maria.rohlinger@za.uni-koeln.de](mailto:maria.rohlinger@za.uni-koeln.de)  
Tel.: +49-221-4769445

Send all communication concerning the organisation to (secretariat):

[spring-seminar@za.uni-koeln.de](mailto:spring-seminar@za.uni-koeln.de)

Tel.: +49-221-47694-46 Fax: +49-221-47694-904

**Registration Form Spring Seminar 2005**

Zentralarchiv für Empirische Sozialforschung  
University of Cologne  
Postfach 41 09 60

50869 Köln  
Germany

**Registration for Spring Seminar 2005 (Please print or type)**

Name .....

Address .....

.....

.....

Phone/Fax .....

E-mail .....

Studied at .....

Examination .....

Present occupation.....

Institution .....

.....

.....

I want to participate in

- The whole Spring Seminar, 14 February – 4 March, 2005
- Parts of the Spring Seminar  
(Please tick appropriate box)
- Dr. **Garnt Dijksterhuis**  
University of Groningen, The Netherlands  
*Dimensions in data: Are all data flat?*  
*Principal Component Analysis and Procrustes Analysis*  
14 – 18 February 2005
- Prof. Dr. **Patrick J.F. Groenen**  
Erasmus University Rotterdam, The Netherlands  
*Visualization of Multidimensional Data*  
21 - 25 February 2005
- Prof. Dr. **Jörg Blasius**  
University of Bonn, Germany  
*Correspondence Analysis and Related Methods*  
28 February - 4 March 2005

Date: ..... Signature: .....

Each course is limited to 40 participants. Registration will be accepted by order of application date.

## Spring Seminar Lecturers 1972 - 2004

In the following, we list lecturers of the past 33 years:

**A**ish (1993), Allerbeck (1972, 1982), Andreß (1985, 1991), Armingier (1983, 1989)

**B**acher (1996, 2002), Blossfeld (1992), Böckenholt (1997), Borg (1990), Brüderl (2003)

**C**oxon (1976, 1996)

**D**e Leeuw, E. (1998), Diekmann (1985, 1997, 2001)

**E**ngel (1999)

**F**rick (1999)

**G**raff (1980), Greenacre (1994, 2002), Groenen (2002)

**H**agenaars (1991, 2000), Harder (1972), Heiser (1994), Herz (1974), Hox (1998), Hummell (1972, 1980)

**I**nglehart (1995)

**J**agodzinski (1979, 1982, 1986), Jöreskog (1993)

**K**aase (1973), Kappelhoff (1988), Klingemann (1972, 1973), Kristof (1990), Küchler (1978, 1981), Kühnel (1993, 2000), Küsters (1989)

**L**angeheine (2001), Lehnen (1981), Lohmöller (1984)

**M**ayer (1973, 1974, 1975), McCutcheon (1989, 1996, 2001), McFarland (1974), Meulemann (1982), Meulman (1994), Miller (1977), Müller (1974)

**N**orpoth (1975)

**P**appi (1972, 1973, 1980, 1983, 1988), Pickles (2004), Pötter (1999)

**R**einecke (2000), Rendtel (1992), Rohwer (1999), Rottleuthner-Lutter (1987), Rudas (2003)

**S**ahner (1972), Sanders (1995), Schmidt (1979, 1984, 1986, 1993), Schupp (1999), Snijders (1998, 2004), Sodeur (1988, 1990), Stein (2004), Stier (1992)

**T**horne (1987), Tilly (1975), Topf (1998)

**van der Ark** (2003), **van der Heijden** (1994), **van Deth** (1995), **Vermunt** (1997, 2003)

**W**eede (1975, 1978, 1979), **Wold** (1984)

**Z**iegler (1972)

## **ZHSF-Methodenseminar 2004**

### **Forschungsmethoden, Datenbankmanagement und Statistik in der Historischen Sozialforschung**

#### ***Repetitorium, Basis- und Aufbaumodule ("ZHSF-Herbstseminar")***

Das in 2003 erneut erweiterte Lehrangebot des ZHSF-Methodenseminars (seit 1980 bekannt als "ZHSF-Herbstseminar") ist modular strukturiert, d.h. es besteht aus in sich geschlossenen Lehreinheiten, die thematisch aufeinander abgestimmt sind. Die Module des ZHSF-Methodenseminars werden teils virtuell im Internet, teils vor Ort in Köln angeboten.

Virtuell werden angeboten: das dreimonatige Basismodul "Forschungsmethoden" (einschließlich der Propädeutika der Vor-Ort-Basis-Module "Datenbankmanagement" und "Deskriptive Statistik") und das zweimonatige "Follow-Up-Seminar" zu den Basismodulen. Das dreimonatige "Repetitorium" findet von März bis Mai im Internet statt.

Vor Ort in Köln werden angeboten: zwei zweitägige Einstiegsseminare zum Repetitorium bzw. zum virtuellen Basismodul, zwei viertägige Basismodule und zwei sechstägige Aufbaumodule.

Die Basismodule vermitteln die Grundlagen der Methodik Historischer Sozialforschung (Forschungsmethoden, Datenbankmanagement und Deskriptive Statistik). Sie können ohne besondere Vorkenntnisse besucht werden.

Die Aufbaumodule befassen sich mit der Theorie und Anwendung der multiplen Regression sowie verwandten Methoden des allgemeinen linearen Modells für metrische und diskrete abhängige Merkmale. Die Teilnahme an den Aufbaumodulen setzt Vorkenntnisse im Bereich der bivariaten Datenanalyse voraus.

#### **ZHSF-Methodenseminar 2004: Repetitorium**

Einstiegsseminar zum virtuellen Repetitorium:

20.02.-22.02.2004; vor Ort in Köln

Modul "Online-Repetitorium":

März bis Mai 2004; virtuell im Internet

**ZHSF-Methodenseminar 2004 ("ZHSF-Herbstseminar"): Basismodule**

Einstiegsseminar zum virtuellen Basismodul I:

07.05.-09.05.2004; vor Ort in Köln

Basismodul I "Forschungsmethoden":

Mai bis Juli 2004; virtuell im Internet

Basismodul II "Datenbankmanagement":

31.07.-04.08.2004; vor Ort in Köln

Basismodul III "Deskriptive Statistik":

04.08.-08.08.2004; vor Ort in Köln

Follow-Up-Seminar zu den Basismodulen:

August-Oktober 2004; virtuell im Internet

**ZHSF-Methodenseminar 2004 ("ZHSF-Herbstseminar"): Aufbaumodule**

Aufbaumodul I: "Allgemeines Lineares Regressionsmodell/Submodelle":

22.08.-27.08.2004; vor Ort in Köln

Aufbaumodul II: "Analyse diskreter abhängiger Merkmale":

30.08.-03.08.2004; vor Ort in Köln

**Gebühren****Repetitorium**

€ 50,- (Studenten und arbeitslose Teilnehmer: € 25,-).

**Basismodule**

€ 120,- für alle Basismodule (Studenten und arbeitslose Teilnehmer: € 60,-).

€ 50,- je Einzelmodul (Studenten und arbeitslose Teilnehmer: € 25,-).

**Aufbaumodule**

€ 120,- für beide Module inkl. Kursmaterialien (Studenten und arbeitslose Teilnehmer: € 60,-)

€ 75,- für ein Einzelmodul inkl. Kursmaterialien (Studenten und arbeitslose Teilnehmer: € 37,-).

**Anmeldung und Kontakt:**

Zentralarchiv für Empirische Sozialforschung

Abt. Zentrum für Historische Sozialforschung (ZHSF)

Liliencronstr. 6, 50931 Köln

+49 221 47694 -56, -35; +49 221 47694 55

zhsf-support@uni-koeln.de; WWW-Info: <http://zhsf.za.uni-koeln.de>

## **Daten und Datenbanken der Historischen Statistik am Zentralarchiv für Empirische Sozialforschung, Abteilung ZHSF**

**Bericht zum Workshop am 9. u. 10. Oktober 2003 in Köln**

**von Volker Müller-Benedict<sup>1</sup>**

Der Workshop unter Leitung von *Manfred Ehling* vom Statistischen Bundesamt und *Rainer Metz* vom ZA/ZHSF diente dazu, Personen zusammenzubringen, die an und mit Datenbanken zu quantitativen historischen Daten, insbesondere Zeitreihen, arbeiten. Dabei war vor allem daran gedacht, "laufende Projekte zur Historischen Statistik vorzustellen, und über Erfahrungen, Probleme und Zielsetzungen bei der Aufbereitung und Archivierung" (Ankündigungstext) zu sprechen.

Ziel des Workshops war weiterhin die Bildung einer Arbeitsgruppe, deren Aufgabe darin bestehen soll, die verstreuten Daten zur historischen Statistik langfristig am ZHSF in vergleichbarer und elektronischer Form zu archivieren und in einem "Handbuch" zu präsentieren. Über die Weitergabe dieser Daten an die Forschungsöffentlichkeit oder an das Statistische Bundesamt bzw. andere Ämter kann die Arbeitsgruppe schließlich auch Empfehlungen abgeben.

Nach einer Einleitung über diese Ziele durch *R. Metz* erinnerte *Manfred Ehling* daran, dass sich das Statistische Bundesamt schon früher einmal für die "Verlängerung" relevanter Zeitreihen nach "hinten", über den Beginn der BRD hinaus, interessiert hat, dass aber neue Anstöße nötig seien, um das Interesse des Bundesamtes an einer historischen Datenbank wieder zu wecken.

Anschließend wurden einzelne Projekte vorgestellt. *Andreas Kunz* (Institut für Europäische Geschichte, Mainz) erläuterte den tragischen, aber typischen Fall des DFG-Schwerpunktes „Quellen und Forschungen zur Historischen Statistik von Deutschland“ (1981-1991), dessen umfangreiche Datenerhebungen seit Projektende

---

<sup>1</sup> Dr. *Volker Müller-Benedict* ist Professor am Institut für Soziologie, Universität Leipzig

nur in derartig inkompatibler Form vorliegen, oder gar nicht veröffentlicht sind, dass die Aussichten, diese Daten jemals vollständig in eine Datenbank zu integrieren, eher skeptisch zu beurteilen sind. **Thomas Rahlf** (Herder-Institut, Marburg) stellte seine umfangreiche Link-Liste zu Daten der historischen Statistik vor. **Claude Diebolt** (Montpellier) erläuterte ein Projekt des französischen Wissenschaftsministers, das ähnliche Ziele wie das ZHSF verfolgt, sich aber erst im Anfangsstadium befindet. **Patrick Kammerer** (Forschungsstelle für Sozial- und Wirtschaftsgeschichte, ETH Zürich) erläuterte eine geplante Datenbankstruktur für ein Forschungsprojekt zur Wirtschaftsgeschichte der Schweiz. **Karin Orth** (Historisches Seminar der Universität Freiburg, Forschergruppe zur Geschichte der DFG) stellte die zur Geschichte der DFG zusammengestellte Datenbank vor, die jedoch noch nicht unmittelbar verfügbar gemacht werden soll. **Nele Bracke** (Universität Gent) präsentierte eine Datenbank, die Landkarten der Region Gent in verschiedenen historischen Stadien mit entsprechenden sozialhistorischen Daten verknüpft und die im Internet verfügbar sein soll. Voraussetzung dafür ist die Anwendung umfangreicher geografischer Software ([www.flwi.ugent.be/hisgis/nl/start\\_nl.htm](http://www.flwi.ugent.be/hisgis/nl/start_nl.htm)). **Holger Zinn** (Wiesbaden) referierte über Methoden zur Abgrenzung des geografischen Raums "Rhein-Main-Gebiet" an Hand von langfristigen Daten über Pendlerströme.

**Maximilian Stempfhuber** vom Informationszentrum Sozialwissenschaften in Bonn stellte die Datenbank ELVIRA vor, die Zeitreihen zur wirtschaftlichen Entwicklung der Elektrik- und Elektronikbranche in einer Vielzahl von Ländern für die Marktforschung beinhaltet. Sie ist z.B. für die Hintergrundrecherche über mögliche Investitionsstandorte anwendbar. Die Struktur dieser Datenbank erschien jedoch durch nur wenige Parameter (Länder, Branchen, Volumen) relativ festgelegt. Danach präsentierte **Alexander Berg** (data-quest \* Suchi & Berg GmbH, Göttingen) die vom ZHSF konzipierte und z.T. schon im Internet implementierte HISTAT-Datenbank ([www.histat.gesis.org](http://www.histat.gesis.org)), in der sich neben zahlreichen Studien auch bereits die Daten aus dem Datenhandbuch zur deutschen Bildungsgeschichte: Das Hochschulstudium in Preußen und Deutschland 1820-1944, Göttingen 1987 von **H. Titze**, **H.-G. Herrlitz**, **V. Müller-Benedict** und **A. Nath** befinden. Anschließend gab **Volker Müller-Benedict** (Universität Leipzig, Institut für Soziologie) zunächst einen kurzen Überblick über die vorhandenen und geplanten Datenbanken aus dem Bereich der historischen Bildungsstatistik und demonstrierte an einem Beispiel die praktischen Möglichkeiten eines Eingabe-Moduls, das im Moment in seinem Forschungsprojekt im Einsatz ist und für die HISTAT-Datenbank verfügbar gemacht werden kann. Der Vortrag von **Helmut Köhler** (Max-Planck-Institut für Bildungsforschung, Berlin) über den Stand der Daten zu Schulen in der BRD fiel aus. Die Beiträge endeten mit der Wunschliste von **Mark Spoerer** (Universität Hohenheim, Lehrstuhl für Wirt-

schafts- und Sozialgeschichte) für eine gute Versorgung mit historischen Daten, in der er eine Art Matrix vorschlug, in der historisch-statistische Merkmale verschiedener Art (Zeit, Region, wirtschaftliche, soziale u.a.) verknüpfbar sein sollen und als deren zentrale Zugriffseinheit die Erhebungsebene (Region) dienen soll.

Die Tagung zeigte, dass der Stand der Verfügbarkeit historischer Daten heute trotz Internet und der damit gegebenen Möglichkeiten noch immer sehr heterogen, zum Teil noch in Papierform und zum anderen Teil in nur schwer oder gar nicht lesbarer elektronischer Form gegeben ist (Beiträge von **Kunz, Rahlf, Orth**). Es zeigte sich aber auch, dass diejenigen, die von vorneherein an die spätere Verfügbarkeit ihrer Daten für die Öffentlichkeit – und d.h. nicht nur ausschließlich die scientific community – denken, ihren Daten eine wesentlich verlängerte Überlebenszeit sichern (Beiträge von **Diebolt, Bracke, Stempfhuber**).

In der Schlussdiskussion wurde vielfach der Wunsch geäußert, die verschiedenen Datenbestände zusammenzubringen und darüber hinaus auch technische Werkzeuge zu entwickeln, die ein späteres Zusammenführen des eigenen Datenbestands mit anderen Daten erleichtern. Auch die mögliche Zusammenarbeit mit Behörden sollte angestrebt und gefördert werden. Insgesamt wurde das ZHSF für die einzig mögliche Institution gehalten, die sich dieser Wünsche annehmen sollte. Daraufhin wurde eine "Arbeitsgemeinschaft Datenbanken der historischen Statistik" gegründet, an der sich einige Anwesende beteiligen werden und zu deren Sprecher **Rainer Metz** gewählt wurde.

Ansprechpartner im Statistischen Bundesamt:

Dr. **Manfred Ehling**, Statistisches Bundesamt, Institut für Forschung und Entwicklung in der Bundesstatistik

65180 Wiesbaden, Tel.: +49 611 752903, Fax: +49 611 753950

Mail: manfred.ehling@destatis.de

Ansprechpartner im ZA/ZHSF:

Prof. Dr. **Rainer Metz**, Zentralarchiv für Empirische Sozialforschung, Abt. ZHSF  
Universität zu Köln, Liliencronstr. 6, D-50931 Köln

Tel.: +49 221 4769436, Fax: +49 221 4769455

Mail: metz@za.uni-koeln.de

## Neue Studien im ZHSF

von **Gabriele Franzmann**

Das ZHSF (Zentrum für Historische Sozialforschung) ist eine Abteilung des Zentralarchivs für Empirische Sozialforschung. Zu seinen Aufgaben gehören die Akquisition, Aufbereitung, Archivierung und Dokumentation maschinenlesbarer Forschungsdaten, die im Rahmen historischer Studien entstanden sind. Die archivierten Studien sind unter der Web-Adresse [www.gesis.org/Datenservice/Suche/Daten/index.htm](http://www.gesis.org/Datenservice/Suche/Daten/index.htm) recherchierbar und können ausgeliehen werden. Handelt es sich bei den Daten der Studie um Zeitreihen zur Historischen Statistik<sup>1</sup>, dann werden diese vom ZHSF in die Internet-Datenbank „HISTAT“<sup>2</sup> importiert, womit die Zeitreihen online recherchierbar und downloadbar sind.

### Übersicht über die neu akquirierten Studien in der Zeit von Dezember 2003 bis Mai 2004

**ZA-Studien-Nr. 8169**

**Studientitel: Physical Stature of Georgia Convicts, 1770-1860.**

**Primärforscher: Komlos, John**

---

1 Zur Definition des Begriffs der „Historischen Statistik“ siehe *Metz, Rainer; Berg, Alexander* „Datenbanken zur Historischen Statistik im ZHSF“. In: *Historical Social Research (HSR)*, Vol. 29, No.2, S. 160-171; *Ehling, Manfred* (1996): *Historische Statistik – Probleme und Perspektive der internationalen Zusammenarbeit*. In: Statistisches Bundesamt (Hg.): *Wirtschaft und Statistik 1996/7*, S. 413-421; *Kunz, Andreas*: *Historische Statistik von Deutschland. Ein Forschungsschwerpunkt der Deutschen Forschungsgemeinschaft 1981-1991*. In: *Historical Social Research 22/2*, 1997, S. 236-249.

2 Der Name „HISTAT“ wurde erstmals Anfang der 1990er Jahre von *Andreas Kunz* und *Ulrike Albrecht* für eine zu entwickelnde Datenbank zur Historischen Statistik von Deutschland vorgeschlagen. Da das ZA/ZHSF an die damaligen Bemühungen anknüpft, erschien es nahe liegend, und auch von der Sache her gerechtfertigt, für unsere Datenbank denselben Namen zu verwenden; vgl. *Albrecht, Ulrike; Kunz, Andreas* (1990): *Building a Databank on German Historical Statistics*. In: *Metz, Rainer; Van Cauwengerghe, Eddy; van der Voort, Roel* (eds.), *Historical Information Systems. Proceedings Tenth International Economic History Congress*, Leuven, August 1990. Leuven: University Press, 77-86. *Kunz, Andreas* (1990): *Eine Datenbank zur Historischen Statistik von Deutschland*. In: *Diederich, Nils; Hölder, Egon; Kunz, Andreas* u.a. (Hg.): *Historische Statistik in der Bundesrepublik Deutschland*. Band 15 der Schriftenreihe des Statistischen Bundesamtes „Forum der Bundesstatistik“. Stuttgart: Metzler-Poeschel, S. 159-163. Eine Beschreibung des Serviceprojektes „Historische Statistik“ sowie der Zugang zur Datenbank HISTAT sind unter folgender Internet-Adresse zu finden: [http://www.gesis.org/Datenservice/ZHSF/Historische\\_Statistik/index.htm](http://www.gesis.org/Datenservice/ZHSF/Historische_Statistik/index.htm).

**Ziel und Inhalt:** Diese Datensammlung stellt die Körperstatur von Strafgefangenen des US-amerikanischen Staates Georgia dar, die sich zwischen 1770 und 1860 in Haft befanden. Die bereitgestellten Informationen umfassen: Alter bei Inhaftierung, Größenabmessungen in Fuß (feet) und Zoll (inches), Bundesstaat, in dem die Verurteilung erfolgte, Beruf, Geburtsort, Straftat, Monat und Jahr des Haftantritts, Monat und Jahr, in dem die Strafe auslief, eventuelle Begnadigungen und deren Datum.

**Schlagworte:** Anthropometrie, Körpermaße, Lebensstandard, Wohlstandsindikatoren, Strafgefangene, USA.

**ZA-Studien-Nr. 8170**

**Studientitel: Age at Menarche of Poor Viennese Women, 1907.**

**Primärforscher: Komlos, John**

**Ziel und Inhalt:** In dieser Studie wird die Relevanz des Ernährungszustandes für die Fruchtbarkeit unter Heranziehung von Daten über Familien, die unter Hungersnot leiden und von Armut geprägt sind, untersucht. Dabei soll getestet werden, ob der Beginn der Geschlechtsreife der Frau sowie die Geburtenrate von der Nahrungsversorgung beeinflusst werden.

Die Datensammlung bezieht sich auf das Alter sozial deklassierter Wiener Frauen beim Eintritt der Menarche<sup>3</sup> und untersucht die Beziehung dieses Zeitpunktes zur Fruchtbarkeit. Die Informationen über die Frauen umfassen: Alter, Familienstand, Geburtsort, Religion, Alter beim Eintritt der Menarche, Regelmäßigkeit und Dauer der Regel, Körpergröße in cm, Gewicht, Puls sowie die Anzahl von bisherigen Schwangerschaften. Weitere Daten beinhalten den Monat, in dem die Mutter in der Entbindungsklinik eintraf, das Geschlecht des Neugeborenen, die Größe des Neugeborenen in cm, das Geburtsgewicht des Neugeborenen in g, die Anzahl an Tagen, die die Mutter nach der Niederkunft in der Klinik verweilte, das Gewicht des Säuglings bei Verlassen des Hospitals und die vorherigen Krankheiten der Mutter.

**Schlagworte:** Anthropometrie, Körpermaße, Lebensstandard, Ernährungszustand, Wohlstandsindikatoren, Armut, Fruchtbarkeit.

---

<sup>3</sup> Fachausdruck für die erste Regelblutung

**ZA-Studien-Nr. 8171, HISTAT**

**Studientitel:** Bevölkerung in Deutschland ab 1815 / Datenkompilation auf der Basis von publizierten Textsammlungen und Untersuchungen, die sich auf amtliches statistisches Quellenmaterial stützen.

**Primärforscher:** *Sensch, Jürgen*

**Ziel und Inhalt:** Die wesentliche Zielsetzung der Datensammlung zur Bevölkerungsstatistik besteht darin, die Entwicklung der natürlichen und der geographischen Bevölkerungsbewegungen in der modernen deutschen Geschichte zusammenfassend nachzuzeichnen. Bevölkerung ist die nach biologischen Merkmalen (z.B. Geschlecht, Alter) und sozialen Gesichtspunkten (z.B. Familienstand, Beruf, Religionszugehörigkeit usw.) gegliederte Bevölkerungszahl innerhalb eines bestimmten, abgegrenzten Raumes zu einem bestimmten Zeitpunkt (Bevölkerungsstand, z.B. Einwohnerzahl zu einem bestimmten Zeitpunkt bzw. Stichtag oder die mittlere Bevölkerung eines bestimmten Jahres). Die Bevölkerungsbewegung bildet die Veränderung von Zahl und Struktur der Bevölkerung mit den Komponenten „generatives Verhalten“ und „Wanderungen“ ab. Das prozesshafte Geschehen wird wesentlich bestimmt durch die Elemente der Fruchtbarkeit und der Sterblichkeit (Heiratshäufigkeit, Geburtenhäufigkeit, Verstorbenen). Neben der natürlichen Bevölkerungsbewegung ist die geographische Bevölkerungsbewegung zu berücksichtigen, d.h. die Zu- und Abwanderungen (Wanderungsbewegungen). Die Datensammlung enthält neben den absoluten Eheschließungs-, Geburten- und Sterbefallzahlen auch die jeweils berechneten Verhältniszahlen, z.B. die Eheschließungs-, Geburten- und Sterbeziffern sowie die Wanderungsbilanzen. Es ist das Ziel der Datensammlung, die von der Forschung akkumulierten Daten zur Bevölkerungszahl, zur Bevölkerungsstruktur und zur natürlichen Bevölkerungsbewegung aus Primär- wie Sekundärquellen zusammenfassend darzustellen.

**Schlagworte:** Bevölkerung, Bevölkerungsbewegung, generatives Verhalten, Wanderung, Fruchtbarkeit, Sterblichkeit, Eheschließungen.

**ZA-Studien-Nr. 8172**

**Studientitel:** Der preußische Reichstagswähler 1912-1924.

**Primärforscher:** *Nöcker, Horst*

**Ziel und Inhalt:** Dem Vergleich zwischen dem jeweiligen Parteien- und Wahlsystem und den Ergebnissen der Wahlen in der Monarchie des deutschen Kaiserreichs einerseits, und in der demokratischen Staats- und Gesellschaftsordnung der Weimarer Republik andererseits, stand seit jeher nicht nur deren strukturelle und funktio-

nale Unterschiedlichkeit entgegen, sondern auch die Unvergleichbarkeit der statistischen Datengrundlagen. Es wird im Rahmen dieser Studie eine Methode der Zusammenstellung und Bearbeitung des publizierten amtlichen statistischen Materials entwickelt, durch die es erstmals möglich wird, direkte quantifizierende Vergleiche zwischen den numerischen Angaben über Wahlen und Wahlbevölkerung einerseits im Kaiserreich, andererseits in der Weimarer Republik anzustellen. Durch die Konstruktion von neuen statistischen Größen – den „Vergleichsgebieten“ – ist es dem Autor gelungen, eine alle wesentlichen Unterschiede in den statistischen Erhebungskriterien regionaler und struktureller Art aufhebende Erhebungsebene zu schaffen, die direkte, epochenübergreifende Vergleiche von Daten der Wahl-, Bevölkerungs-, Konfessions- und Berufsstatistik erlauben. Die insgesamt 277 Vergleichsgebiete wurden auf der Basis von Verwaltungskreisen des preußischen Staates in den Grenzen von 1925 gebildet. Somit wurde das Vergleichsgebiet als statistische Grundgröße einer neuen Aggregatebene zum ersten Mal in die historische Wahlforschung eingeführt.

**Schlagworte:** Parteiensysteme, Wahlsysteme, Reichstagswahlen, Monarchie, Deutsches Kaiserreich, Weimarer Republik, Wahlforschung.

**ZA-Studien-Nr. 8173**

**Studientitel: Ökonomik des privaten Haushalts anhand von Haushaltsrechnungen.**

**Primärforscher: *Pierenkämper, Toni***

**Ziel und Inhalt:** Die Ausgabenrechnungen privater Haushalte dienen als Instrumente zur Bestimmung der Veränderung des Lebensstandards der Bevölkerung und als sozialhistorische Quelle. Die wissenschaftliche Auseinandersetzung mit privaten Haushaltsbudgets beginnt schon Ende des 18. Jahrhunderts, um die Lage von Familien der ländlichen Unterschicht in England zu untersuchen. Erst ca. 60 Jahre nach den ersten Arbeiten in England begann auch auf dem europäischen Kontinent eine systematische Erhebung von Haushaltsrechnungen. Dabei wurden sehr detaillierte Familienbudgets erfasst.

Die Studie „Haushaltsrechnungen im 19. und 20. Jahrhundert“ dokumentiert Angaben zu den Budgets von 3696 Haushalten in der Zeit von 1850 bis 1927 aus verschiedenen Quellen. Die Daten beinhalten eine Beschreibung der aufgenommenen Haushaltsrechnung (das Jahr, die Währung, den Wohnort des Haushaltes, Art der Erhebung der Haushaltsrechnung: aus Befragungen oder aus privat geführten Rechnungsbüchern der Haushalte und die Erhebungshäufigkeit). Dem folgt eine Beschreibung des befragten Haushaltes (Zahl der Erwachsenen, Zahl der Kinder, Zahl

der arbeitenden Kinder, Anzahl sonstiger zum Haushalt gehörenden Personen, Beruf des Mannes, Beruf der Frau, Berufe der Kinder, Angaben zur Selbstversorgung z.B. durch Viehhaltung, Gartenbau, etc.). Anschließend werden die Einnahmen des Haushaltes aufgelistet sowie die Verwendung des Haushaltseinkommens für verschiedene Ausgabenposten (Nahrung, Miete, Kleidung, Kultur, Steuern, Freizeit, etc.) dokumentiert.

**Schlagworte:** private Haushaltsrechnung, Haushaltsausgaben, Haushaltsstatistik, Lebensstandard, Familienbudgets, Armut.

**ZA-Studien-Nr. 8174, HISTAT**

**Studientitel: Öffentliche Bildungsausgaben im internationalen Vergleich**

**Primärforscher: Diebolt, Claude**

**Ziel und Inhalt:** Auf der Grundlage von sozioökonomischen Zeitreihen (Bildungsausgaben, Wirtschaftswachstum und Bevölkerungszahl in Deutschland, 1829-1989, und Frankreich, 1820-1996) wird versucht, fraktional integrierte Modelle der Zeitreihenanalyse anzuwenden, um vorhandene Zyklen in den ausgewählten Zeitreihen zu identifizieren. Dieser Analyseansatz, der auf die vergleichende Analyse von Bildung, Wirtschaftswachstum und demographischen Reihen für Frankreich und Deutschland im 19. Jh. angewendet wird, ermöglicht die Identifizierung verschiedener zyklischer Komponenten der vorliegenden Zeitreihen. In der Studie wurden jeweils für Deutschland und Frankreich folgende Variablen erhoben: die legale Bevölkerungszahl (d.h. Bevölkerung mit deutscher bzw. französischer Staatsangehörigkeit), die Bildungsausgaben (laufend und in konstanten Preisen), die Schülerzahlen und das Bruttoinlandsprodukt.

**Schlagworte:** Zeitreihenanalyse, öffentliche Bildungsausgaben, Humankapital, Wirtschaftswachstum, Vergleich Deutschland – Frankreich.

**ZA-Studien-Nr. 8175, HISTAT**

**Studientitel: Die Expansion der Universitäten 1850-1992.**

**Primärforscher: Windolf, Paul**

**Ziel und Inhalt:** Die Expansion der Universitäten und ihr Strukturwandel während der letzten hundert Jahre sind das zentrale Thema der Untersuchung. Die Expansion der Universitäten zwischen 1870 und 1990 wird für fünf Länder analysiert: Deutschland, Italien, Frankreich, USA und Japan. Diese Länder haben während der letzten hundert Jahre den Prozess der Industrialisierung und Modernisierung ange-

führt und weisen ein in diesem Zeitraum stark expandierendes Hochschulsystem auf. Zwei Hypothesen, die das Wachstum erklären, werden getestet: Die Humankapital-Theorie und die Theorie der individuellen Statuskonkurrenz. Die Analyse konzentriert sich auf die Interaktion zwischen dem Bildungssystem und dem Wirtschaftssystem, insbesondere während einer ökonomischen Krise. Es wird festgestellt, dass die Bildungsexpansion während der letzten hundert Jahre als ein zyklischer Prozess verlief, in dem Auf- und Abschwünge einander ablösten.

Als Indikator für die Expansion des Bildungssystems dienten die Einschreibungsraten an den Universitäten. Die Schwingungen der Bildungsexpansion wurden von demographischen Effekten bereinigt, d.h. es wurde nur mit Bildungsquoten gerechnet. Die Bildungsquote wird definiert als der Anteil eines Jahrgangs, der die Hochschule besucht.

**Schlagworte:** Bildungssystem, Wirtschaftssystem, Bildungsexpansion, Konjunktur, Universitäten, Vergleich Deutschland – Frankreich – Italien – USA – Japan, Humankapital, individuelle, Statuskonkurrenz, Zeitreihenanalyse.

**ZA-Studien-Nr. 8176, HISTAT**

**Studientitel: Die Entwicklung der Aktienkreditbanken 1848-1913.**

**Primärforscher: *Burhop, Carsten***

**Ziel und Inhalt:** Die historische Entwicklung der deutschen Aktienkreditbanken vor Beginn des Ersten Weltkrieges wird sowohl auf nationaler als auch auf internationaler Ebene immer wieder diskutiert. Die traditionelle Wirtschaftsgeschichtsschreibung sieht in den großen Aktienkreditbanken eine entscheidende Institution für die Industrialisierung Deutschlands. Diese Hypothese konnte bisher lediglich für die letzten Dekaden des Kaiserreichs quantitativ belegt oder verworfen werden. Ziel der vorliegenden Arbeit ist, eine erstmals bis 1848 zurückreichende Statistik der deutschen Aktienkreditbanken zu erstellen. Bisher lagen lediglich Bilanzdaten mit einem begrenzten Zeitraum für die Jahre 1883 bis 1913 vor. Ausgehend vom Gründungsjahr der ersten deutschen Aktienkreditbank, des "Abraham Schaaffhausensche Bankvereins" in Köln im Jahr 1848, sollen die vorhandenen Datenreihen zur Entwicklung der Aktienkreditbanken ergänzt werden. Die auf mikroökonomischer Ebene erhobenen Daten werden zu einer gesamtwirtschaftlichen Statistik aggregiert. Die bereits publizierten Daten, vor allem die Zahlenreihen von *Walther G. Hoffmann* werden hier ergänzt.

**Schlagworte:** Banken, Aktienkreditbanken, Erster Weltkrieg, Industrialisierung.

**ZA-Studien-Nr. 8178, HISTAT****Studientitel: Bildungsausgaben in Berlin 1829-1914.****Primärforscher: *Diebolt, Claude***

**Ziel und Inhalt:** Die vorliegende Studie enthält die öffentlichen Bildungsausgaben für Berlin für die Zeit von 1829-1914. Die Bevölkerungszahl, die Netto-Ausgaben insgesamt sowie für den Posten Bildung werden aufgelistet. Die Bildungsausgaben werden untergliedert in Ausgaben für Primär- und Sekundärschulen. Bei der Berechnung der öffentlichen Bildungsausgaben in konstanten (d.h. inflationsbereinigten) Preisen wurde auf den Index für Großhandelspreise und auf den Preisindex für Lebenshaltungskosten zurückgegriffen, da die Bildungsausgaben sich zu einem großen Teil (ca. zwei Drittel) aus Löhnen zusammensetzen.

**Schlagworte:** Bildung, Bildungsausgaben, Bevölkerung, Berlin

Ansprechpartnerin:

***Gabriele Franzmann***

Zentralarchiv für Empirische Sozialforschung

Abteilung: Zentrum für Historische Sozialforschung, Köln

franzmann@za.uni-koeln.de

## **The Central and Eastern Eurobarometer 1990 - 1997: TRENDS CEEB 1-8**

### **The cumulative Data set is now available on CD-ROM**

**by Brigitte Hausstein, Wolfgang Zenk-Möltgen and Meinhard Moschner**

#### *Introduction: The Central and Eastern Eurobarometer Survey Series*

The Central and Eastern Eurobarometer (CEEB) survey series has been carried out on behalf of the European Commission (DG X) between 1990 and 1997 under the direction of *Karlheinz Reif* (until 1995) and *George Cunningham*, in total eight surveys. Once a year the CEEB surveys have been monitoring economic and political changes, and attitudes towards Europe and the European Union in up to 20 countries of the region.

The regular CEEB sample size is about 1000 respondents per country (see table 1). The samples are drawn among the citizens of the respective country, of 15 years and over. Respondents are interviewed face-to-face, in their private residences. Starting with CEEB 2 a multi-stage random probability sample design has been applied for all countries (except Hungary which adopted a quota sampling technique for CEEB 2). Starting with CEEB 5 (1994) the fieldwork and data integration was coordinated by GfK EUROPE Ad hoc Research together with regional partners, thereby replacing GALLUP UK.

The main survey results are published in the European Commission's CEEB reports. They are available at the ZA web page:

[http://www.gesis.org/en/datas\\_service/eurobarometer/ceeb/topics.htm](http://www.gesis.org/en/datas_service/eurobarometer/ceeb/topics.htm)

(please see references). The raw data are made available for social science research purposes by the ZA data service.

In October 2001, the European Commission has launched a new series of surveys in the 13 countries that are applying for European Union membership under the heading Candidate Eurobarometer (initially named "Applicant Countries Eurobarometer"), ordered and co-ordinated by the Directorate-General for Press and Communication (Public Opinion Analysis). The first CCEB survey was carried out in

**Table 1** Participating Countries and Number of Respondents

	CEEB 1 1990 ZA2493	CEEB 2 1991 ZA2251	CEEB 3 1992 ZA2321	CEEB 4 1993 ZA2474	CEEB 5 1994 ZA2577	CEEB 6 1995 ZA2802	CEEB 7 1996 ZA2924	CEEB 8 1997 ZA3068
Albania		1000	1049	1054	1034	1003	1013	
Mace- donia/FYROM			1002	1097	1000	1000	1000	
Slovenia			1063	1000	1086	1164	1114	1042
Bulgaria	1492	989	1312	1198	1045	1098	1035	1014
Romania		1000	1000	1176	1281	1141	1195	1044
Hungary	989	987	1000	972	1018	1004	1002	1080
Czechoslovakia January 1990 October 1990	1478 1490	1076						
Czech Republic			924	817	1062	1075	1021	967
Slovakia			734	684	995	1137	1066	1160
Poland	1014	1000	999	1004	1004	1000	1004	997
GDR May 1990 GDR May 1991	837 1085							
Greater Moscow	504							
European USSR	1561							
European Russia		975	1000	1377				
Russian Federation					1000	1178	1065	
Armenia			918	1000	1000	1000	1000	
Georgia		x	1038		1000	1059	1006	
Moldova			1000					
Belarus			1030	1143	1099	1021	1061	
Ukraine			1400	1171	1200	1199	1200	
Estonia		999	1000	1011	1002	1001	1071	1007
Latvia		999	1000	992	1000	1094	1017	1025
Lithuania		1000	1000	1020	1008	1003	1012	1013
Kazakhstan					1000	1000	1002	
Croatia						990	988	
Yugoslavia (Serbia/ Montenegro)							993	

x = The data are not available.

Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, Slovenia, Turkey and in the Republic of Cyprus.

(Electronic reports: [http://europa.eu.int/comm/public\\_opinion/cceb\\_en.htm](http://europa.eu.int/comm/public_opinion/cceb_en.htm))

*Available documentation*

CEEB1 (ZA2493): Surveys have not been based on a common basic questionnaire. The English codebook was based on the English version of the original country questionnaires as supplied by the Center for European Survey Analyses and Studies (ZEUS), University of Mannheim.

Original field questionnaires.

The integrated data set of the CEEB 1 was created by ZA with kind support of ZEUS, University of Mannheim.

CEEB2 (ZA2251): English Codebook provided by ZEUS.

The integrated data set of the CEEB 2 was created by ZEUS, University of Mannheim.

CEEB3 (ZA2321): English version of the Armenian questionnaire.

The integrated data set of the CEEB 3 was created by ZEUS, University of Mannheim

CEEB4 (ZA2474): Basic “short” Questionnaire (provided by GALLUP)

English version of the Hungarian questionnaire.

Original field questionnaires.

The integrated data set of the CEEB 4 was created by ZA with kind support of ZEUS, University of Mannheim

CEEB5 (ZA2577): Basic English questionnaire (GfK Edition) and field questionnaires.

The integrated data set of the CEEB 5 was created by GfK EUROPE (International Coordinator).

CEEB6 (ZA2802): Basic English questionnaire (GfK Edition) and field questionnaires.

The integrated data set of the CEEB 6 was created by GfK EUROPE (International Coordinator).

CEEB7 (ZA2924): Basic English questionnaire (GfK Edition) and field questionnaires.

The integrated data set of the CEEB 7 was created by GfK EUROPE (International Coordinator).

CEEB8 (ZA3068): Basic English questionnaire (GfK Edition) and field questionnaires.

The integrated data set of the CEEB 8 was created by GfK EUROPE (International Coordinator).

### *The CEEB TREND File 1990-1997 – the cumulative CEEB data set*

The CEEB TREND file consists of 125.875 cases and 280 variables. It is archived under the ZA Study number ZA3648. The trend variables for the cumulative data set were chosen in an iterative selection process. The selection was based on the CEEB 1 - 8 data sets which contain the integrated data of all participating countries per year<sup>1</sup>.

Starting point for the selection process were all those questions defined as a TREND by the European Commission's Eurobarometer unit. Furthermore only those questions were defined as "TRENDS" and included in the CEEB TREND File which were asked at least three times (years) in identical/similar wording. Country specific questions (asked only in a single country) and technical variables were not included (except some demographic, voting and weighting variables).

By omitting redundancies and simple transformations of original variables, demographics were reduced. Questions with unclear content and construction have been excluded. The complete documentation in the ZA Codebook Explorer also contains the available documentation of those dropped variables.

**Table 2** Dropped Variables

CEEB No.	Number of Variables	Number of dropped Variables
CEEB 1 (ZA2493) <sup>2</sup>	191	127
CEEB 2 (ZA2251)	246	169
CEEB 3 (ZA2321)	230	115
CEEB 4 (ZA2474)	85	9
CEEB 5 (ZA2577)	115	0
CEEB 6 (ZA2802)	145	16
CEEB 7 (ZA2924)	174	41
CEEB 8 (ZA3068)	130	33

1 The data file construction is based upon the works of *Meinhard Moschner* (ZA Cologne) and of the staff of the former ZEUS Institute at Mannheim University which was responsible for the yearly integration of the CEEB. The production of the TREND file and the codebook documentation was supported by the research assistance of *Daniel Seidel* (ZA Cologne).

2 CEEB 1 Surveys have not been based on a common basic questionnaire. Only a few common questions have been asked in all participating countries. Also the demographic part was mainly non-uniform. The integrated data set of the CEEB1 was created by ZA with kind support of ZEUS, University of Mannheim. The single country files (ZA Study Nos. ZA2252 to ZA2260) may include additional questions which have been asked in one country only.

**Table 3** Trends

	<b>Trend variable</b>	<b>Variable Name CEEB TREND File</b>
1	Country	V3
2	Year	V4
3	CEEB Number	V6
4	Permanently Resident	V8
5	Development in general	V9
6	Economic situation last 12 months	V10
7	Economic situation next 12 months	V11
8	Financial situation last 12 months	V12
9	Financial situation next 12 months	V13
10	Free Market Economy - right/ wrong	V14
11	Polit. discussion - convince friends	V15
12	Polit. discussion - frequency	V16
13	Economic reforms speed	V17
14	Satisfaction: Democracy development	V18
15	Satisfaction: Democracy working	V19
16	Human rights: Respect nowadays	V20
17	Frequency: Feeling European	V21
18	Impression: European Community/Union	V22
19	EC/ EU Info source	V23 to V57
20	Countries future tied up with	V58
21	EU membership: for/ against	V59
22	EC membership: favour/ opposed	V60
23	EC membership: Time period	V61
24	NATO membership: for/ against	V62
25	Countries future tied up with	V58
26	EC ties benefits	V63 to V72
27	EU relationship: Who benefits	V73
28	Left-Right self placement	V74 to V76
29	Right to vote in general elections	V77
30	Vote intention next election: Yes/ No	V78
31	Vote intention	V79 to V100
32	Vote intention inclined to	V101 to V117
33	Party voted for last national elections	V118
34	Sex	V119
35	Age of respondent	V120 to V124
36	Marital status	V125
37	Household size	V126
38	Education level	V127
39	Age education	V128
40	Occupation sector	V129 to V131
41	Occupation of respondent	V132
42	Occupation status	V133
43	Income - HH monthly net	V134 to V 198
44	Income - Respondent monthly	V199 to V206
45	National/ Ethnic Background	V207
46	Language spoken	V208 to V214
47	Region	V215 to V236
48	Type of Community	V237
49	Size of community	V238 to V242

As a result, the CEEB TREND File contains a total of 49 TRENDS (see table 3). The question wordings of all TRENDS were checked systematically and are documented in the ZA CodebookExplorer (as far as the documentation was available, see above). After selection of the TRENDS, the corresponding variables were identified. The number of selected trend questions has summed up to in total 280 variables (including technical questions: V1 ZA Study number, V2 ID, V5 Original ZA Study number; V7 Gallup Identification Number, V243 Date of interview, V244 Start of interview-time, V245 Duration of interview, V246 N of persons present during interview, V247 Respondent cooperation, V248 Original survey number: country, V249 Sampling point, V250 to V280 Weighting factors).

### ***Standardisation and Harmonisation of the TREND variables***

Some of the extracted TREND variables are based on instruments which differ from one survey to another and from one country to another. To integrate these variables into the CEEB TREND File they had to be standardized and harmonized with respect to:

1. identical variables names and labels
2. identical coding over time
3. identical coding across countries
4. coding of missing values according to the ZA standard.

Country specific variables were harmonized as far as it seemed to be appropriate. Recoded variables are documented in the ZA CodebookExplorer. After checking equivalence of country specific codes of the demographic trends the country specific categories were integrated. Slight changes of the codes are documented as well.

### ***ZA CodebookExplorer***

The data and meta data of the CEEB TREND File as well as the single integrated CEEB 1 to CEEB 8 files are provided on a CD-ROM. With the help of the exploration tool "ZA Codebook Explorer" the meta information at study and variable level is presented. This tool is founded on a data base and supports the ability to print codebooks, tables and graphs. First analyses and the creation of tables or graphs for frequencies, cross tabulations and comparative analyses are possible as well as several statistical procedures.



The integration of the CEEB TREND File into the ZA Codebook Explorer enables the user to structure the meta information in different “Explorer Views”. The CEEB TREND File database consists of four Explorer Views:

 **Studies**

 **CEEB TRENDS**

 **CEEB TRENDS: Countries**

 **User Defined Entries.**

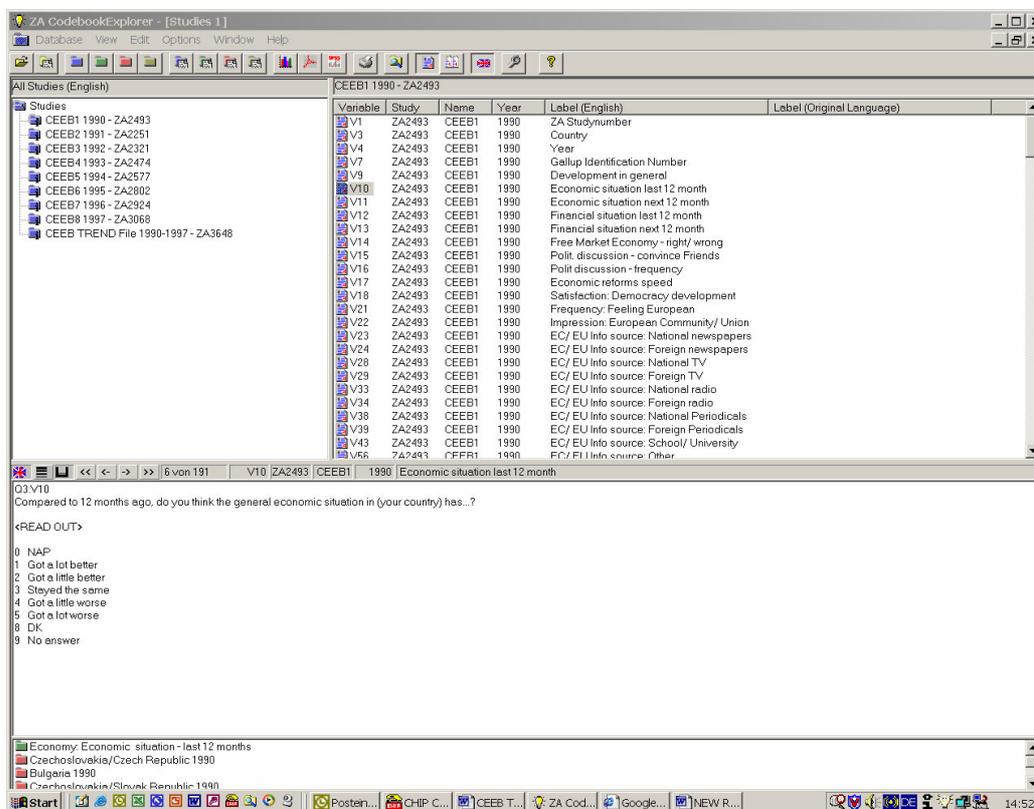
The *Studies View* provides a comprehensive overview of the questions which were asked in each CEEB as well as questions that were included in the CEEB TREND File. This *Study View* is composed of four elements which provide additional information on study and variable level (see figure 1):

1. List of included studies (CEEBS 1 to CEEBS 8 and CEEB TREND File) with the corresponding ZA Study numbers (on the left hand side of the window)
2. List of variables (e.g. V1 to V407) assigned to the marked study (e.g. CEEB1 1990 - ZA2493)
3. Variables window displays the answer and question texts for the marked variable (e.g. V10; see figure 1). Country specifics and other changes of the variables are documented here.
4. Below the Variable Window a list of other Explorer Views appears to which the variable is assigned (e. g. V10 in CEEB1 is assigned to the TREND View and Countries View)

Additional information on study level can be retrieved from the respective nodes within the Explorer View by opening the associated Acrobat Reader PDF-Files: ZA Study descriptions, English basic questionnaires (1990-1997) and selected original field questionnaires. Background information on the CEEB Study Series and the CEEB TREND File are available as well.

The Explorer View “*CEEB TRENDS*” contains only variables of the CEEB TREND File, the 49 TRENDS. The TRENDS are structured in 9 basic categories reflecting the main topics of the CEEB TREND File. With the help of the “*Countries View*” the user can identify the TRENDS according to countries and years in which they were conducted. The composition of both the CEEB TRENDS and the Countries is similar to the *Study View* except for the left part of the display. Instead of the list of studies the *CEEB TRENDS View* provides a list of the TREND variables and the *Countries View* displays the countries and years in which a question was asked.

**Figure 1** Explorer “View Studies”: CEEB1 1990 – ZA2251/V10



The “*User Defined View*” can be modified according to the user’s needs (entries can be added or renamed; the order of entries can be changed as well.).

Further features of the ZA Codebook Explorer are: simple and extended searches in a user-friendly manner and first analyses of the data. The analysis features includes besides other components a comparative analysis of variables and separate analysis of the integrated data sets according to country. Additionally the SPSS files are provided for a more detailed statistical analysis. Comprehensive information on the functionality offered by the Analysis feature can be found on the CD-ROM.

The CD-ROM is available at the GESIS Service Agency Eastern Europe in Berlin and ZA Cologne for the price of 50 EURO. For ordering please contact:



**Brigitte Hausstein**, Schiffbauerdamm 19, 10117 Berlin

Email: [hausstein@berlin.iz-soz.de](mailto:hausstein@berlin.iz-soz.de)

Phone: +49 (0)30 233611314

Or: **Meinhard Moschner**, Bachemer Str. 40, 50931 Köln

Email: [moschner@za.uni-koeln.de](mailto:moschner@za.uni-koeln.de)

Phone: +49 (0)221 4769421

### References:

**Hofrichter, J. and Weller, I.**, 1991: Central and Eastern Eurobarometer No.1. Public Opinion in Central and Eastern Europe – 1990, ZEUS Mannheim.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb1\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb1_report.pdf)

**Commission of the European Communities**, 1992: Central and Eastern Eurobarometer No.2. Public Opinion about the European Community. Ten Countries Survey - Autumn 1991. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb2\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb2_report.pdf)

**Commission of the European Communities**, 1993: Central and Eastern Eurobarometer No.3. Public Opinion about the European Community. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb3\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb3_report.pdf)

**European Commission**, 1994: Central and Eastern Eurobarometer No.4. Public Opinion and the European Union. 16 Countries' Survey 1993. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb4\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb4_report.pdf)

**European Commission**, 1995: Central and Eastern Eurobarometer No.5. Public Opinion and the European Union. 18 Countries' Survey 1994. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb5\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb5_report.pdf)

**European Commission**, 1996: Central and Eastern Eurobarometer No.6. Public Opinion and the European Union. 19 Countries' Survey 1995. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb6\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb6_report.pdf)

**European Commission**, 1997: Central and Eastern Eurobarometer No.7. Public Opinion and the European Union. 20 Countries' Survey 1996. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb7\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb7_report.pdf)

**European Commission**, 1998: Central and Eastern Eurobarometer No.8. Public Opinion and the European Union. 10 Countries' Survey 1997. Brussels.

[http://www.za.uni-koeln.de/data/en/eurobarometer/ec\\_reports/ceeb8\\_report.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/ec_reports/ceeb8_report.pdf)

**Luijckx, R., Brislinger, E. and Zenk-Möltgen, W.**, 2003: European Values Study 1999/2000 – A Third Wave: Data, Documentation and Database on CD-ROM, in: ZA-Information 52.

**Moschner, M.**, 1999: Central and Eastern Eurobarometer Variable Synopsis. Zentralarchiv Köln. [http://www.gesis.org/en/data\\_service/eurobarometer/ceeb/ceeb1-8\\_variables.pdf](http://www.gesis.org/en/data_service/eurobarometer/ceeb/ceeb1-8_variables.pdf) [31.03.2004]

**Scholz, E., Schmitt, H.**, 2001: The Mannheimer Eurobarometer Trend File 1970-1999. Codebook and Un-weighted Frequency Distribution.

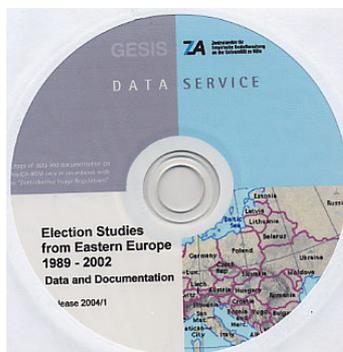
[http://www.za.uni-koeln.de/data/en/eurobarometer/eb\\_trends/eb\\_trends\\_1970-1999\\_cod.pdf](http://www.za.uni-koeln.de/data/en/eurobarometer/eb_trends/eb_trends_1970-1999_cod.pdf) [31.03.2004]

**ZA CEEB** web pages: [http://www.gesis.org/en/data\\_service/eurobarometer/ceeb](http://www.gesis.org/en/data_service/eurobarometer/ceeb) [31.03.2004]

**Zenk-Möltgen, W. and Mochmann, E.** 2000: Der Continuity-Guide der deutschen Wahlforschung und der ZA CodebookExplorer, in: Klein, M. et al. (Hrsg.): 50 Jahre empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten. Wiesbaden: Westdeutscher Verlag.

# Election Studies from Eastern Europe 1989 - 2002 New Release

## Data and Documentation on CD-ROM



The CD-ROM contains the data and documentation of 63 studies on elections and political attitudes from 15 East European countries, spanning 1989-2002. The data are prepared for secondary analysis according to the archival standards and are available with the documentation (study descriptions, questionnaires, codebooks, and information on publications) as well as back ground information on election dates in Eastern Europe. In addition it is possible to search the questionnaires and codebooks by using the Adobe Acrobat Index available on the CD.

The CD-ROM is designed to provide the community with a comprehensive product which is a source of readily understandable and accessible data and background information on election studies from Eastern Europe.

The interface is designed in HTML format and it allows the user to browse the study descriptions, codebooks and questionnaires just with a mouse click. With the exception of the study descriptions (they are in HTML format) all documents

 The screenshot shows the GESIS website interface. The main heading is "Election Studies from Eastern Europe 1989 - 2002 Data and Documentation". Below this, there is a table listing various studies. The table has columns for Title, Sample, Variables, ZA No., Access, and Documents.
 

Title	Sample	Variables	ZA No.	Access	Documents
Final Month after Referendum Albania 1995	1210	3	ZA2011	A	Study Description Questionnaire
Belarus					
Parliamentary Election in Belarus 1995	1000	118	ZA2017	A	Study Description Questionnaire
Bulgaria					Country Report
Individual Modernization in Bulgaria 1988	1943	125	ZA1944	A	Study Description Questionnaire
Public Opinion in Bulgaria - November 1990	3216	136	ZA1945	A	Study Description

are in PDF format and can be viewed and printed with the Adobe Acrobat Reader available on the CD-ROM. The data are available as SPSS files (system files as well as portable files) ready-made for further analyses.

The CD-ROM is available at the GESIS Service Agency Berlin and the ZA Cologne for the price of 25 EURO. For ordering, please contact:

**Brigitte Hausstein**, Schiffbauerdamm 19, 10117 Berlin

Email: hausstein@berlin.iz-soz.de

Phone: +49 (0)30 233611314

Or: **Eberhard Riedel**, ZA, Bachemer Str. 40, 50931 Köln

Email: riedel@za.uni-koeln.de

Phone: +49 (0)221 4769464

---

## **ZA-Direktor wird Mitglied des Center of Excellence in Nishinomiya**

**Wolfgang Jagodzinski** ist zum auswärtigen Mitglied des Center of Excellence (CoE) an der Sozialwissenschaftlichen Fakultät der Kwansei Gakuin Universität in Nishinomiya ernannt worden, das vom japanischen Ministerium für Bildung, Kultur, Sport, Wissenschaft und Technik eingerichtet worden ist. Er hat in dieser Eigenschaft eine vierwöchige Gastprofessur in Japan wahrgenommen und Vorträge auf dem ersten Symposium des CoE, an der Sozialwissenschaftlichen Fakultät und an der Betriebswirtschaftlichen Fakultät in Nishinomiya, im Graduiertenkolleg für Internationale Studien an der Universität Kobe und im Institut für Mathematische Statistik in Tokio gehalten.

## Der European Social Survey (ESS) 2002/2003

von Katja Neller<sup>1</sup>

### 1 Ziele des ESS

Das generelle Ziel des ESS ist die Entwicklung, inhaltliche Konzeptualisierung und Durchführung einer in Bezug auf die verwendeten Erhebungsinstrumente in der aktuellen Forschung verankerten und methodisch nach den strengstmöglichen Qualitätsstandards organisierten Studie zum Einstellungs- und Wertewandel in verschiedenen Ländern. Im Mittelpunkt stehen repräsentative Bevölkerungsbefragungen, die idealerweise im Zweijahresrhythmus wiederholt werden sollen. Die Langfristforschungsperspektive des Projektes besteht darin, die Interaktion zwischen den sich wandelnden politischen und ökonomischen Institutionen und den Einstellungen, Überzeugungen und Verhaltensmustern der Bevölkerungen der jeweiligen Länder zu beschreiben und zu erklären.

Um eine optimale Vergleichbarkeit der Erhebungsergebnisse zu erreichen, wurden alle Teilschritte des Projektes, von der Stichprobenziehung über die Übersetzung des englischen Ausgangsfragebogens in die Sprachen der Teilnehmerländer bis zur konkreten Durchführung der Erhebungen im Feld, soweit möglich standardisiert, von einer zentralen Steuerungsinstanz (dem Central Coordinating Team, s.u.) gegengeprüft und – einschließlich der auftretenden Abweichungen – umfassend dokumentiert.

An der ersten Welle des ESS, deren Befragungen im Zeitraum September 2002 bis Dezember 2003 stattfanden, nahmen 22 Nationen teil. Neben Deutschland waren dies Belgien, Dänemark, Finnland, Frankreich, Griechenland, Großbritannien, Irland, Israel, Italien, Luxemburg, Niederlande, Norwegen, Österreich, Polen, Portugal, Schweden, Schweiz, Slowenien, Spanien, die Tschechische Republik und Ungarn. Mit Ausnahme Frankreichs sind inzwischen für alle Teilnehmerländer Daten verfügbar. Diese sind für alle Interessierten über eine Projekthomepage (s.u.) frei zugänglich.

---

<sup>1</sup> *Katja Neller* ist Geschäftsführerin für die abgeschlossene erste sowie die laufende zweite Welle der deutschen Teilstudie des European Social Survey (<http://www.europeansocialsurvey.de>). Universität Stuttgart, Institut für Sozialwissenschaften, Politikwissenschaft I, Breitscheidstr. 2, 70174 Stuttgart. Email: [katja.neller@po.pol.uni-stuttgart.de](mailto:katja.neller@po.pol.uni-stuttgart.de).

lich. Da in Frankreich die Feldarbeit im Dezember 2003 abgeschlossen wurde, ist die Komplettierung des internationalen Datensatzes für den Sommer 2004 zu erwarten.

## 2 Projektidee und Organisation

Die Idee zum European Social Survey-Projekt (ESS) entstand in der European Science Foundation (ESF). Dank der finanziellen und inhaltlichen Unterstützung der Europäischen Kommission, der ESF und nationaler forschungsfördernder Institutionen entwickelte sich das Projekt rasch weiter. Das Projekt soll in den nächsten Jahren fortgesetzt werden. Die Finanzierung wird durch das 5. Rahmenprogramm der Europäischen Kommission, die ESF und nationale Institutionen zur Forschungsförderung, im Falle von Deutschland die Deutsche Forschungsgemeinschaft (DFG), gesichert.

Wesentliche Institutionen innerhalb des ESS sind u.a. das international besetzte Central Coordinating Team (CCT) unter Leitung von *Roger Jowell* (City University, London), das allgemeine Koordinations- und Organisationsaufgaben wahrnimmt, und die sog. National Coordinators, die für die Durchführung der Studien in den einzelnen Ländern zuständig sind. Die erhobenen Daten wurden vom norwegischen Datenarchiv "Norwegian Social Science Data Services" (NSD) aufbereitet.

## 3 Themen

Im Rahmen des ESS wird ein einheitlicher Fragebogen zu verschiedenen Aspekten des politischen und gesellschaftlichen Zusammenlebens verwendet, der durch länderspezifische Fragen ergänzt wird. Der Fragebogen der ersten Welle bestand aus insgesamt vier Modulen, von denen zwei den festen Kern des ESS ausmachen. Zwei Module waren als variable, d.h. von Welle zu Welle wechselnde, Komponenten konzipiert. Während der Kernfragebogen die Beobachtung von Stabilität und Wandel für eine Reihe von sozio-ökonomischen, sozio-politischen, sozialpsychologischen und sozio-demographischen Variablen ermöglicht, dienen die variablen Module dazu, spezielle Forschungsfragen oder aktuelle politische Themen in die Befragungen einzubinden. Die Auswahl der variablen Module des ESS erfolgte über ein vom CCT ausgeschriebenes Wettbewerbsverfahren, bei dem sich international besetzte Wissenschaftlergruppen mit Themenvorschlägen und Fragebogenentwürfen bewerben konnten.

Die beiden Kernmodule bestehen zum größten Teil aus bereits erprobten, theoretisch relevanten Fragen nach verschiedenen individuellen Orientierungen und Verhaltensweisen (z.B. Mediennutzung, soziales Vertrauen, politisches Interesse, politische Kompetenz, politisches Vertrauen, Wahlbeteiligung und andere Formen poli-

tischer Partizipation, Parteibindung, soziopolitische Orientierungen, subjektives Wohlbefinden, soziale Exklusion, Religion, subjektiv empfundene Diskriminierung, nationale und ethnische Identität).

Die variablen Module der ersten Welle waren "Citizenship, Involvement, Democracy" und „Immigration“. Die Items des Wechselmoduls Citizenship, Involvement, Democracy, das von *Ken Newton, José Ramon Montero, Anders Westholm, Hanspeter Kriesi* und *Sigrid Roßteutscher* entwickelt wurde, lehnen sich in weiten Teilen an die Frageformulierungen aus dem Projekt "Citizenship, Involvement, Democracy" an<sup>2</sup>. Das Modul umfasst Instrumente zur differenzierten Erhebung von Mitgliedschaft und Involvierungsgrad in verschiedenen Freiwilligenorganisationen, zur Erfassung familiärer und freundschaftlicher Bindungen sowie Fragen zum Profil eines "guten Bürgers" und zum Arbeitsumfeld der Befragten.

Das Wechselmodul "Immigration", das von *Ian Preston, Thomas Bauer, David Card, Christian Dustmann* und *James Nazroo* entwickelt wurde, beinhaltet sehr differenzierte Fragen zu Einwanderungs- und Asylthemen, u.a. Einstellungen zu, Wahrnehmungen von und Vorurteile gegenüber Zuwanderern sowie entsprechende Politikpräferenzen der Befragten.

Der als Selbstausfüller administrierte Zusatzfragebogen umfasste neben den Fragen der *Schwartz*-Werteskala wechselnde Testfragen für Methodenexperimente in insgesamt sechs Splitversionen.

#### **4 Soziodemographie, Kodierungsstandards, abgeleitete Indizes, Gewichtung**

Generell ist die Erhebung soziodemographischer Variablen im Rahmen des ESS sehr differenziert. Enthalten sind unter anderem Angaben zur Haushaltszusammensetzung, Wohngegend, Bildung und Beruf des/der Befragten sowie seines/ihrer Partners, seiner/ihrer Eltern, Gewerkschaftsmitgliedschaft, Haushaltseinkommen, Familienstand. Für den internationalen Vergleich des Bildungsniveaus stehen neben den national unterschiedlichen Items Variablen mit den ISCED-Levels, sowohl für den/die Befragte(n) selbst, ggf. Partner(in), Vater und Mutter zur Verfügung. Daneben wurde eine Berufskodierung gemäß der internationalen Standardklassifikation der Berufe ISCO-88 sowie eine Klassifikation von wirtschaftlichen Tätigkeiten nach Wirtschaftszweigen nach dem Standard NACE REV. 1 vorgenommen.

---

<sup>2</sup> Weitere Informationen unter <http://www.mzes.uni-mannheim.de/projekte/cid/>.

Für die Gewichtung der Daten stehen im internationalen Datensatz zwei Variablen zur Verfügung. Das sog. Designgewicht (DWEIGHT) korrigiert die Daten im Hinblick auf kleine Unterschiede bei den Auswahlwahrscheinlichkeiten für die Stichprobenziehung, das Population Size-Gewicht (PWEIGHT) passt die Daten bei der Analyse mehrerer Länder im Hinblick auf ihre proportionale Bevölkerungsgröße an.

## **5 Studienbeschreibung der deutschen Teilstudie des ESS**

**Studientitel: Gesellschaft und Demokratie in Europa.**

**Deutsche Teilstudie im Projekt "European Social Survey", 2002/2003**

### **Primärforscher**

National Coordinating Team Deutschland (NCT);

*Jan W. van Deth* (National Coordinator, Universität Mannheim),

*Oscar W. Gabriel* (Universität Stuttgart),

*Heiner Meulemann* (Universität zu Köln),

*Edeltraud Roller* (Johannes-Gutenberg-Universität Mainz),

*Katja Neller* (Geschäftsführerin, Universität Stuttgart).

### **Erhebungszeitraum**

Der Feldeinsatz erfolgte gestaffelt. Die erste Staffel mit einem Umfang von einem Drittel der Sample-Points ging - vornehmlich in Ostdeutschland - seit dem 23.11.2002 ins Feld. Die Hauptstaffel der Sample-Points ging Mitte Januar 2003 ins Feld. Feldende war der 16. Mai 2003.

### **Datenerhebung**

infas (Bonn)

### **Grundgesamtheit und Stichprobenziehung**

Untersuchungsgebiet: Bundesrepublik Deutschland.

Personenstichprobe: Personen ab 15 Jahren (ohne Altersobergrenze), die in einem privaten Haushalt in Deutschland leben, ungeachtet ihrer Nationalität, Staatsbürgerschaft, Sprache oder Rechtsstellung. Zweistufige, disproportional geschichtete Zufallsauswahl in Westdeutschland (inkl. Westberlin) und Ostdeutschland (inkl. Ostberlin). Oversampling für Ostdeutschland. In der ersten Auswahlstufe wurden Gemeinden in West- und Ostdeutschland ausgewählt. Für die Auswahl der Sample-Points wurden die Gemeinden nach der Kombination von Kreisen und 10 BIK-Gemeindegrößenklassen (Stand: Gemeindeverzeichnis 2001) geschichtet. Innerhalb dieser Schichtungszellen erfolgte die Auswahl entsprechend des Bedeutungsge-

wichtiges der Gemeinden, das sich aus der Zahl der 15-jährigen und älteren Bevölkerung je Gemeinde ergibt. In der zweiten Auswahlstufe wurden Personenadressen aus den Einwohnermeldekarteien zufällig gezogen.

### **Erhebungsverfahren**

Mündliche Befragung mit standardisiertem Frageprogramm (CAPI) inkl. Listenheft und schriftliche Zusatzbefragung (drop off mit sechs Splitversionen).

### **Fallzahl**

Im Datensatz der deutschen Teilstudie sind insgesamt 2919 Fälle enthalten. 1821 Interviews wurden in Westdeutschland geführt, 1098 in Ostdeutschland. Der Zusatzfragebogen wurde von insgesamt 2833 Befragten bearbeitet (Ausschöpfung: 97% der CAPI-Fälle).

### **Ausschöpfungsquote**

Die Ausschöpfungsquote beträgt bei Berechnung auf der Basis der insgesamt realisierten 2995 Interviews 59,6%, auf der Basis der 2919 gültigen Interviews 58,1% (Bruttostichprobe: 5796, stichprobenneutrale Ausfälle: 774, bereinigte Stichprobe: 5022, Verweigerungen durch Zielperson: 1271, Niemanden angetroffen: 422, sonstige Ausfallgründe: 334).

### **Incentives**

Als Incentive wurde ein mit dem Logo des European Social Survey versehener Taschenrechner, der jeweils am Ende eines vollständig realisierten Interviews übergeben wurde, eingesetzt. Als zusätzlicher Anreiz wurden am Ende der Feldphase unter den Teilnehmern 5 Städtereisen in europäische Hauptstädte verlost. Auf das Incentive und die Verlosung wurden die Zielpersonen schon im Anschreiben hingewiesen.

### **Interviewerschulungen**

Alle im ESS eingesetzten Interviewer wurden eigens für die Studie inhaltlich geschult. Der größte Teil der Interviewer wurde im Beisein von Vertretern der ESS-Projektgruppe in insgesamt fünf zentralen Workshops bei Infas in Bonn geschult. Zwei Schulungen fanden im November 2002 lokal vor Ort in Ostdeutschland statt. Vierzehn nacheingesetzte Interviewer wurden im Verlauf des Feldprozesses telefonisch geschult.

### **Interviewer**

Insgesamt waren 182 Interviewer im Einsatz. Aus feldstrategischen Gründen war es stellenweise nötig, Interviewer aus bestimmten Regionen mehrfach einzusetzen.

Auch waren Interviewer, die aufgrund ihrer Qualitäten für die Nachbearbeitungsphase besonders geeignet waren und entsprechend eingesetzt wurden, zum Teil in mehreren Gemeinden aktiv. Die durchschnittliche Anzahl von Interviews pro Interviewer liegt bei 15,6.

### **Interviewdauer**

Die Interviewdauer lag zwischen 40 und 271 Minuten. Die durchschnittliche Interviewdauer betrug 76 Minuten.

### **Themen**

Neben dem ESS-Standardfragebogen (siehe oben) enthielt die deutsche Teilstudie ein Modul mit deutschlandspezifischen Fragen. Dies waren u.a. Items zur Erhebung des Wohnortes vor 1990; zur Bewertung verschiedener Lebensbereiche in Ost- bzw. Westdeutschland vor der Wiedervereinigung; Fragen zur Einstellung zur ehemaligen DDR, zur Idee des Sozialismus und zur Leistungsbereitschaft der Bürger in den neuen Bundesländern; zur Verbundenheit mit der eigenen Stadt bzw. Gemeinde/Region/Bundesland/ehemaligen DDR bzw. BRD/mit Deutschland als Ganzem; politisches Vertrauen (Bundesregierung, öffentliche Verwaltung, Parteien) sowie Fragen zum Thema gesellschaftliche Gerechtigkeit.

## **6 Dokumentation und Zusatzinformationen**

### **Fragebögen, Codebücher und sonstige Studienmaterialien**

Sämtliche Fragebögen in den jeweiligen Landessprachen stehen zum Download auf der Projekthomepage (s.u.) bereit. Die Fragebögen für alle beteiligten Länder sowie sämtliche übrigen Studienmaterialien können auf der zentralen Projekthomepage heruntergeladen werden: [http://ess.nsd.uib.no/2003\\_questionnaires.jsp](http://ess.nsd.uib.no/2003_questionnaires.jsp). Die Materialien zur deutschen Teilstudie finden Sie darüber hinaus auch auf der deutschen Projekthomepage unter <http://www.europeansocialsurvey.de>. Einen Kurzüberblick über die verfügbaren Länder und ihre Feldzeiten sowie Abweichungen vom vorgegebenen Fragebogen findet sich unter: [http://ess.nsd.uib.no/2003\\_Fworksummary.jsp](http://ess.nsd.uib.no/2003_Fworksummary.jsp). Eine ausführliche Dokumentation der Studie beinhaltet die Seite [http://ess.nsd.uib.no/2003\\_documentation.jsp](http://ess.nsd.uib.no/2003_documentation.jsp). Darüber hinaus wird auf der Projekthomepage des NSD die Möglichkeit eines Online-Codebooks angeboten. Für ausgewählte Länder bzw. Variablen können Grundauszählungen durchgeführt werden: <http://ess.nsd.uib.no>. Außerdem stehen für methodische Analysen für die Teilnehmerländer des ESS codierte Informationen aus den Interviewerkontaktprotokollen inkl. eines Quartiersfragebogens zur Verfügung, die ebenfalls auf der Projekthomepage des NSD heruntergeladen werden können.

### **Verfügbare Makro- und Kontextdaten**

Im Rahmen des ESS wurde für die Teilnehmerländer eine Sammlung von ausgewählten Aggregatdaten erstellt. Diese umfassen neben diversen Bevölkerungsstatistiken (z.B. Verteilung auf regionale Einheiten, Geschlecht, Alter, Bildung) z.B. Angaben zum Bruttosozialprodukt, zur Preisentwicklung oder zur durchschnittlichen Lebenserwartung. Diese Informationen sind über die Projekthomepage zugänglich. Download unter: [http://ess.nsd.uib.no/2003\\_other.jsp](http://ess.nsd.uib.no/2003_other.jsp). Für die deutsche Teilstudie besteht darüber hinaus die Möglichkeit, über eine Variable mit Bundestagswahlkreisziiffern die Individualdaten mit Aggregatdaten zu verknüpfen.

### **Ereignisdatensammlung für die ESS-Teilnehmerländer**

Im Rahmen des ESS wurden von den Teilnehmerländern Ereignisdatensammlungen angefertigt, die Geschehnisse während der Feldzeit (sowie einen Monat vor Beginn der Feldzeit) dokumentieren sollen, denen ein potentieller Einfluss auf die erhobenen Daten zugeschrieben werden kann. Die Ereignisdatensammlungen für *alle* beteiligten Länder finden Sie unter:

[http://www.scp.nl/users/stoop/ess\\_events/events\\_overview.htm](http://www.scp.nl/users/stoop/ess_events/events_overview.htm).

Für die deutsche Teilstudie des ESS wurde von August 2002 bis Mai 2003 eine Ereignisdatensammlung nach den Vorgaben des CCT erstellt. Hierfür wurden die Titelseiten der "Frankfurter Allgemeinen Zeitung" und die Hauptnachrichtensendungen von ARD und ZDF herangezogen. Dabei wurden aufgrund bestimmter Kriterien ausgewählte Ereignisse erfasst sowie deren Dauer bzw. das Ausmaß ihrer Medienpräsenz festgehalten. Auswahlkriterien waren:

Bedeutende nationale (bzw. internationale) Ereignisse, die deutschlandweit und über Deutschland hinaus über einen längeren Zeitraum Beachtung in den Medien bzw. der Öffentlichkeit finden (z.B. Konflikt zwischen Israel und Palästina, Selbstmordattentäter in Erfurt; Terroraktionen der ETA u.Ä.). Bedeutende nationale Ereignisse, die einige Zeit (an mind. 3 aufeinanderfolgenden Tagen) die Titelseiten/Schlagzeilen beherrschen und zu einer breiten öffentlichen Diskussion und/oder zu einer Zunahme der Medienberichterstattung über diese Themen führen (z.B. Streiks, Diskussion über Arbeitsmarktpolitik usw.), Wahlen, wichtigste Themen des Wahlkampfes, Veränderungen der politischen Landschaft, z.B. im Bereich der Parteien (z.B. Schill-Partei usw.)

### **Weitere Projektinformationen unter:**

Allgemeine Informationen: <http://www.europeansocialsurvey.org>.

Datendownload, Online-Codebook, Dokumentation: <http://ess.nsd.uib.no>.

Deutsche Teilstudie: [www.europeansocialsurvey.de](http://www.europeansocialsurvey.de).

## **The Establishment of the Social Science Data Archive in Bulgaria (SSDA)**

by **Yantsislav Yanakiev<sup>1</sup>** and **Brigitte Hausstein**

After the democratic changes in Bulgaria in 1989 the number of empirical social surveys has rapidly increased. In addition to the existing public research institutes many private sociological and marketing agencies have been established. Both actors in public and private sectors have produced a large number of quantitative data sets. At the same time, the data from social surveys has not been analysed and used thoroughly by the research community as well as decision makers. There are many reasons for this situation. Among the most important are difficult and unregulated data access conditions, lack of willingness to exchange data and information among the research institutions, significant data losses, etc.

On the other hand there exists a growing need for secondary analysis of data from social surveys and opinion polls for educational purposes as well as the development of theoretical and methodological research.

The establishment of the Social Science Data Archive (SSDA) in Bulgaria will contribute to the

- accessibility of survey data,
- development of secondary data analysis
- data exchange with researchers abroad.

Potential users of the data are the academic and educational institutions in Bulgaria and abroad, various NGOs, think tanks, etc.

### **Objectives of the data archive project are:**

- To increase the researchers' awareness and understanding of the importance to set up a Social Science Data Archive in Bulgaria.
- To make available data from international comparative surveys for secondary data analysis in Bulgaria and thus, to promote a data sharing culture.

---

1 *Yantsislav Yanakiev* is associate professor and vice-president of the Research Centre on Regional and Global Development (REGLO) in Sofia, Bulgaria.

- To assess current problems and possibilities for consolidation and integration of sociological data as well as to identify prospects for cooperation among Bulgarian sociological research institutions.
- To define common criteria for data storage in the archive and
- To promote international co-operation and exchange of experience in social science data archiving, and particularly in standards for data documentation, training of personnel and motivation of researchers to deposit data in the archive.

**The main activities are:**

- Desk research on the experience of already existing social science data archives in Central and East European Countries
- Needs assessment: Expert survey among main actors in the field of social empirical research in Bulgaria as well as main expected users
- Specialized training in social science data archiving
- Establishment of the Information and Data Dissemination Centre and installation of two work stations for secondary data analysis
- Establishment of a database containing data from international surveys in which Bulgaria has taken part and making it available to Bulgarian researchers
- Indexing existing data files and the thematic catalogue of surveys carried out in Bulgaria during the last decade
- Translation of questionnaires, data descriptions, etc. into English, with priority to recent surveys carried out in the last two-three years
- Design and creation of a web page
- Dissemination of information materials about SSDA
- Organisation of an international workshop to discuss the prospects for consolidation and integration of sociological data in Bulgaria and the further development of SSDA. It is envisaged that the workshop will be integrated into the International Summer School "Best Practices in Archiving and Use of Data from Comparative Studies", organized by UNESCO and International Social Science Council.

The project is sponsored by the UNESCO Participation Programme and is carried out at the Research Centre Regional and Global Development (REGLO) during the period October 2002 - October 2004.

For more information, please contact Yantsislav Yanakiev (project director)  
reglo-office@techno-link.com

## Buchhinweise

**Buttler, Günter; Fickel, Norman**  
**Statistik mit Stichproben**

**Paperback, ISBN: 3-499-55653-7**  
**Reinbek bei Hamburg:**  
**rowohlt 2002, 352 Seiten,**  
**12,90 Euro**

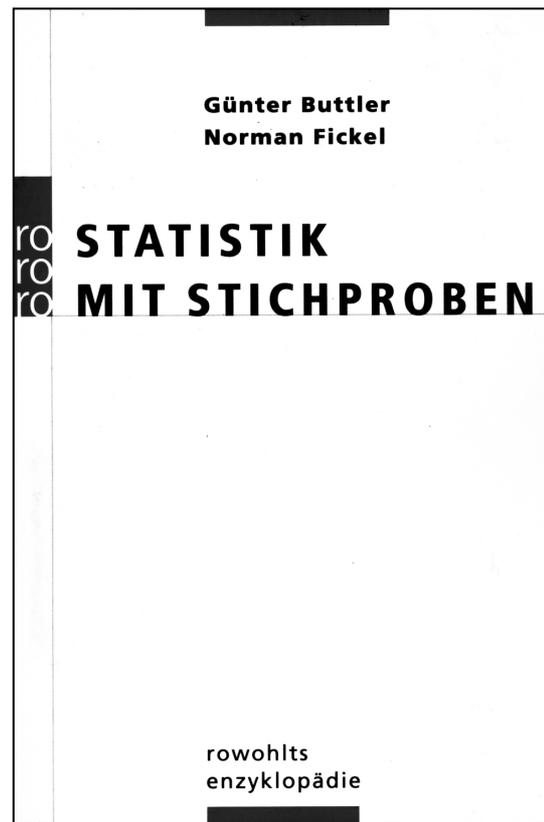
Das Buch beschreibt eine Reihe von Stichprobenverfahren und zeigt, wie unbekannte Maßzahlen der Gesamtheit geschätzt und wie Hypothesen anhand von Stichproben überprüft werden können.

Aus dem Inhalt:

- Möglichkeiten der Datengewinnung
- Arten von Stichproben
- Schätzverfahren
- Statistische Testverfahren
- Techniken der Zufallsauswahl
- Genauigkeit von statistischen Daten
- Datenbeschaffung im Internet
- Beispiele für konkrete Stichproben

Das Buch vermittelt ein profundes Verständnis dafür, wie Stichproben durchgeführt werden und wie aus den so ermittelten Daten Informationen gewonnen werden, die gehobenen Qualitätsansprüchen genügen. Es ist verständlich geschrieben und erläutert die Grundbegriffe sehr anschaulich. Am Ende jedes Kapitels stehen Fragen oder Rechenaufgaben, die der Überprüfung des zuvor Gelesenen dienen. Die Auflösungen finden sich im Anhang.

Auswahlkonzepte und Besonderheiten von Datenbasen – Einkommens- und Verbraucherstichprobe, Mikrozensus, ALLBUS und DeutschlandTREND –, die für Sozialforscher von speziellem Interesse sind, werden dargestellt. Die Beschreibung von Techniken der Auswahl, wie z.B. Zufallsauswahl aus Registern, geschichtete Zufallsauswahl oder die ADM-Stichprobe ist durch typische Beispiele angereichert. Ein Buch, das nicht nur Studierenden der Sozialforschung einen Einstieg in das Thema bietet.



**Zöfel, Peter**  
**SPSS-Syntax: Die ideale Ergänzung**  
**für effiziente Datenanalyse**

**ISBN: 3-8273-7052-3**  
**München: Pearson Studium 2002,**  
**210 Seiten, CD, 24,95 Euro**

*A help for old friends* nannte sich in einer der frühen Ausgaben der SPSS-Handbücher ein Kapitel, das den Programmnutzern der ersten Stunde die Neuerungen der erweiterten Version von SPSS kompakt vermittelte. An dieses Kapitel fühlt man sich erinnert, wenn **Peter Zöfel** ein SPSS-Buch für die Freunde der Syntax-Programmierung vorlegt. ‚Alte Hasen‘ haben immer darauf gesetzt, dass Syntax-Programmierung mit SPSS von großem Vorteil ist. Dokumentiert der Syntaxfile doch geradezu beiläufig alle Analyseschritte und vor allem auch die Modifikation der Daten z.B. durch recodes und computes. Dieser File ist ein unverzichtbarer Beleg, der die in der Analysephase häufig wechselnden Veränderungsschritte in der letztendlich verwendeten Form dokumentiert. Eigentlich *die* Garantie für eine wissenschaftlich saubere Analysearbeit. Und: Alle Auswertungen können jederzeit bequem wiederholt werden.

**Aus dem Inhalt:** Prinzipielles zu Syntax, Dateneingabe über eine Textdatei, Datenaufbereitung, Berechnen von neuen Variablen, Auswahl von Fällen, Auflisten von Variablen, Arbeiten mit Systemdateien, Sortieren und Gewichten, Makros, Programmabläufe, Prozeduren, Matrizenrechnung.

Das Buch beschreibt die vielfältigen Einsatzmöglichkeiten der SPSS-Syntax anhand von zahlreichen Beispielen. Die zugehörigen Daten und Programme sind auf einer CD beigegeben. Auch das Schreiben und Testen von Programmen, ohne dass bereits Daten vorliegen, wird aufgezeigt.

Das Buch verzichtet auf die Beschreibung statistischer Methoden. Es konzentriert sich auf eine kompakte Darstellung der Befehle für eine effiziente Programmierung und wendet sich somit an die alten Freunde von SPSS, die für eine schnelle Programmierung ein Brevier suchen, das auf aktuellstem Stand ist.

