SSOAR

Open Access Repository
www.ssoar.info

Qualitätsmanagement und Qualitätssicherung: 7. Wissenschaftliche Tagung

König, Christian (Ed.); Stahl, Matthias (Ed.); Wiegand, Erich (Ed.)

Veröffentlichungsversion / Published Version
Konferenzband / conference proceedings

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
König, C., Stahl, M., & Wiegand, E. (Hrsg.). (2007). *Qualitätsmanagement und Qualitätssicherung: 7. Wissenschaftliche Tagung* (GESIS-Tagungsberichte, 1). Bonn: GESIS-IZ Sozialwissenschaften. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-261115

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Qualitätsmanagement und Qualitätssicherung

7. Wissenschaftliche Tagung

Christian König, Matthias Stahl, Erich Wiegand (Hrsg.)

im Auftrag
des Statistischen Bundesamtes, Wiesbaden,
der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI)
und des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.

GESIS-Tagungsberichte, Band 1

gesis

Qualitätsmanagement und Qualitätssicherung

GESIS-Tagungsberichte

Herausgegeben von GESIS, Bonn
Band 1

GESIS ist ein Institut der Leibniz-Gemeinschaft.

Qualitätsmanagement und Qualitätssicherung

7. Wissenschaftliche Tagung

Christian König, Matthias Stahl, Erich Wiegand (Hrsg.)

im Auftrag

des Statistischen Bundesamtes, Wiesbaden,
des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.
und der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI)

GESIS-Tagungsberichte, Band 1

GESIS, Bonn 2007

Bibliographische Information Die Deutsche Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliothek; detaillierte bibliographische Daten sind im Internet über www.ddb.de abrufbar.

ISBN-13 978-3-86819-000-7
ISSN 1865-7559

Herausgeber, Druck und Vertrieb:
GESIS
Lennéstraße 30, 53113 Bonn
Tel.: 02 28 - 22 81 - 0
Printed in Germany

© 2007 GESIS, Bonn. Alle Rechte vorbehalten. Insbesondere ist die Überführung in maschinenlesbare Form sowie das Speichern in Informationssystemen, auch auszugsweise, nur mit schriftlicher Einwilligung gestattet.

Inhalt

Begrüßung Präsident Walter Radermacher 7

Manfred Ehling
Qualitätsmanagement und Qualitätssicherung: Einführung 11

Holger Mühlbauer
Normung 33

Erich Wiegand
Standards und Normen in der Markt- und Sozialforschung 43

Olaf Hofmann
Qualitätsstandards bei Online-Access Panels. 51

Silke I. Keil, Katja Neller
Ex-ante-Qualitätskontrolle im European Social Survey. 63

Michael Pusler
Qualitätsanforderungen der Auftraggeber an die Umfrageforschung aus Sicht der Medien- bzw. Verlagsmarktforschung 77

Ulrike Timm, Andrea Kron
Bewertung von Datenqualität in der amtlichen Statistik 89

Martina Hahn
Der Verhaltenskodex des Europäischen Statistischen Systems 99

Christina Buchwald, Katja Lukanow
Qualitätskontrolle im Telefoninterview 111

Hartmut Scheffler
Aus- und Weiterbildung als Instrument der Qualitätssicherung 155

Hariet Köstner
Anforderungen und Nutzen der Zertifizierung in der Markt- und Sozialforschung 163

Referentinnen und Referenten 177

Teilnehmerliste 183

Begrüßung Präsident Walter Radermacher

Die wissenschaftliche Tagung über „Qualitätsmanagement und Qualitätssicherung“, zu der ich Sie heute hier in Wiesbaden herzlich begrüße, ist die inzwischen siebte gemeinsame wissenschaftliche Veranstaltung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes.

Ich möchte nicht mit einem Rückblick auf die Themen unserer gemeinsamen Veranstaltungsreihe seit dem Jahre 1995 beginnen, sondern Ihre Aufmerksamkeit auf ein Buch richten, das für mich wie wenig andere Bücher mit dem Qualitätsbegriff verbunden ist, Robert M. Pirsig’s 1974 erschienenes Werk „Zen and the Art of Motorcycle Maintenance“[1]. Pirsig beschreibt darin eine Motorradreise durch die USA, in welcher der Qualität, dargestellt anhand der Wartung und Pflege des Motorrades, eine zentrale Bedeutung zukommt. Die Geschichte spielt auf verschiedenen Ebenen und trägt teilweise autobiografische Züge, in denen Pirsig auch seine persönliche Lebensgeschichte und seine Erfahrungen mit psychiatrischen Therapieeinrichtungen verarbeitet. Das Motorrad und seine Wartung sind darin Archetypen für Technologie und Pirsig’s Qualitätsdiskussion beschreibt die alte Debatte zwischen rational logischer Sicht auf der einen und einer subjektiven Wahrnehmung auf der anderen Seite. Qualität als metaphysischer Begriff steht in der Philosophie von Pirsig über dem Dualismus von Subjektivität und Objektivität. Er beschreibt Qualitätsbewusstsein als etwas nie vollständig Erfassbares, das er dem üblichen Verständnis von Bewusstsein als intellektuellem Erfassen der Realität voranstellt. Der britische Observer beschreibt in seiner Ausgabe vom 19. November 2006 Pirsig’s „Zen and the Art of Motorcycle Maintenance“ als das am besten verkaufte philosophische Buch aller Zeiten; ich kann es jedem am Qualitätsbegriff Interessierten nur empfehlen. Lassen Sie mich jetzt aber doch ein klein wenig in die Geschichte unserer gemeinsamen Tagungen blicken. Wenn wir auf die vergangenen sechs gemeinsamen Tagungen zurückblicken, so können wir zweierlei erkennen:

1. dass der zeitliche Abstand zwischen den Tagungsthemen, die neuen Techniken gewidmet sind, kürzer geworden ist und

1 Deutsche Ausgabe: Robert M. Pirsig, Zen und die Kunst ein Motorrad zu warten, Fischer Taschenbuchverlag 1978

2. dass Qualitätsaspekte quasi die Klammer um alle Themen gebildet haben und in diesem Jahr mit unserem Thema „Qualitätsmanagement und Qualitätssicherung“ erstmals das explizite Thema unserer wissenschaftlichen Tagung darstellen.

Beides hängt natürlich eng miteinander zusammen: viele Aspekte des Qualitätsmanagements und der Qualitätssicherung wären ohne die neuen technischen Entwicklungen gar nicht vorstellbar gewesen.

Dass Qualitätsaspekte quasi die Klammer um alle unsere bisherigen Tagungen gebildet haben, verwundert auch nicht, denn die Herausforderungen, vor der die kommerziellen Marktforschungsinstitute, die sozialwissenschaftliche Forschung und die amtliche Statistik stehen, sind – wie wir alle inzwischen wissen – sehr wohl vergleichbar. Während sich die kommerziellen Institute im Wettbewerb am Markt behaupten müssen, Gewinne erwirtschaften, Kosten senken und Marktanteile sichern müssen, zwingen die stets enger werdenden Spielräume in den öffentlichen Haushalten die akademische Forschung wie auch die amtliche Statistik dazu, mit geringeren Ressourcen auszukommen und dennoch die eigenen Produkte ständig insbesondere auch unter den Qualitätsaspekten Aktualität und Genauigkeit zu verbessern. Alle Beteiligten können diesen Herausforderungen nur durch ein effizientes Qualitätsmanagement und eine hervorragende Qualitätssicherung unter dem Einsatz modernster Technologien gerecht werden.

Bevor wir mit den Fachvorträgen beginnen, ist es mir ein echtes Anliegen, allen an dieser Veranstaltung Beteiligten herzlich für die gute Zusammenarbeit zu danken. Die Arbeitsteilung dieser wissenschaftlichen Veranstaltungen, bei der auch in diesem Jahr die ASI wieder für die Produktion des Tagungsbandes verantwortlich zeichnen wird, der ADM das Catering des abendlichen „get-together“ hier im Crowne Plaza übernimmt und das Bundesamt für die Organisation der Tagung verantwortlich ist, hat sich hervorragend bewährt. Auch die inhaltliche Vorbereitung läuft dank der seit vielen Jahren eingespielten vertrauensvollen Zusammenarbeit in einem kleinen Team aus Vertretern von ASI, ADM und Bundesamt reibungslos. Die Tatsache, dass Herr Scheffler als Vorstandsvorsitzender des ADM und Geschäftsführer der TNS Infratest Holding, nach der Übernahme der Moderation unserer wissenschaftlichen Tagung vor zwei Jahren auch diesmal als Referent unsere Veranstaltung aktiv unterstützt, unterstreicht den Stellenwert, den diese gemeinsame Veranstaltungsreihe auch für die kommerzielle Marktforschung besitzt und dies freut uns sehr.

Es sind aber bei weitem nicht nur diese gemeinsamen wissenschaftlichen Tagungen, in denen sich die gute Zusammenarbeit zeigt. Seit über 20 Jahren treffen sich jedes Jahr Ende November Vertreter von ASI, ADM und dem Statistischen Bundesamt zu einem äußerst konstruktiven Gedankenaustausch hier in Wiesbaden. Unsere wissenschaftlichen Tagungen sind ein gemeinsames Produkt unter vielen, die ihren Ausgangspunkt in den gemeinsamen November-Sitzungen ge-

nommen haben. Ohne andere Projekte zurücksetzen zu wollen, möchte ich exemplarisch die Verabschiedung gemeinsamer demographischer und regionaler Standards und auch die Integration der BIK-Regional-Informationen in das Einzelmaterial des Mikrozensus erwähnen. Neben allen konkreten Kooperationsprojekten hat die Vielzahl der gemeinsamen Aktivitäten und der gegenseitige Austausch in den vielen Jahren der Kooperationen zu einem vertrauensvollen Verhältnis nicht nur zwischen den beteiligten Institutionen, sondern auch zwischen vielen Marktforschern, Sozialwissenschaftlern und Statistikern auf der Arbeitsebene geführt. Ein solcher Umgang miteinander ist ein hohes Gut und gerade heute leider nicht mehr überall selbstverständlich. An dieser guten Zusammenarbeit haben Sie alle, meine Damen und Herren, Ihren Anteil und dafür danke ich Ihnen. Diese vertrauensvolle Zusammenarbeit ist mir sehr wichtig und ich wünsche mir, dass wir noch eine Vielzahl weiterer gemeinsamer Projekte in diesem hervorragenden gemeinsamen Arbeitsklima realisieren können.

Meine Damen und Herren, ich möchte abschließend schon jetzt allen Referentinnen und Referenten für ihre Beiträge danken und Herrn Dr. Ehling für die Übernahme der Moderation. Ich wünsche Ihnen auch in diesem Jahr einen interessanten Tagungsverlauf mit spannenden Diskussionen und anregenden Gesprächen am Rande der Veranstaltung.

Qualitätsmanagement und Qualitätssicherung: Einführung

Manfred Ehling

1 Einleitung

Die Diskussion um die Qualität statistischer Daten hat in den vergangenen Jahren erhebliche Veränderungen erfahren. Stand in der Vergangenheit das Dreieck aus Genauigkeit – Vollständigkeit – Aktualität (vgl. Chlumsky/Wiegert 1993) im Mittelpunkt der Debatte, wird Datenqualität jetzt unter der Triade institutioneller Rahmen – statistische Prozesse – statistische Produkte diskutiert (vgl. Kopsch/Köhler/Körner 2006, S. 793ff). Institutionelle und organisatorische Faktoren wirken sich danach genauso wie die Erhebung, Verarbeitung und Analyse auf die erreichte Datenqualität aus. Ein übergreifendes Qualitätsmanagement der „erhebenden Stelle" ist Voraussetzung für eine effiziente Erhebungsorganisation, die in allen Phasen der Statistikerstellung prozessbezogene Qualitätsstandards erfüllt. Schließlich definiert sich die Qualität statistischer Produkte über unterschiedliche Komponenten, wobei in der deutschen amtlichen Statistik einheitlich die Qualitätskriterien Relevanz, Genauigkeit, Aktualität und Pünktlichkeit, Verfügbarkeit und Transparenz, Vergleichbarkeit sowie Kohärenz angewendet werden, die im Europäischen Statistischen System (ESS) entwickelt wurden.

Um in die Thematik „Qualitätsmanagement und Qualitätssicherung" aus unterschiedlicher Perspektive einzuführen, sollen die Qualitätsansprüche der Markt- und Sozialforschung und der amtlichen Statistik vergleichend dargestellt werden. Dazu werden Qualitätsstandards, Qualitätskriterien und Maßnahmen der Qualitätssicherung aus dem Umfeld der an der Tagung beteiligten Organisationen gegenübergestellt. Voraussetzung und Rahmen für eine zielgerichtete Qualitätssicherung ist ein umfassendes Qualitätsmanagement, das sich nicht nur auf den Erstellungsprozess von Daten bezieht, sondern auch die Unterstützungsprozesse und die Führung mit betrachtet. Im nächsten Kapitel wird dazu der Qualitätsmanagementansatz des Statistischen Bundesamtes kurz dargestellt.

2 Qualitätsmanagement im Statistischen Bundesamt

Umfassendes Qualitätsmanagement als ganzheitlicher Ansatz zur Leitung eines Unternehmens oder einer Behörde wird definiert als „eine auf die Mitwirkung al-

ler ihrer Mitglieder gestützte Führungsmethode einer Organisation, die Qualität in den Mittelpunkt stellt und durch Zufriedenstellung der Kunden auf langfristigen Geschäftserfolg sowie auf Nutzen für die Mitglieder der Organisation und für die Gesellschaft zielt" (DIN ISO 8402). Dieser Managementansatz erfordert eine von allen Beschäftigten gelebte Qualitätsphilosophie. Der Qualitätsbegriff wird dabei umfassend verstanden: nicht nur die Produkt- und die Servicequalität, sondern auch die Qualität im Hinblick auf die Belange der Mitarbeiter, die Kundenzufriedenheit, den langfristigen Geschäftserfolg sowie den gesellschaftlichen Nutzen werden in den Mittelpunkt gerückt. Um diese umfassenden Anforderungen zu kennzeichnen, wird das Wort „Qualität“ oft durch den englischen Begriff der „Excellence“ ersetzt.

Das Statistische Bundesamt hat 1999 mit der Qualitätsoffensive Statistik begonnen, ein umfassendes Qualitätsmanagement im Sinne des Total Quality Management (TQM) einzuführen. Ausgangspunkt war die Notwendigkeit, bei stetig steigenden Bedürfnissen der Kunden und Nutzer und gleichzeitiger Abnahme des Budgets ein weiterhin hohes Qualitätsniveau durch den Einsatz neuer Informations- und Kommunikationstechniken zu erreichen. Die Ausrichtung auf ein modernes Amt orientiert sich damit an den Leitgedanken eines umfassenden Qualitätsmanagements und basiert auf wenigen und klaren Prinzipien, die mit den Begriffen Innovation, Veränderung, Partnerschaft und kontinuierliche Verbesserung umschrieben werden können.

Übersicht 1: Die qualitätsoffensive Statistik: Zieldimensionen und -kategorien

Das umfassende Qualitätsmanagement im Statistischen Bundesamt baut auf den folgenden Leitsätzen auf (vgl. Übersicht 1):

- Ergebnisorientierte Führung: Die Vorgesetzten einigen sich mit den Mitarbeiterinnen und Mitarbeitern zu Beginn einer Arbeit auf deren Endprodukt; detaillierte Anweisungen und Kontrollen entfallen.
- Verantwortung, Kreativität und Eigeninitiative der Mitarbeiterinnen und Mitarbeiter werden gestärkt, z. B. durch flexible Arbeitszeit, Ideenbörse und Budgetierung.

Die Qualitätsziele des Statistischen Bundesamtes umfassen:

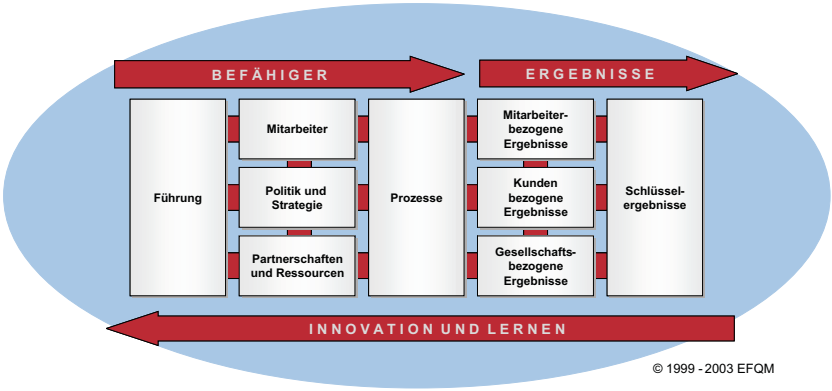
- Führungsverpflichtung und Mitarbeiterbeteiligung,
- Kundenzufriedenheit und Produktverbesserung,
- kontinuierliche Prozessoptimierung und
- verbesserte Wirtschaftlichkeit.

Damit beinhaltet Qualitätsmanagement im Statistischen Bundesamt die Bausteine:

- Organisationsqualität,
- Prozessqualität,
- Produktqualität und
- Nachfrageorientierung.

Die Umsetzung dieses umfassenden Qualitätsmanagements (TQM) erfolgt auf der Basis des EFQM-Modells, das 1988 von der European Foundation for Quality Management (EFQM) entwickelt wurde. Nach Schätzungen arbeiten heute etwa 10.000 Unternehmen und Verwaltungen danach. Das EFQM-Modell ist ein Werkzeug, das Hilfestellung gibt für den Aufbau und die kontinuierliche Weiterentwicklung eines umfassenden Managementsystems und das aufzeigt, wo man sich auf dem Weg zu exzellentem Management / Spitzenleistungen befindet. Es hilft, eigene Stärken, Schwächen und Verbesserungspotentiale zu erkennen und die eigene Strategie darauf auszurichten. Die Entscheidung für das EFQM-Modell als konzeptionellen Rahmen der Qualitätsoffensive in der Bundesstatistik begründet einen auf Dauer angelegten Entwicklungsprozess, bei dem auf allen relevanten Feldern systematisch (d. h. geplant und mit Messungen nachgeprüft) Verbesserungen umgesetzt werden.

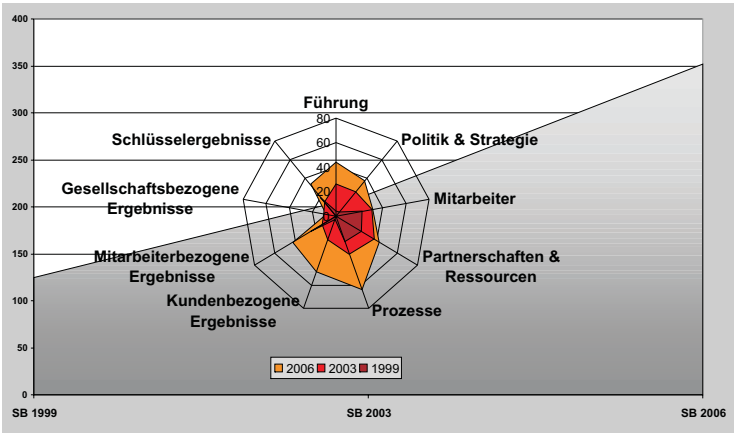
Übersicht 2: Das EFQM-Modell für Excellence



Exzellente *Ergebnisse* im Hinblick auf *Leistung, Kunden, Mitarbeiter und Gesellschaft* werden durch eine *Führung* erzielt, die *Politik und Strategie* mit Hilfe der *Mitarbeiter*, *Partnerschaften* und *Ressourcen* sowie der *Prozesse* umsetzt.

EFQM-Selbstbewertungen sind ein Diagnoseinstrument des Managements zur systematischen Bewertung des Statistischen Bundesamtes. Das EFQM-Konzept verlangt eine regelmäßige ganzheitliche Betrachtung der Organisation zu sämtlichen Kriterien des Modells. Neben der umfassenden und systematischen Überprüfung der Ergebnisse wie auch der Mittel und Wege können so Stärken und Verbesserungspotentiale ermittelt und Schritte auf dem Weg zu einer kontinuierlichen Verbesserung geplant, umgesetzt und evaluiert werden.

Übersicht 3: Ergebnisse der EFQM-Selbstbewertungen



EFQM-Selbstbewertungen wurden im Statistischen Bundesamt in den Jahren 1999, 2003 und 2006 durchgeführt. Übersicht 3 zeigt zum einen den Erfüllungsgrad der neun Kriterien des EFQM-Modells, in den Jahren, in denen Selbstbewertungen durchgeführt wurden, in der Form eines Spinnennetzes. Auf der Ordinate ist die erzielte Gesamtpunktzahl in den Bewertungsjahren (SB = Selbstbewertung) abzulesen. Die Erfüllungsgrade in den neun Kriterien werden mit unterschiedlichen Faktoren gewichtet, so dass im EFQM-Modell max. 1.000 Punkte erreicht werden können, was bisher keinem Unternehmen bzw. keiner öffentlichen Verwaltung gelungen ist. Teilnahmevoraussetzung für den europäischen EFQM-Qualitätspreis und den nationalen Ludwig-Erhard-Preis ist eine Bewertung mit mindestens 500 Punkten. In der letzten Selbstbewertung erzielte das Statistische Bundesamt rund 350 Punkte.

Die Selbstbewertung 1999 verfolgte das Ziel, einen ersten Eindruck über den Stand des Qualitätsmanagements zu liefern. Das EFQM-Modell diente dabei als Orientierungsrahmen. In 2003 und 2006 wurden mit dem strukturierten Ansatz des EFQM-Modells alle Aspekte des Qualitätssystems betrachtet, um daraus konkrete Verbesserungsmaßnahmen abzuleiten. In der Selbstbewertung 1999 wurden von den insgesamt 32 Bewertungskriterien des EFQM-Modells 14 Kriterien betrachtet, während in 2003 und 2006 alle Kriterien und Teilkriterien des EFQM-Modells bewertet wurden. Aus diesen Gründen sind die Ergebnisse von 1999 einerseits und 2003 und 2006 andererseits nicht miteinander vergleichbar. Gleichwohl lässt sich über den Ergebnisvergleich die positive Entwicklungsrichtung des Statistischen Bundesamtes dokumentieren (siehe Übersicht 3). Ohne ins Detail zu gehen, ist aus der Gegenüberstellung der Selbstbewertungen 2006, 2003 sowie 1999 herauslesbar: Das Statistische Bundesamt hat seine Qualität gesteigert.

Der Vergleich mit den Ergebnissen der früheren Selbstbewertungen zeigt deutliche Verbesserungen bei den Kriterien Führung, Prozesse, Kundenbezogene Ergebnisse, Mitarbeiterbezogene Ergebnisse und bei den Schlüsselergebnissen. Allerdings haben die Selbstbewertungen auch gezeigt, dass in einigen Bereichen noch Verbesserungspotential für das Statistische Bundesamt besteht. Für die Zukunft gilt es, die vorhandenen Stärken zu erhalten und Verbesserungspotentiale gezielt in Angriff zu nehmen. Dabei wird sich das Statistische Bundesamt zunächst auf die Felder konzentrieren, in denen in den nächsten Jahren die größten Verbesserungen erreichbar erscheinen. Dies sind insbesondere die Felder Partnerschaften & Ressourcen, Prozesse, Kundenbezogene Ergebnisse und Schlüsselergebnisse.

Das Statistische Bundesamt hat seit 1999 eine Reihe von Maßnahmen im Rahmen der Qualitätsoffensive Statistik durchgeführt. Dazu zählt u.a. die Festlegung von strategischen Zielen, die in den verschiedenen fit-Programmen (fit 2005, fit 2008, in Kürze fit 2012; fit = fachkompetent, innovativ, topaktuell) fixiert sind.

Weitere Schritte sind die Einführung und Weiterentwicklung von Strategie- und Programmplanungen, Prozessanalysen, Controlling, Budgetierung, Corporate Design, Mitarbeiterbefragungen, Kundenbefragungen sowie Projekte zur Verbesserung der Datenqualität.

Ohne auf die Punkte des Verbesserungskatalogs im Einzelnen einzugehen, soll im Folgenden ein systematischer Überblick über ausgewählte Maßnahmen zu den EFQM-Kritierien gegeben werden.

Übersicht 4: Ausgewählte Maßnahmen zu einzelnen EFQM-Kritierien

1. Führung
- Leitbild

2. Politik und Strategie
- fit 2008, 2012
- Programmplanung
- Jahresarbeitsplanung
- Masterplan zur Reform der amtlichen Statistik

3. Mitarbeiter
- Personalentwicklung/Fortbildung
- Personalführungsgespräche
- Ideenbörse

4. Partnerschaften und Ressourcen
- Kosten- und Leistungsrechnung
- Produkthaushalt
- IT-Planung und -Koordinierung
- Informations- und Wissensmanagement

5. Prozesse
- Standardisierung des Produktionsprozesses der Statistikerstellung
- Prozessmodell
- Leitfaden für Projektmanagement
- Controlling

6. Kundenergebnisse
- Kundenbefragungen
- Kundenmanagementsystem

7. Mitarbeiterergebnisse
- Vorgesetztenrückmeldung

- Mitarbeiterbefragungen

9. Schlüsselergebnisse

- Veröffentlichung von Qualitätsberichten
- Strategie- und Programmplan
- EFQM-Selbstbewertungen

3 Qualitätskriterien der Umfrageforschung

Zu Qualitätskriterien der Umfrageforschung hat die Deutsche Forschungsgemeinschaft (DFG) eine Denkschrift veröffentlicht (Kaase 1999). Die DFG dient gemäß ihrer Satzung der Wissenschaft. Sie gewährt finanzielle Unterstützung von Forschungsaufgaben, fördert den wissenschaftlichen Nachwuchs sowie die Zusammenarbeit unter Forschern mit Partnern aus dem In- und Ausland. Eine für die Erfüllung dieser Aufgaben immer wichtiger werdende Grundlage ist die Qualitätssicherung der Forschung. „Lege artis“ zu arbeiten, definiert eine 1998 erschienene Empfehlung der DFG als eine Grundvoraussetzung für gute wissenschaftliche Praxis (Deutsche Forschungsgemeinschaft 1998).

Ein wichtiges Instrument sozialwissenschaftlicher Forschung sind Umfragen, die die Grundlage für vielfältige politische und gesellschaftliche Entscheidungen bilden. Die DFG hat in ihrer von Kaase herausgegebenen Denkschrift zu Qualitätskriterien der Umfrageforschung dazu u. a. Qualitätsmaßstäbe formuliert. Diese Veröffentlichung behandelt Qualität in unterschiedlichen Dimensionen von der „handwerklichen" Korrektheit der methodischen Vorgehensweise über forschungsethische Grundsätze des Umgangs mit Befragten und ihren Aussagen bis hin zu Fragen der Dokumentation und Offenlegung der aus Umfragen hervorgegangenen Datensätze (vgl. Kaase 1999, S.11ff). Diese Qualitätskriterien bilden einen Maßstab für die akademische Sozialforschung. Da die Umfrageforschung in Deutschland seit längerem durch eine konstruktive Kooperation und einen kontinuierlichen Dialog zwischen akademisch und privatwirtschaftlich verorteten Forschern sowie der amtlichen Statistik gekennzeichnet ist, sind alle Gruppen sowohl in die Erstellung einbezogen worden als auch mögliche Zielgruppe der Denkschrift.

Die Publikation der DFG beschreibt den aktuellen Stand der Umfrageforschung und setzt Maßstäbe für das, was das Prädikat wissenschaftlich verdient. Kern der Umfrageforschung ist danach das Instrument der repräsentativen Stichprobe in Verbindung mit einer standardisierten Befragung. Qualitätskriterien für die Güte einer Umfrage sind daraus abgeleitet

- die verzerrungsfreie Abbildung einer definierten Grundgesamtheit durch die Stichprobe und

- die gültige Messung der erhobenen Sachverhalte durch die Befragungsinstrumente.

Die Schwierigkeit besteht darin, im konkreten Fall festzustellen, ob die Kriterien erfüllt sind und falls nicht, wo die Gründe dafür liegen. Um die Qualität von Umfragen zu bewerten, kann man sich auf zwei Indizien stützen. Die erste Frage, die zu beantworten ist, richtet sich auf die Ergebnisse der Umfrage: Sind die Ergebnisse richtig (im Sinne von wahr)? Die zweite Frage richtet sich auf die eingesetzten Methoden und Verfahren: Sind die korrekten Verfahren genutzt worden? Beide Punkte hängen zusammen, können aber in ihrer Bedeutung unterschiedlich gewichtet werden. Bilden Umfrageergebnisse die Grundlage für praktische Schlussfolgerungen und Entscheidungen, wird das erste Qualitätskriterium im Vordergrund stehen, also die richtigen und brauchbaren Ergebnisse. Methodische Anforderungen und Qualitätsstandards können sehr unterschiedlich sein; der methodische Aufwand soll jedoch in angemessenem Verhältnis zum Untersuchungsziel stehen. Im akademischen Bereich orientiert man sich stärker an den korrekten Verfahren. Forschungsgegenstand sind hier oft Einstellungen oder komplexe Konstrukte, für die die „wahre" Verteilung in der Grundgesamtheit nicht bekannt ist. Dabei geht man davon aus, dass bei Anwendung wissenschaftlich begründeter Verfahren die Umfrage den untersuchten Sachverhalt innerhalb gewisser Fehlergrenzen richtig abbildet.

Bei der Denkschrift handelt es sich insgesamt weniger um eine Festlegung von Qualitätsstandards als vielmehr um eine Methodenbeschreibung im Sinne der Darstellung von „best practices" der Umfrageforschung. Die Denkschrift formuliert Qualitätsmaßstäbe u. a. zu den folgenden Themen:

- Stichproben,
- Fragebogen,
- Feldarbeit,
- prozessorientierte Qualitätskontrollen und –standards,
- international vergleichende Umfrageforschung,
- Datenschutz und
- Veröffentlichung von Umfrageergebnissen.

Die Denkschrift hat maßgeblich dazu beigetragen, die sozialwissenschaftliche Qualitätsdebatte zu beleben (vgl beispielsweise Mueller/Nauck/Diekmann 2000, Ehling/Rendtel 2004). Bei Qualitätsbewertungen können, wie oben dargestellt, zwei Ebenen unterschieden werden: erstens die Qualität der eingesetzten Methoden und Verfahren sowie zweitens die erzielte Qualität einer Erhebung. Eine Umfrage, die methodisch höchste Qualitätsstandards aufweist, kann in ihrer Realisierung gleichwohl Qualitätsmängel aufweisen, während eine methodisch einfach durchgeführte Erhebung im Rahmen ihrer Möglichkeiten gute Ergebnisse liefern

kann. Qualitätsmaße und Qualitätsmaßstäbe müssen der verwendeten Methodik angemessen sein. Die sozialwissenschaftliche Qualitätsdebatte sollte das gesamte Spektrum möglicher Qualitätsniveaus einbeziehen. Aus finanziellen Gründen kann es sich die sozialwissenschaftliche Forschung nicht leisten, immer auf dem höchsten Qualitätsniveau zu bestehen, andererseits dürfen gerade vereinfachende Verfahren nicht von einer kritischen Qualitätsbewertung ausgenommen werden.

Die Sozialwissenschaften müssen gezielt neue Verfahren und Techniken der Umfrageforschung einbeziehen, um sowohl theoretisch als auch praktisch die Fortentwicklung der empirischen Forschung zu unterstützen. Sozialwissenschaftliche Lehrmeinungen und Qualitätsvorstellungen dürfen nicht der Entwicklung der Umfragepraxis hinterherhinken. Die Qualitätsbewertung von Umfragemethodik benötigt neben der theoretischen Fundierung auch eine empirische Basis in vermehrter Methodenforschung. Vergleichende Methodenexperimente und -evaluationen müssen gezielt gefördert werden. Mit dem DFG Schwerpunktprogramm „Survey-Methodologie“, das ab dem 1.1.2008 starten soll, sind jetzt auch Fördermöglichkeiten für die wissenschaftliche Weiterentwicklung der Umfrageforschung geschaffen worden.

4 Standards zur Qualitätssicherung in der Markt- und Sozialforschung

Vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. und dem Berufsverband Deutscher Markt- und Sozialforscher e.V. sind gemeinsame Standards zur Qualitätssicherung in der Markt- und Sozialforschung herausgegeben worden, die einen Orientierungsrahmen für die angestrebte Datenqualität bilden. Diese Standards beschreiben über den Erhebungsprozess hinaus auch die Beziehungen zwischen Instituten und Auftraggeber. Sie sind ein Katalog von verbindlichen Vorgaben, um eine hohe Ergebnisqualität zu erreichen, wobei die Art und Weise der Zielerreichung unter Anwendung anerkannter wissenschaftlicher, methodischer oder technischer Maßnahmen freigestellt ist. Die Untersuchungsergebnisse sind hier auch von den Ansprüchen, Erwartungen und Vorgaben des Auftraggebers abhängig, wobei immer nach wissenschaftlichen Grundsätzen, kostengünstig und mit angemessener Qualität gearbeitet werden soll.

Empfehlungen zur Qualitätssicherung für Online-Befragungen ergänzen die grundlegenden Standards. Ergebnis daraus ist u. a. die im Rahmen dieser Veranstaltung vorgestellte ISO Norm 20252 für Markt-, Meinungs- und Sozialforschung.

Markt- und Sozialforschung bewegen sich zumeist im Spannungsfeld zwischen wissenschaftlicher Qualität und zeitlicher Verfügbarkeit der Ergebnisse so-

wie den dafür zur Verfügung stehenden Budgets. Die Qualität der Untersuchungsergebnisse ist mithin auch entscheidend von den Ansprüchen und Erwartungen des Auftraggebers abhängig. Die Qualitätsstandards sollen bewirken, dass die Institute Forschung nach anerkannten wissenschaftlichen Grundsätzen, kostengünstig, mit angemessener Qualität und entsprechend den mit dem Auftraggeber vereinbarten Vorgaben durchführen.

In den Qualitätsstandards werden zuerst allgemeine Aspekte der Qualitätssicherung angesprochen wie z. B. ein übergreifendes Qualitätsmanagement, Mitarbeiterschulungen und die Einhaltung von Vertraulichkeit, Beratung und Zusammenarbeit mit dem Auftraggeber und die Studienkonzeption und -begleitung. Im Bereich der Datenerhebung werden Qualitätssicherungsmaßnahmen zur Anwerbung, Schulung und Kontrolle der Interviewer, der Feldsteuerung sowie der Kontrolle der Fragebogen und Daten beschrieben. Aus dem Bereich der Datenverarbeitung werden Verfahren für die Durchführung und Kontrolle der Codierung sowie der Datenerfassung, die Bereinigung der Rohdaten und deren Gewichtung beschrieben, die helfen, qualitativ hochwertige Daten zu liefern. Dies gilt auch für die Analyse und Darstellung der Ergebnisse und die Übermittlung der Untersuchungsergebnisse an den Auftraggeber.

5 Qualitätsstandards der amtlichen Statistik[1]

Um ihrem Informationsauftrag gerecht werden zu können, unterliegt gerade die Erstellung und Verbreitung amtlicher Statistiken strengen Qualitätsanforderungen. Zur Information von Nutzern, Auftraggebern und Befragten haben die Statistischen Ämter des Bundes und der Länder ihre Qualitätsstandards veröffentlicht. Die Publikation dokumentiert die grundlegenden Qualitätsleitlinien der amtlichen Statistik, die Maßnahmen zur Qualitätssicherung für alle Schritte der Statistikproduktion sowie den organisatorischen Rahmen der Qualitätssicherung.

Die Einhaltung der Standards gewährleistet, dass amtliche Statistiken unabhängig von politischer Einflussnahme und unter Nutzung anerkannter wissenschaftlicher Verfahren erstellt werden. In den Standards spiegelt sich die lange Tradition von Qualität in der amtlichen deutschen Statistik wider. Sie befinden sich in Übereinstimmung mit internationalen Grundsätzen, Leitlinien und Normen. Zu nennen sind insbesondere die gemeinsamen Qualitätsstandards, die die Statistischen Ämter in den Mitgliedstaaten der Europäischen Union und das Statistische Amt der Europäischen Gemeinschaften (Eurostat) – zusammenge-

1 Die Ausführungen dieses Kapitels basieren auf der Veröffentlichung Statistische Ämter des Bundes und der Länder (Hrsg.): Die Qualitätsstandards der amtlichen Statistik, Wiesbaden 2006

schlossen im Europäischen Statistischen System (ESS) – im Rahmen des Verhaltenskodex Europäische Statistiken (Code of Practice) vereinbart haben.

Übersicht 5: Verhaltenskodex europäische Statistiken

I. Der institutionelle Rahmen
- Grundsatz 1: Fachliche Unabhängigkeit
- Grundsatz 2: Auftrag zur Datenerhebung
- Grundsatz 3: Angemessene Ressourcen
- Grundsatz 4: Verpflichtung zur Qualität
- Grundsatz 5: Statistische Geheimhaltung
- Grundsatz 6: Unparteilichkeit und Objektivität

II. Die statistischen Prozesse
- Grundsatz 7: Eine solide Methodik
- Grundsatz 8: Geeignete statistische Verfahren
- Grundsatz 9: Vermeidung einer übermäßigen Belastung der Auskunftgebenden
- Grundsatz 10: Wirtschaftlichkeit

III. Die statistischen Produkte
- Grundsatz 11: Relevanz
- Grundsatz 12: Genauigkeit und Zuverlässigkeit
- Grundsatz 13: Aktualität und Pünktlichkeit
- Grundsatz 14: Kohärenz und Vergleichbarkeit
- Grundsatz 15: Zugänglichkeit und Klarheit

Der Verhaltenskodex dient der Sicherstellung eines einheitlichen Qualitätsniveaus in der Europäischen Union. Er findet in der deutschen amtlichen Statistik uneingeschränkte Anwendung. Institutionelle und organisatorische Faktoren wirken sich maßgeblich auf die Effizienz und Glaubwürdigkeit der amtlichen Statistik aus. Die in diesem Zusammenhang relevanten Qualitätskriterien haben den Charakter von unverzichtbaren Grundsätzen. Diese Grundsätze leiten sich überwiegend direkt aus dem gesetzlichen Auftrag der statistischen Ämter ab und sind vor allem im Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz) und den Landesstatistikgesetzen festgeschrieben.

5.1 Qualitätsstandards für den institutionellen Rahmen

Das Bundesstatistikgesetz schreibt für die amtliche Statistik die Grundsätze der wissenschaftlichen Unabhängigkeit, Neutralität und Objektivität fest. Das Gesetz garantiert damit, dass amtliche Statistiken unabhängig von politischer und anderer externer Einflussnahme erstellt und verbreitet werden.

- Im Rahmen der die Statistiken anordnenden Gesetze und EU-Verordnungen tragen die statistischen Ämter die alleinige Verantwortung für die Festlegung der statistischen Methoden und Verfahren sowie für den Inhalt und den Zeitplan statistischer Veröffentlichungen. Die Festlegung erfolgt aufgrund statistischer Überlegungen unter Beachtung der rechtlichen, finanziellen und organisatorischen Rahmenbedingungen. Die angewandten Methoden und Verfahren werden offengelegt.
- Die Veröffentlichung und Kommentierung statistischer Ergebnisse erfolgt objektiv, unparteilich und getrennt von politischen bzw. Grundsatzerklärungen.
- Die Ergebnisse der amtlichen Statistik sind allgemein zugänglich, d.h. alle Nutzer haben gleichberechtigten Zugang zu den Ergebnissen. Grunddaten sind zudem in jedem Fall kostenlos erhältlich. Sie werden jedermann als öffentliches Gut über die Medien, das Internet oder auf Anfrage zur Verfügung gestellt. Neben der allgemeinen Grundversorgung und den nachfrage- oder zielgruppenorientierten Standardangeboten wird zudem eine nach individuellen Bedürfnissen maßgeschneiderte Datenaufbereitung gegen Kostenerstattung angeboten.

Die Entscheidung, welche statistischen Erhebungen durchgeführt werden, liegt in Deutschland beim Gesetzgeber und nicht bei der Bundesregierung, der jeweiligen Landesregierung oder bei den statistischen Ämtern selbst. Gemäß dem Prinzip der Gesetzmäßigkeit müssen alle statistischen Erhebungen vom Gesetzgeber angeordnet werden.

- Für alle Erhebungen der amtlichen Statistik gibt es eine gesetzliche Grundlage, in der im Einzelnen die Parameter einer Erhebung, wie z. B. Grundgesamtheit, Erhebungsumfang und Erhebungseinheiten, Erhebungsmerkmale sowie Erhebungstermin, festgelegt sind.
- Sofern dies für die Erzielung zuverlässiger Ergebnisse erforderlich ist, sind die Befragten zur Auskunft verpflichtet. Die Auskunftspflicht wird in diesem Fall vom Gesetzgeber festgeschrieben. Sie garantiert eine hohe Rücklaufquote sowie aussagefähige Ergebnisse auch bei tiefer fachlicher und regionaler Gliederung.
- Das Bundesstatistikgesetz regelt auch die Möglichkeit der Verwendung von Verwaltungsdaten für statistische Zwecke, wodurch ggf. aufwändige Direkterhebungen vermieden werden können.

Auf die Einhaltung des Datenschutzes legen die statistischen Ämter größten Wert. Die Geheimhaltung von Einzelangaben der Befragten und deren ausschließliche Verwendung für statistische Zwecke sind in jedem Fall gewährleistet und zudem gesetzlich vorgeschrieben.

- Bei jeder Veröffentlichung wird überprüft, ob die statistische Geheimhaltung gewährleistet ist.
- Die entsprechenden Verfahren zur Wahrung des Statistikgeheimnisses sind dokumentiert und werden von allen statistischen Ämtern angewendet. Damit garantieren die statistischen Ämter den Befragten den Schutz vor Offenlegung ihrer Einzelangaben. So können diese wahrheitsgemäß antworten, ohne befürchten zu müssen, dass ihre Angaben von anderen Behörden, wie z. B. von der Steuerverwaltung, oder von einem konkurrierenden Unternehmen verwendet werden.
- Für externe Datennutzer aus anerkannten wissenschaftlichen Einrichtungen besteht die Möglichkeit, statistische Einzeldaten zu Forschungszwecken auszuwerten. Das Bundesstatistikgesetz enthält in diesem Fall jedoch strenge Vorschriften, die die Geheimhaltung und den Datenschutz sicherstellen.
- Die Sicherung und wenn möglich Verbesserung des erreichten Qualitätsniveaus bestimmt das Handeln der statistischen Ämter. Die statistischen Ämter ergreifen alle Maßnahmen, die zum Erreichen dieses Ziels erforderlich sind.
- Die statistischen Ämter wenden bei der Statistikerstellung einheitlich sachgerechte und anerkannte Methoden an, die ein hohes Qualitätsniveau und volle Vergleichbarkeit der dezentral erhobenen Daten gewährleisten.
- Die Produktqualität wird regelmäßig anhand der im ESS vereinbarten Qualitätskriterien überwacht. Für alle Statistiken werden Qualitätsberichte erstellt und veröffentlicht, die Informationen über die Methodik und alle relevanten Qualitätsaspekte enthalten.
- Die Leiter der Nationalen Statistischen Ämter im ESS haben darüber hinaus gemeinsame Qualitätsgrundsätze für alle wesentlichen Aspekte der statistischen Arbeit aufgestellt, die in der „Qualitätserklärung des ESS“ festgeschrieben sind.

Die Bundesregierung und die Landesregierungen stellen ihren statistischen Ämtern personelle und finanzielle Ressourcen in einem Umfang zur Verfügung, der ausreicht, ihren Auftrag zu erfüllen. Im Rahmen der jährlichen Haushaltsverhandlungen besteht Gelegenheit, die Ressourcen den sich wandelnden Anforderungen und Prioritäten in der amtlichen Statistik anzupassen.

5.2 Qualitätsstandards für die statistischen Prozesse

Prozessbezogene Qualitätsstandards betreffen alle Phasen der Statistikerstellung: von der Ermittlung des Datenbedarfs über die Vorbereitung, die Datengewinnung, die Aufbereitung der Daten, die Hochrechnung, die Auswertung und Analyse der Ergebnisse bis zur Verbreitung der Ergebnisse. In jedem Arbeitsschritt sind spezielle Qualitätsanforderungen zu erfüllen. Zugleich gibt es übergreifende Qualitätsstandards, die in allen Prozessen der Statistikerstellung angewendet werden.

Qualitativ hochwertige Statistiken müssen auf einer adäquaten Methodik basieren. Dies erfordert den Einsatz geeigneter Instrumente und Verfahren sowie ein entsprechendes Know-how. Die amtliche Statistik orientiert sich dabei an den aktuellen wissenschaftlichen Erkenntnissen, internationalen Normen und Standards, Leitlinien und vorbildlichen Lösungen sowie ihrem reichen Erfahrungsschatz.

- Die statistischen Methoden, die von den statistischen Ämtern eingesetzt werden, entsprechen dem aktuellen Stand der wissenschaftlichen Forschung. Internationale Normen und Standards werden – soweit sinnvoll – durchgehend beachtet. Zugleich werden mit anderen statistischen Ämtern und der Wissenschaft Erfahrungen ausgetauscht und vorbildliche Ansätze übertragen.
- Die statistischen Ämter gewährleisten die korrekte Umsetzung und Einhaltung der vereinbarten Methoden und Verfahren durch systematische Qualitätssicherungsmaßnahmen in allen Phasen der Erstellung und Verbreitung von Statistiken.
- Die verwendete statistische Methodik wird in der nationalen und internationalen Fachöffentlichkeit zur Diskussion gestellt sowie in anerkannten Fachpublikationen veröffentlicht.
- Weiterentwicklungen der statistischen Methodik werden – soweit erforderlich – in Zusammenarbeit mit der Wissenschaft durchgeführt.
- Durch eine individuelle Fortbildungsplanung und Personalentwicklung stellen die statistischen Ämter sicher, dass alle Mitarbeiterinnen und Mitarbeiter über die für die jeweiligen Aufgaben erforderlichen Kenntnisse und Fähigkeiten verfügen.

Die statistischen Ämter setzen die Ressourcen wirtschaftlich ein. Insbesondere im Zuge der Einführung neuer Technologien, der ämterübergreifenden Aufgabenerledigung sowie der Nutzung von Daten aus Verwaltungsregistern werden alle Möglichkeiten zur Steigerung der Produktivität genutzt.

- Die amtliche Statistik schöpft das Produktivitätspotential, das moderne Informations- und Kommunikationstechnologien bieten, in allen Phasen der Statistikerstellung so weit wie möglich aus.

- Die statistischen Ämter gewährleisten die Effizienz des statistischen Systems in Deutschland unter anderem durch arbeitsteilige Zusammenarbeit, indem ein Amt für alle anderen die Software für die statistische Aufbereitung entwickelt.
- Im Rahmen der „Optimierten Kooperation" wird die Arbeitsteilung derzeit wesentlich ausgebaut. Dabei übernimmt beispielsweise ein Amt gegen Kostenerstattung die Statistikaufbereitungen für andere Ämter. Das Amt, das bei einer Ausschreibung die wirtschaftlichste Erledigung der jeweiligen Aufgabe anbietet, führt die Arbeiten aus.
- In der Verwaltung bereits vorhandene Daten werden so weit wie möglich auch für statistische Zwecke genutzt, um aufwändige zusätzliche Datenerhebungen zu vermeiden.
- Die statistischen Ämter wenden für wichtige Teilprozesse der Statistikproduktion verstärkt standardisierte Verfahren und Techniken an, die eine einheitliche Bearbeitung ermöglichen.

Die statistischen Ämter halten die durch statistische Erhebungen bedingte Belastung der Unternehmen sowie der Bürgerinnen und Bürger so gering wie möglich.

- Im Rahmen der Vorbereitung und Durchführung von Statistiken wird von der amtlichen Statistik stets geprüft, inwieweit mit Blick auf die Ergebnisqualität und die Praktikabilität der Vorgehensweise die Belastung der Befragten so gering wie möglich gehalten werden kann.
- Zur Minimierung der Belastung werden insbesondere folgende Maßnahmen eingesetzt: verstärkte Nutzung von Verwaltungsdaten, Einsatz von Stichprobenverfahren, Rotation von Stichprobeneinheiten bei periodisch wiederholten Erhebungen, direkte elektronische Übernahme von Unternehmensdaten aus den Buchführungsunterlagen, Einsatz verbesserter Schätzverfahren zur Minimierung des Umfangs von Stichproben, Vermeidung von Mehrfacherfassungen gleicher Merkmale.

5.3 Die Qualitätsstandards für die statistischen Produkte

Die Qualität statistischer Produkte definiert sich über ein breites Spektrum unterschiedlicher Kriterien. In der amtlichen Statistik werden – wie bereits einleitend erläutert – einheitlich die Qualitätskriterien Relevanz, Genauigkeit, Aktualität und Pünktlichkeit, Verfügbarkeit und Transparenz, Vergleichbarkeit sowie Kohärenz angewendet, die im ESS entwickelt wurden.

Für alle Bundesstatistiken stehen Qualitätsberichte zur Verfügung, die (neben Informationen zu den angewendeten Methoden) Angaben hinsichtlich sämtlicher Qualitätskriterien des ESS enthalten. Die Qualitätsberichte stellen alle für eine sachgerechte Nutzung der Ergebnisse erforderlichen Informationen bereit. Sie

sind über das Internet zugänglich und zusätzlich in alle Fachserien des Statistischen Bundesamtes integriert.

Relevanz

Die amtlichen Statistiken müssen den Anforderungen der Nutzer entsprechen.

- Die Hauptnutzer haben durch ihre Beteiligung an Fachausschüssen und Arbeitskreisen des Statistischen Beirats die Möglichkeit, bei verändertem oder neu auftretendem Informationsbedarf Einfluss auf die Statistik anordnende Gesetzgebung zu nehmen und so ihre Anforderungen in das statistische Programm einfließen zu lassen.
- Zukünftige Informationsanforderungen werden von der amtlichen Statistik frühzeitig aufgegriffen und in Vorschläge zur Anpassung des statistischen Programms integriert.
- Es werden regelmäßig Befragungen zur Nutzerzufriedenheit durchgeführt.

Genauigkeit

Statistiken sollen die Realität möglichst genau und zuverlässig widerspiegeln. Bei ihrer Interpretation ist aber zu beachten, dass statistische Ergebnisse in der Regel – trotz aller Sorgfalt und Gründlichkeit bei der Ermittlung – mit einem Unschärfebereich behaftet sind, der in der Statistik als Fehler bezeichnet wird. Unterschieden werden stichprobenbedingte Fehler und nicht-stichprobenbedingte Fehler. Stichprobenbedingte Fehler umfassen zum einen den Stichprobenzufallsfehler, zum anderen mögliche Verzerrungen auf Grund des Hochrechnungsverfahrens. Nicht-stichprobenbedingte Fehler umfassen Fehler durch die Erfassungsgrundlage, Fehler durch Antwortausfälle, Messfehler und Aufbereitungsfehler.

- Für die statistischen Ergebnisse wird stets eine Genauigkeit angestrebt, die sich am Verwendungszweck der Ergebnisse orientiert. Durch Einsatz adäquater Methoden und Verfahren wird sichergestellt, dass die angestrebte Genauigkeit erreicht wird.
- Die amtliche Statistik analysiert – soweit möglich – stichprobenbedingte wie auch nicht-stichprobenbedingte Fehler, um eine missbräuchliche und fehlerhafte Nutzung der Ergebnisse zu vermeiden. Die Ergebnisse der Fehleranalysen werden veröffentlicht.
- Bei vorläufigen Ergebnissen wird das Revisionsverhalten analysiert und der mittlere Revisionsbedarf zum endgültigen Ergebnis veröffentlicht.

Aktualität und Pünktlichkeit

Statistiken müssen aktuell sein und pünktlich verbreitet werden.

- Die statistischen Ämter erstellen die Statistiken möglichst zeitnah und veröffentlichen die Ergebnisse zu zuvor festgelegten Terminen. Die amtliche Statistik erfüllt die höchsten europäischen und internationalen Aktualitätsanforderungen.
- Für Statistiken, bei denen die Aktualität der Ergebnisse besonders im Vordergrund steht (z. B. für Konjunkturstatistiken), werden vorläufige Ergebnisse veröffentlicht, sofern diese eine akzeptable Genauigkeit aufweisen.
- Die Periodizität der Veröffentlichungen trägt dem Nutzerbedarf weitestgehend Rechnung.

Verfügbarkeit und Transparenz

Statistische Ergebnisse müssen für den Nutzer leicht zugänglich und hinsichtlich Konzept und Methoden vollständig dokumentiert sein.

- Die statistischen Ämter verbreiten ihre Produkte nach einem gemeinsamen Marketingkonzept, das sich an den Anforderungen der Nutzer orientiert. Dabei wird die Vielfalt der heute möglichen Verbreitungsformen, wie Internet, Datenbanken (GENESIS), gedruckte Veröffentlichungen, CD-ROMs und Telekommunikation (Telefon, Telefax) genutzt.
- Die Nutzer werden in Fachveröffentlichungen und im Internet fortlaufend über die Methoden und die Qualität der statistischen Produkte informiert. Statistische Metadaten werden im Einklang mit internationalen Standards dokumentiert.
- Standardinformationen werden kostenlos zur Verfügung gestellt. Maßgeschneiderte Aufbereitungen werden gegen Kostenerstattung erstellt.
- Für alle Bundesstatistiken informieren Qualitätsberichte über die Datenqualität sowie die angewendete Erhebungsmethodik.
- Für viele Statistiken ist der Zugang zu Mikrodaten für Forschungszwecke gestattet. Dieser ist jedoch beschränkt auf anerkannte wissenschaftliche Einrichtungen und unterliegt zugleich strengen Geheimhaltungsvorschriften.

Vergleichbarkeit

Statistische Ergebnisse sollen zuverlässige zeitliche, räumliche und fachliche Vergleiche ermöglichen.

- Um eine möglichst hohe Vergleichbarkeit zu erreichen, verwendet die amtliche Statistik bei der Erstellung von Statistiken einheitliche – wenn möglich internationale – Standards in Bezug auf die Definitionen, die Einheiten, die Merkmale und die Klassifikationen.

- Unumgängliche Änderungen, wie z. B. Wechsel von Klassifikationen, Änderungen der Erhebungsmerkmale, werden dokumentiert.

Kohärenz

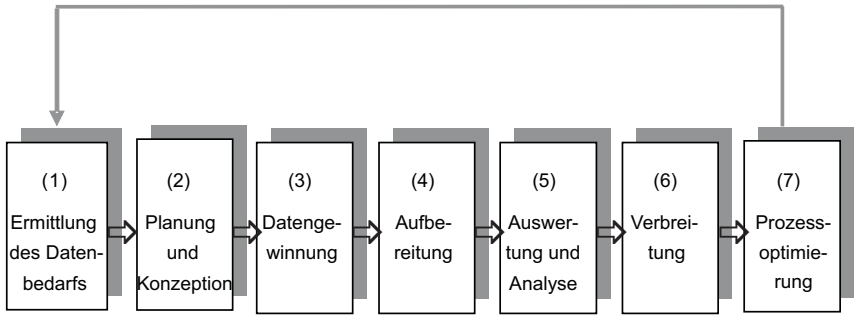
Statistische Ergebnisse, die sich auf die gleiche Grundgesamtheit beziehen, aber aus unterschiedlichen Statistiken stammen, sollen möglichst kohärent sein, d. h. widerspruchsfrei untereinander in Beziehung gesetzt werden können. Es sollte möglich sein, miteinander zusammenhängende Daten aus unterschiedlichen Quellen zu kombinieren und zusammen zu verwenden.

- Die amtliche Statistik prüft relevante Statistiken auf Kohärenz, z. B. zwischen jährlichen und unterjährlichen Statistiken, vorläufigen und endgültigen Ergebnissen.
- Auftretende Unterschiede und ihre Ursachen werden – soweit möglich – erläutert und quantifiziert.

Einige dieser Qualitätskriterien stehen in Konkurrenz zueinander, so dass nicht alle Kriterien gleichzeitig optimiert werden können. Dies trifft insbesondere auf das Verhältnis von Genauigkeit und Aktualität zu. So hängt es u. a. von den Anforderungen der Nutzer ab, ob ein Verlust an Genauigkeit, z. B. bei der Bereitstellung von vorläufigen Ergebnissen, zugunsten einer schnelleren Datenbereitstellung in Kauf genommen werden kann oder sollte. Allerdings muss eine Verbesserung der Aktualität nicht zwangsläufig mit einem Verlust an Genauigkeit verbunden sein, denn im Rahmen einer Effizienzsteigerung der verschiedenen Prozesse der Statistikerstellung können Aktualitätsgewinne bei gleich bleibender Genauigkeit erzielt werden. Stichworte sind in diesem Zusammenhang der Einsatz neuer Erhebungstechniken, die Nutzung flexiblerer Auswertungssoftware sowie moderner Kommunikationsmittel und die Bereitstellung von vorläufigen Ergebnissen.

5.4 Qualitätssicherung in der amtlichen Statistik

Amtliche statistische Informationen sind das Ergebnis einer komplexen Kette zahlreicher Arbeitsschritte. Die einzelnen Schritte müssen ineinander greifen und aufeinander abgestimmt sein, damit die Nutzeranforderungen optimal erfüllt werden können. Der Weg zum Endprodukt kann in sieben „Grundprozesse“ gegliedert werden, die in der Regel arbeitsteilig von den Statistischen Ämtern der Länder und dem Statistischen Bundesamt durchgeführt werden. In all diesen Grundprozessen greifen die jeweiligen Maßnahmen der Qualitätssicherung. Die Darstellung der Verfahren zur Qualitätssicherung orientiert sich daher am Ablauf einer Statistik.



Übersicht 6: Grundprozesse der Statistikerstellung

(1) Ermittlung des Datenbedarfs: Durch eigene Beobachtungen und durch Kontakte zu vielfältigen gesellschaftlich relevanten Institutionen und Gruppen ermitteln die statistischen Ämter neu entstehenden Datenbedarf. Kann dieser auf Grundlage bereits vorhandener Daten nicht abgedeckt werden, macht die amtliche Statistik Vorschläge, wie dem neuen Informationsbedarf begegnet werden kann. Dazu kann auch gehören, eine neue Erhebung vorzuschlagen. Eine wichtige Rolle bei der Ermittlung neuen Datenbedarfs wie bei der Weiterentwicklung des Statistischen Programms insgesamt spielen der Statistische Beirat und dessen Fachausschüsse, in denen zentrale Nutzergruppen vertreten sind.

(2) Planung und Konzeption: Auf der Basis des ermittelten Datenbedarfs wird die Datengewinnung vorbereitet. Hierzu zählen die Mitwirkung bei der Entwicklung der Rechtsgrundlage durch Beratung und Stellungnahme, die Festlegung der Erhebungs- beziehungsweise der Auswahlgrundlage, die Auswahl einer Erhebungsmethode, die Umsetzung der zu erhebenden Merkmale in ein Erhebungsinstrument, der Test des Erhebungsinstrumentes, die Stichprobenplanung, die Organisation der Feldarbeit, die Planung der Aufbereitung, Auswertung, Analyse und Verbreitung sowie der Qualitätsüberprüfung im Rahmen der Prozessoptimierung.

(3) Datengewinnung: Diese Phase enthält die praktischen Schritte der Datenerhebung im Rahmen der Feldarbeit bei Befragungen oder der Nutzung von Verwaltungsdaten.

(4) Aufbereitung: Im Rahmen der Aufbereitung werden die Daten erfasst (falls diese noch nicht in elektronischer Form vorliegen), Fehler in den Daten durch Plausibilitätskontrollen ermittelt und zum Beispiel durch Rückfragen oder mathematische Verfahren bereinigt. Auf Grundlage der plausibilisierten Daten folgt dann die Ergebniserstellung. Bei Stichprobenerhebungen geschieht dies mittels

einer Hochrechnung, ergänzt durch eine Fehlerrechnung. Am Ende der Aufbereitung stehen Standardtabellen und ein Datensatz für die Durchführung weiterer Auswertungen.

(5) Auswertung und Analyse: Im Anschluss an die Aufbereitung werden weiterführende Datenanalysen durchgeführt. Zusätzlich fließen die Ergebnisse aus zahlreichen Erhebungen in statistische Gesamtsysteme ein. Hinzu kommen die Dokumentation der Erhebungen und deren Datenqualität sowie die Analyse und Interpretation der Daten.

(6) Verbreitung: Die Verbreitung statistischer Informationen stellt den Abschluss der Erhebungs- und Analysetätigkeiten dar. Grundlage hierfür ist das Marketingkonzept der statistischen Ämter. Je nach Kundeninteresse werden die statistischen Informationen als kostenlose Grundversorgung, nachfrage- und zielgruppenorientierte Standardangebote sowie als maßgeschneiderte Beratung und Aufbereitung angeboten. Zusätzlich zu den Daten selbst werden für alle Statistiken Qualitätsberichte mit Angaben zu Methoden, Definitionen und Datenqualität bereitgestellt. Mit Internetangeboten, gedruckten Veröffentlichungen, über den Infoservice und auf Presseveranstaltungen werden die Daten über eine Vielzahl von Zugangswegen präsentiert.

(7) Prozessoptimierung: Ein wesentliches Ziel ist die ständige Verbesserung der Datenqualität und die Steigerung der Effizienz durch die kontinuierliche Analyse und Verbesserung aller oben genannten Arbeitsprozesse und ihrer Ergebnisse. Hierzu besteht in allen Ämtern ein System von Qualitätskontrollen.

6 Schlusswort

Mit der Einführung sollte der Gesamtrahmen der Veranstaltung angesprochen werden. Im Rahmen der Tagung ist geplant, einige Schlaglichter auf das Qualitätsmanagement und die Qualitätssicherung zu legen. Die Schwerpunkte der Konferenz liegen weniger auf Aspekten und Maßnahmen des Managements von Datenqualität. Die folgenden Themen stehen dabei im Mittelpunkt:

- Standards und Normen,
- Qualitätskontrolle,
- Qualitätsbewertung und
- Qualitätssicherung.

Wie auch in den vergangenen gemeinsamen Veranstaltungen vom Arbeitskreis deutscher Markt- und Sozialforschungsinstitute e.V. (ADM), der Arbeitsgemeinschaft sozialwissenschaftlicher Institute e.V. (ASI) und dem Statistischen Bundesamt kommen auch bei dieser Tagung Referenten aus allen beteiligten Institu-

ten zu Wort, um den gegenseitigen Dialog und das Lernen voneinander zu fördern.

Literatur

Chlumsky, J., Wiegert, R. u.a. (1993): Qualität statistischer Daten. Beiträge zum wissenschaftlichen Kolloquium am 12./13. November 1992 in Wiesbaden, Wiesbaden

Deutsche Forschungsgemeinschaft (1998): Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft", Weinheim

Ehling, M., Rendtel, U. et al. (2004): Harmonisation of Panel Surveys and Data Quality, Wiesbaden

Kaase, M. (Hrsg.) (1999): Qualitätskriterien der Umfrageforschung. Denkschrift, Berlin

Kopsch, G., Köhler, S., Körner, T. (2006): Der Verhaltenskodex Europäische Statistiken (Code of Practice), in: Wirtschaft und Statistik 8/2006, S. 793-804

Mueller, U., Nauck, B., Diekmann, A. (Hrsg.) (2000): Handbuch der Demographie. Band 1. Berlin, Heidelberg, New York

Statistische Ämter des Bundes und der Länder (Hrsg.) 2006: Die Qualitätsstandards der amtlichen Statistik, Wiesbaden

Normung

Holger Mühlbauer

1 DIN, CEN, ISO - Drei Buchstaben und ihre Bedeutung

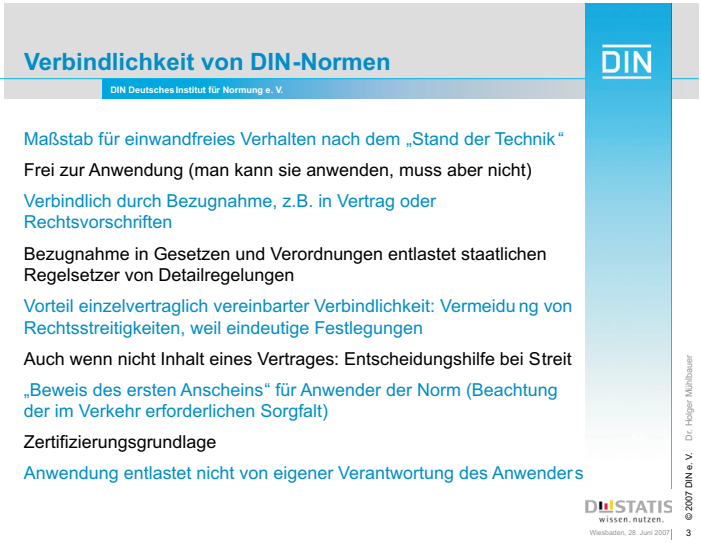
Das DIN ist ein eingetragener gemeinnütziger Verein mit Sitz in Berlin (DIN Deutsches Institut für Normung e. V., gegründet 1917). Das DIN ist die für die Normungsarbeit zuständige Institution in Deutschland. Dieser Status wurde im Vertrag mit der Bundesrepublik Deutschland am 5. Juni 1975 anerkannt.

Das DIN organisiert die deutsche Interessenvertretung in den europäischen und internationalen Normungsorganisationen CEN bzw. ISO. Alle drei Organisationen sind in ihrer Rechtsform privatrechtliche Institutionen, also keine öffentlich-rechtlichen Stellen mit Rechtsetzungsbefugnis. Jedoch sind sie ihrer Art nach mit herausgehobener Autorität ausgestattet.

Das DIN ist der „runde Tisch“, an dem sich Hersteller, Handel, Verbraucher, Handwerk, Dienstleistungsunternehmen, Wissenschaft, technische Überwachung, Staat, d.h. jedermann, der ein Interesse an der Normung hat, zusammensetzen, um den „state of the art“ zu ermitteln und unter Berücksichtigung neuer Erkenntnisse in Normen niederzuschreiben, und zwar möglichst im Konsens.

DIN-Normen bilden einen Maßstab für einwandfreies, normgerechtes Verhalten, und sind im Rahmen der Rechtsordnung mittelbar von Bedeutung.

DIN-Normen stehen jedermann zur Anwendung frei. Das heißt, man kann sie anwenden, muss es aber nicht. DIN-Normen werden verbindlich durch Bezugnahme, z. B. in einem privatrechtlichen Vertrag zwischen privaten Parteien, bei Ausschreibungen oder in Gesetzen und Verordnungen. Der Vorteil der einzelvertraglich vereinbarten Verbindlichkeit von Normen kann darin liegen, dass sich inhaltliche Rechtsstreitigkeiten vermeiden lassen, wenn die in Bezug genommenen Normen eindeutige Festlegungen enthalten. Die Bezugnahme in Gesetzen und Verordnungen entlastet den Staat und die Bürger von rechtlichen Detailregelungen.



Auch in den Fällen, in denen DIN-Normen von Vertragsparteien nicht zum Inhalt eines Vertrages gemacht worden sind, dienen DIN-Normen im Streitfall als Entscheidungshilfe, wenn es im Kauf- und Werkvertragsrecht um Sachmängel geht. Hier spricht der Beweis des ersten Anscheins für den Anwender der Norm in dem Sinne, dass er die im Verkehr erforderliche Sorgfalt beachtet hat.

DIN-Normen sind jedoch keine Lehrbücher. Deshalb muss jemand, der sie anwendet, soviel Sachverstand haben, dass er die Verantwortung für sein Handeln selbst übernehmen kann.

Das geschützte Zeichen „DIN“ steht für „Deutsches Institut für Normung e.V.“.

Normung

DIN Deutsches Institut für Normung e. V.

Auf Antrag

Für eine Normungsaufgabe jeweils ein Arbeitsgremium zuständig

Im Regelfall mehrere fachlich verwandte Arbeitsgremien zu einem DIN-“Normenausschuss“ zusammengefasst

3.176 Arbeitsgremien in 74 Normenausschüssen
26.100 externe Experten
400 DIN-Mitarbeiter
30.000 Normen

Finanzierung anteilig aus Fördermitteln der interessierten Kreise und DIN-Eigenmitteln

DIN

© 2007 DIN e. V. Dr. Holger Mühlbauer

DESTATIS wissen. nutzen.

Normungsverfahren (national)

DIN Deutsches Institut für Normung e. V.

Antrag
↓
Arbeitsgremium
↓
Entwurf
↓
Entwurfspublikation
↓
Einspruchsverfahren
↓
Schlussfassung
↓
Veröffentlichung

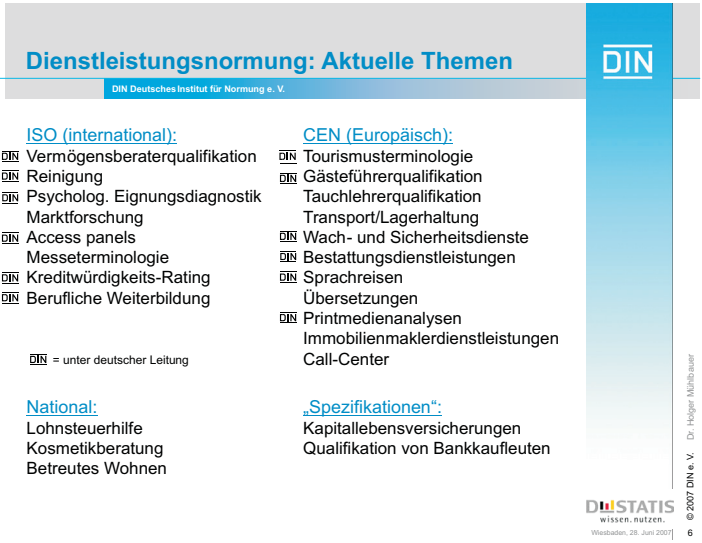
DIN

© 2007 DIN e. V. Dr. Holger Mühlbauer

DESTATIS wissen. nutzen.

2 Dienstleistungsnormung

Seit etwa 1994 widmet sich die Normung verstärkt dem Dienstleistungsbereich. Wie schon in der traditionellen Produktnormung, ist das DIN auch hier engagiert und zusammen mit anderen Ländern erfolgreich in der Etablierung von entsprechenden Normungsvorhaben. Verantwortlich für die Mehrzahl der Projekte ist auf deutscher Seite der Normenausschuss Gebrauchstauglichkeit und Dienstleistungen (NAGD) im DIN.

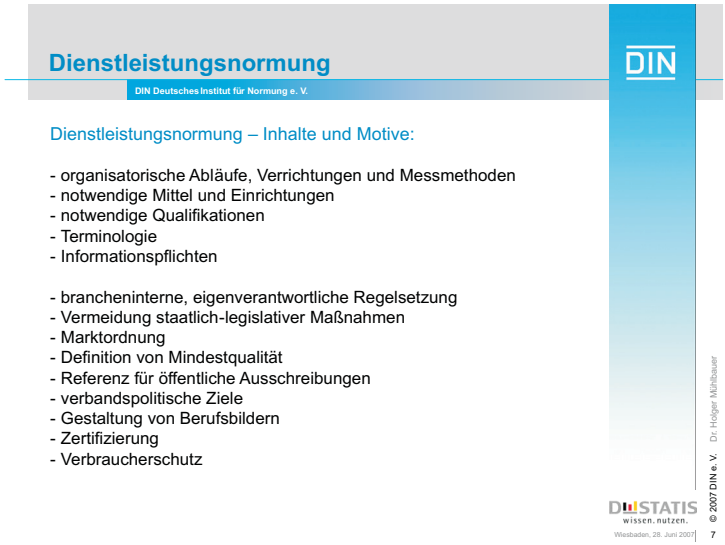


Neben der europäischen Normung als flankierende Maßnahme zur Ausgestaltung des europäischen Binnenmarktes wird Dienstleistungsnormung zunehmend auf internationaler, d.h. ISO-Ebene verankert. Die rein nationale Regelsetzung dient demgegenüber oftmals nur noch als Vorbereitung auf künftige internationalisierte Projekte.

Die Problemlage ist oftmals gleich: keine oder keine ausreichende gesetzliche Zugangsregelung, eine unbestimmte Zahl „schwarzer Schafe”, ungeregelte Geschäftspraktiken zum Schaden des Kunden bzw. zum Schaden des seriösen Teils der Branche. Aus diesem Grund stehen hinter vielen der Projekte die jeweiligen Verbände als Initiatoren. Bei fast allen nationalen Normungsvorhaben wird deshalb auch die Überlegung angestellt, die DIN-Norm als deutsche Position in ein mögliches künftiges transnationales Normungsvorhaben einfließen zu lassen.

Dienstleistungsnormung ist die Festlegung von Anforderungen an die

- organisatorischen Abläufe, Prozesse, Verrichtungen und Messmethoden im Rahmen einer Dienstleistung;
- für die Dienstleistung notwendigen Mittel und Einrichtungen, ohne detaillierte Produktanforderungen;
- für die Dienstleistung notwendigen Qualifikationen;
- im Rahmen der Dienstleistung verwendete Terminologie;
- dem Empfänger der Dienstleistung gegenüber zu unterbreitenden Informationen.



Im Unterschied zu Produktnormung, bei der präferierte technische Lösungen im Vordergrund stehen, treten in der Dienstleistungsnormung bei den interessierten Kreisen andere Beweggründe verstärkt hervor.

Dienstleistungsnormen im hier verwendeten Verständnis dienen der

- Vermeidung staatlich-legislativer Maßnahmen durch brancheninterne, eigenverantwortliche Regelsetzung;
- Schaffung von Marktzugangsempfehlungen;
- Definition von Mindestqualität;
- Erarbeitung von Referenzdokumenten für öffentliche Ausschreibungen;
- Umsetzung legitimer verbandspolitischer Ziele;
- Gestaltung von Berufsbildern;
- Generierung von Zertifizierungsgrundlagen

und ganz allgemein dem Verbraucherschutz.

In jedem Fall ist Dienstleistungsnormung eingebettet in politische bzw. verbands-, industrie- bzw. wirtschaftspolitische Bestrebungen auf nationaler, europäischer oder internationaler Ebene.

3 Aktuelle Beispiele

Im Jahr 2007 nahm eine Reihe von neu etablierten ISO-Komitees ihre Tätigkeit auf. Die Gremien widmen sich einerseits sehr spezifischen, gleichzeitig aber komplexen ISO-Projekten, mit denen der Internationalisierung von Dienstleistungen Rechnung getragen wird. Einige der Komitees bilden die neu geschaffene ISO-Organisationsform „Project Committee“ („PC“) ab. „Project Committees“ sind flexible Arbeitsplattformen für innovative Einzelprojekte. Sie agieren grundsätzlich wie die bisherigen klassischen „Technischen Komitees“ („TC“).

Alle Vorhaben gehen auf Initiativen Deutschlands bzw. des DIN zurück, um an die Stelle nur theoretischer Erörterungen zur Bedeutung von „Dienstleistungsnormung“ konkrete Aktivitäten treten zu lassen. Deutschland stellt in allen Fällen den Vorsitz:

ISO/PC 230 „Psychologische Eignungsdiagnostik“

Thema der künftigen ISO-Norm, deren Erarbeitung in Berlin beschlossen wurde, sind „Verfahren und Methoden für die Eignungsbeurteilung von Menschen in Beschäftigungssituationen“. Die beteiligten Experten vertreten psychologische Berufsorganisationen und Forschungseinrichtungen oder gehören einschlägigen Dienstleistungsanbietern und Fachverlagen an. Inhaltliche Schwerpunkte der künftigen Norm sind die Auswahl, Integration, Implementierung und Bewertung von Eignungsbeurteilungsverfahren. Darüber hinaus soll sie die Interpretation der Beurteilungsergebnisse und die daraus abgeleiteten Entscheidungen sowie Qualifikationsanforderungen an die ausführenden Personen und ethische Prinzipien des Verfahrens zum Gegenstand haben. Vor dem Hintergrund wachsender Bedeutung psychologischer Eignungstests in der Personalauswahl ist es unerlässlich, die dafür benötigten Instrumente und Prozesse einer genauen Prüfung zu unterziehen. Eine professionell durchgeführte Planung, Ausführung sowie Ergebnisinterpretation solcher Eignungstests ist sowohl für den Kunden als auch für den Anbieter von erheblicher wirtschaftlicher Bedeutung.

ISO/PC 231 „Markenwertmessung“

Auf der ersten Sitzung des Komitees wurde die Erarbeitung einer ISO-Norm mit Anforderungen, Methoden und Verfahren zur monetären Markenwertmessung

beschlossen. In einer globalisierten Wirtschaft kommt es immer häufiger zu Unternehmensfusionen und -übernahmen. Auch der härter werdende globale Wettbewerb macht eine zuverlässige und vergleichbare monetäre Bestimmung des Unternehmenserfolgs notwendig. Der Markenwertmessung kommt in beiden Fällen eine wichtige Funktion in der Beurteilung unternehmerischen Potenzials zu. Bisher war es jedoch schwierig, vergleichbare Ergebnisse der Markenwertmessung zu erzielen, da es keine allgemein akzeptierten Standards für die Evaluierung gab und die Resultate verschiedener Untersuchungen häufig erheblich variierten. Diese Unsicherheit übertrug sich auf die Unternehmen, die solchen Evaluierungen aus den genannten Gründen eher zurückhaltend gegenüberstanden. Mit der künftigen Norm zur Markenwertmessung sollen diese Probleme gelöst werden. Der Projektantrag bei ISO beruhte auf Vorgaben, die im DIN-Normenausschuss Gebrauchstauglichkeit und Dienstleistungen in Kooperation mit dem deutschen Markenverband und zahlreichen interessierten Kreisen erbracht wurden. Anlass war die Einführung der so genannten International Accounting Standards, die unter bestimmten Voraussetzungen die Aktivierung von Markenwerten in der Bilanz ermöglichen. Die Mitwirkenden im Komitee sind Vertreter einschlägiger Dienstleistungsanbieter, Forschungseinrichtungen sowie namhafter Industrieunternehmen. Besonders die rege Beteiligung der Industrie unterstreicht die Praxisrelevanz des Themas.

ISO/PC 235 „Bonitätsbeurteilung“

Die Beurteilung der Bonität von Kreditnehmern ist seit Jahrzehnten ein umstrittenes Thema, besonders in Verbindung mit der Prüfung geeigneter Finanzierungsinstrumente in Kapitalmärkten. Seit der Einführung der neuen internationalen Eigenkapitalvorschriften („Basel II“) und mit den Plänen zu Regularien für Sicherheitsreserven von Versicherungsgesellschaften („Solvenz II“) ist Rating jedoch zu einem unabdingbaren Teil der Kreditnahme geworden. Durch die wachsende Bedeutung des Ratings wird auch seine Qualität sowie deren Definition immer relevanter. Die geplante Norm soll die Definitionen und Dienstleistungsanforderungen an professionelle Bonitätsbeurteilung bzw. Rating festlegen. Der Projektantrag basierte auf Vorarbeiten, die im Normenausschuss Gebrauchstauglichkeit und Dienstleistungen in Kooperation mit dem Bundesverband der Ratinganalysten und Ratingadviser e. V. geleistet wurden.

ISO/TC 232 „Bildungsdienstleistungen“

Wissen ist in der globalisierten Gesellschaft zu einem Schlüsselfaktor erfolgreicher Unternehmensführung geworden. Qualitätssicherung in der wirtschaftsorientierten Weiterbildung ist daher weltweit ein zentrales Thema, das eine internationale Zusammenarbeit erfordert. 34 Experten aus Europa, den USA, Asien

und Australien kamen am 26. und 27. März 2007 in Berlin zur Auftaktsitzung des neuen ISO-Komitees ISO/TC 232 „Bildungsdienstleistungen für Weiterbildung und Training“ zusammen. Im Mittelpunkt ihrer Arbeit stehen die berufliche Weiterbildung sowie innerbetriebliche Schulungen und Anforderungen an die Qualifikationsvermittlung. Die Normung wird unter anderem Aspekte der Organisation, der methodisch-didaktischen Konzeption von Bildungsmaßnahmen, Anforderungen an Lehrende, Methoden und Techniken des Wissenserwerbs und der Wissensvermittlung sowie zur Messung der Kundenzufriedenheit zum Inhalt haben. Damit soll zukünftig die Qualität von Bildungsanbietern und ihrer Angebote transparenter und international vergleichbar gemacht, der Wissenstransfer verbessert und die Nachhaltigkeit von Bildung sichergestellt werden. Das TC kann künftig auch als Basis für einschlägige bereichspezifische Normung dienen.

ISO/PC 233 „Reinigungsdienstleistungen“

Im Mittelpunkt stehen Anforderungen und Empfehlungen für Reinigungsdienstleistungen. Die spätere ISO-Norm soll Referenz für Leistungsausschreibungen sein und sowohl die Gestaltung als auch die Bewertung von Leistungsangeboten erleichtern.

Die hinter diesen Projekten stehenden interessierten Wirtschaftskreise in Deutschland kamen mit dem DIN frühzeitig überein, dass ein rein nationaler Regelungsansatz nicht sinnvoll ist, ebenso wenig wie regionale Normung, sondern dass nur internationale Festlegungen in Betracht kommen.

4 Normung in der Marktforschung: DIN ISO 20252

Die Entwicklung von brancheninternen Qualitätsstandards hin zu Qualitätsnormen steht in Deutschland in engem Zusammenhang mit der Internationalisierung von Qualitätsstandards und Normen für die Marktforschung. Zeitlich parallel mit der Entwicklung der „Standards zur Qualitätssicherung in der Markt- und Sozialforschung“ durch die deutschen Verbände hat die European Federation of Associations of Market Research Organisations (EFAMRO) europäische Qualitätsstandards für die Marktforschung entwickelt. Von EFAMRO ging auch die Initiative zur Entwicklung einer internationalen Norm für Markt-, Meinungs- und Sozialforschung aus. Wiederum zeitlich parallel mit der Beantragung dieses Normungsvorhabens bei ISO wurde in Deutschland bzw. im DIN in einem eigens dafür konstituierten Arbeitsausschuss zunächst die DIN-Norm 77500 „Markt- und Sozialforschungs-Dienstleistungen“ entwickelt. Durch Zusammenarbeit der Verbände der deutschen Markt- und Sozialforschung mit dem Deutschen Institut für Normung e.V. (DIN) ist es gelungen, die darin reflektierten Positionen der deutschen Marktforschung in die Verhandlungen zur internationalen Norm ISO

20252 erfolgreich einzubringen. DIN ISO 20252 ist als deutsche Norm übernommen worden und an die Stelle von DIN 77500 getreten.

Standards und Normen in der Markt- und Sozialforschung

Erich Wiegand

Der Artikel 1 des „IHK/ESOMAR Internationaler Kodex für die Praxis der Markt- und Sozialforschung“ (in der Fassung von 1995) fordert, dass Marktforschung immer objektiv und nach anerkannten wissenschaftlichen Grundsätzen durchgeführt werden muss. Gemäß Artikel 12 muss sie darüber hinaus kosteneffizient und von angemessener Qualität sein und wie mit dem Auftraggeber vereinbart durchgeführt werden. Qualitätsstandards bzw. Qualitätsnormen sind ein probates Mittel, um die entsprechenden Nachweise erbringen zu können.

1 Standards und Normen als Instrument der Selbstregulierung

Die Selbstregulierung der Markt- und Sozialforschung definiert die Prinzipien des berufsethischen Verhaltens (in den Berufsgrundsätzen), interpretiert die Anwendung der relevanten Gesetze und rechtlichen Bestimmungen (in den Standesregeln) und formuliert die methodischen Anforderungen an die Markt- und Sozialforschung als (angewandte) wissenschaftliche Forschung (in den Qualitätsstandards).

Den „IHK/ESOMAR Internationaler Kodex für die Praxis der Markt- und Sozialforschung“ als berufsständischen Verhaltenskodex gibt es seit 1948. Seine grundlegenden berufsethischen Verhaltensregeln bilden inzwischen in Deutschland und vielen anderen Ländern den Kern der Selbstregulierung der Markt- und Sozialforschung. Die Verbände der deutschen Markt- und Sozialforschung haben den IHK/ESOMAR Kodex mit einer vorangestellten „Erklärung für das Gebiet der Bundesrepublik“ anerkannt. In dieser Erklärung werden die in den Verhaltensregeln des IHK/ESOMAR Kodex formulierten berufsethischen Grundprinzipien der Markt- und Sozialforschung nochmals hervorgehoben.

Die Verbände der deutschen Markt- und Sozialforschung haben in Ergänzung des IHK/ESOMAR Kodex gemeinsam eine Reihe von Richtlinien herausgegeben, in denen die grundlegenden berufsethischen Verhaltensregeln des Kodex im Hinblick auf bestimmte Forschungsbereiche oder Forschungsmethoden konkretisiert werden und die Anwendung der relevanten rechtlichen Bestimmungen spezifiziert wird. Das geschah und geschieht auch in einer Reihe anderer Länder.

Für die wissenschaftlich-methodischen Anforderungen der empirischen Markt- und Sozialforschung kommt der im Jahr 2006 in Kraft getretenen und veröffentlichten internationalen Norm DIN ISO 20252 „Markt-, Meinungs- und Sozialforschung – Begriffe und Dienstleistungsanforderungen“ eine herausragende Bedeutung zu. Das kann man feststellen, ohne damit die Bedeutung der zahlreich vorliegenden nationalen Qualitätsstandards und Qualitätsnormen zu schmälern. Letztere bildeten die Grundlage der Entwicklung der DIN ISO 20252.

Selbstregulierung ohne wirksame Selbstkontrolle ist ein „zahnloser Tiger“. Deshalb haben die Verbände der deutschen Markt- und Sozialforschung – ADM, ASI, BVM und D.G.O.F. – mit dem „Rat der Deutschen Markt- und Sozialforschung e.V.“ im Jahr 2001 eine verbandsübergreifende Beschwerdestelle etabliert, die die Einhaltung der Berufsgrundsätze, Standesregeln und Qualitätsstandards sanktionsbewehrt durchsetzen kann.

2 Funktion und Nutzen von Standards und Normen

Gerade in einer Situation, in der verschiedentlich der wissenschaftliche Charakter der Marktforschung in Frage gestellt wird, können Qualitätsstandards und Qualitätsnormen erheblich dazu beitragen, ihre politische und rechtliche Anerkennung als angewandte wissenschaftliche Forschung zu sichern bzw. zu fördern. Dazu gehört auch ihr potenzieller Beitrag zur Erhaltung bzw. Verbesserung des Images der Markt- und Sozialforschung in der Öffentlichkeit und die damit verbundene Erhöhung der Teilnahmebereitschaft.

Durch Qualitätsstandards und Qualitätsnormen wird die Effizienz des Forschungsprozesses verbessert. Da die meisten Forschungsprojekte im Spannungsfeld von Kosten und Qualität durchgeführt werden, erlauben sie es den Marktforschungsinstituten damit, ihren privaten und öffentlichen Auftraggebern zu denselben Kosten eine höhere Forschungsqualität zu liefern. Zugleich verbessern Qualitätsstandards und Qualitätsnormen auch die Transparenz des Forschungsprozesses. Beides wirkt sich positiv auf das Vertrauen der Auftraggeber in die Forschungsergebnisse aus. Zudem wird der genuine und im Kern nicht substituierbare Nutzen der Markt- und Sozialforschung als Grundlage für wirtschaftliche und politische Entscheidungen verdeutlicht. Damit schärft die Markt- und Sozialforschung ihr Profil und grenzt ihre Leistungen klarer von anderen Informationsanbietern ab.

Schließlich dienen Qualitätsstandards und Qualitätsnormen auch der branchenexternen und brancheninternen „Selbstreinigung“. Branchenextern tragen sie dazu bei, dass branchenfremden Anbietern von Forschungsdienstleistungen – vor allem Unternehmensberatern und Werbeagenturen – die methodischen Anforderungen an die Markt- und Sozialforschung bewusst werden. Branchenfrem-

de Anbieter stehen dann zu Recht vor der Entscheidung, sich entweder die notwendigen Kenntnisse anzueignen und damit Teil der „market research community" zu werden oder ihre Forschungsaktivitäten wieder einzustellen.

Branchenintern schaffen Qualitätsstandards und Qualitätsnormen die Möglichkeit, die „schwarzen Schafe" der Markt- und Sozialforschung, die es wie in jeder anderen Branche gibt, wirkungsvoll zu sanktionieren. Das kann durch die Selbstreinigungskräfte des Marktes geschehen. Wenn aber das dazu notwendige Qualitätsbewusstsein der Auftraggeber allein nicht ausreicht, dann kann ein Beschwerdeverfahren beim „Rat der Deutschen Markt- und Sozialforschung e.V." für Abhilfe sorgen. Schließlich ist letztgenannter nicht nur für die Durchsetzung der Berufsgrundsätze und Standesregeln, sondern auch der Qualitätsstandards zuständig.

3 Kurze Geschichte der Standards und Normen

Von Anfang an nahm die englische Marktforschung bei der Entwicklung von Qualitätsstandards und Qualitätsnormen eine führende Rolle ein. Dabei spielte die MRQSA – die Market Research Quality Standards Association – eine entscheidende Rolle. An deren Gründung waren eine ganze Reihe nationaler englischer Verbände beteiligt, die im Bereich der Markt- und Sozialforschung im weiteren Sinne tätig sind.

Mit dem „MRQSA Service Standard for Market Research" wurde im Jahr 1996 der erste nationale Qualitätsstandard für die Markt- und Sozialforschung veröffentlicht. Vorausgegangen war ein so genanntes „Interviewer Quality Control Scheme" (zu dessen allgemeiner Funktion siehe Abschnitt 4). Zwei Jahre nach dem ersten branchenintern entwickelten Qualitätsstandard folgte im Jahr 1998 mit der Norm BSI 7911 „Specification for Organisations Conducting Market Research" die Veröffentlichung der ersten offiziellen Qualitätsnorm für die Markt- und Sozialforschung.

Die führende Rolle der englischen Marktforschung bei der Entwicklung von Qualitätsstandards und Qualitätsnormen für die Markt- und Sozialforschung setzte sich auch im anschließenden Schritt der Konformitätsbewertung, d.h. bei der Zertifizierung von Forschungsinstituten gemäß dieser Normen, fort.

In einer Reihe anderer Länder wie – als unvollständige Aufzählung – Frankreich, Italien, Japan, Spanien folgte die Entwicklung von Qualitätsstandards und Qualitätsnormen für die Markt- und Sozialforschung mehr oder weniger zeitversetzt dem englischen Beispiel. Die ersten Qualitätsstandards mit einem länderübergreifenden Gültigkeitsanspruch waren die „European Market Research Quality Standards", die von EFAMRO – der European Federation of Associations of

Market Research Organisations – entwickelt und im Jahr 1999 veröffentlicht wurden.

In Deutschland begann die Entwicklung von Qualitätsstandards für die Markt- und Sozialforschung zu einem relativ späten Zeitpunkt. Das hatte allerdings den Vorteil, von den Erfahrungen anderer Länder profitieren zu können. Damit konnte der zeitliche Vorsprung einer Reihe von Ländern im Lauf der Jahre aufgeholt werden.

Im Jahr 1999 veröffentlichten die Verbände der deutschen Markt- und Sozialforschung gemeinsam die unter der Federführung des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. entwickelten „Standards zur Qualitätssicherung in der Markt- und Sozialforschung". Als Ergänzung folgten zwei Jahre später im Jahr 2001 die „Standards zur Qualitätssicherung für Online-Befragungen". Und weitere zwei Jahr später im Jahr 2003 wurde die Norm DIN 77500 „Markt- und Sozialforschungs-Dienstleistungen" fertig gestellt.

Die Norm DIN 77500 ist entstanden als Transformation der in den Standards zur Qualitätssicherung enthaltenen methodischen Anforderungen in eine normengerechte Sprache. Mit der Veröffentlichung der internationalen Norm DIN ISO 20252 „Markt-, Meinungs- und Sozialforschung – Begriffe und Dienstleistungsanforderungen" bzw. ihrer Übernahme in das nationale Normenwerk wurde die Norm DIN 77500 als logische Konsequenz wieder zurückgezogen. Heute stellt die Norm DIN ISO 20252 in Deutschland wie in anderen Ländern die bei weitem wichtigste Qualitätsnorm in der Markt-, Meinungs- und Sozialforschung dar.

4 Allgemeines Muster der Entwicklung von Standards und Normen

Die Entwicklung von Standards und Normen in der Markt- und Sozialforschung folgt einem allgemeinen Entwicklungsmuster, das mit zeitlicher Verschiebung und unterschiedlicher Betonung der einzelnen Phasen in den meisten Ländern beobachtet werden kann. Am Anfang dieser Entwicklung steht in der Regel die Forderung nach verbindlichen Mindestanforderungen hinsichtlich Anwerbung, Schulung und Kontrolle der Interviewer. Gewöhnlich folgt dieser Forderung aber relativ schnell die Erkenntnis, dass dieser Ansatz zu kurz greift, da der gesamte Forschungsprozess als Kette verstanden werden muss, die bekanntlich nur so stark ist wie ihr schwächstes Glied.

Als Konsequenz folgt daraus in der Regel die Formulierung von Qualitätsstandards, die als Katalog von qualitätsrelevanten Maßnahmen den gesamten Forschungsprozess der Markt- und Sozialforschung umfassen. Die einzelnen qualitätsrelevanten Maßnahmen müssen dabei als Zielvorgaben formuliert werden, die

der Pluralität der Forschungsansätze gerecht werden und die Fortschritte und Innovationen in den Methoden und Techniken der Forschung nicht behindern.

Bis zu diesem Punkt der Entwicklung von Qualitätsstandards sind die Vertreter der Wirtschaftsverbände (für die Anbieter von Forschungsdienstleistungen) und der Berufsverbände (für die Forscher und teilweise auch für die Auftraggeber) der Markt- und Sozialforschung unter sich. Das ändert sich, wenn im nächsten – logischen – Schritt die branchenintern entwickelten Qualitätsstandards in eine offizielle Qualitätsnorm überführt werden. Das ist nur in enger Kooperation mit den jeweiligen nationalen Normungsinstituten möglich. Außerdem sind ab dieser Phase der Entwicklung auch branchenfremde Stakeholder wie zum Beispiel die Verbraucherorganisationen und natürlich auch die Zertifizierungsstellen involviert.

Mit der Fertigstellung der nationalen Qualitätsnormen endet zunächst das eingangs erwähnte allgemeine Entwicklungsmuster. In einer Reihe von Ländern haben die Wirtschaftsverbände der Markt- und Sozialforschungsinstitute die Zertifizierung der Institute gemäß der nationalen Qualitätsnorm zur Voraussetzung der Verbandsmitgliedschaft gemacht. In anderen Ländern, zu denen Deutschland gehört, wird diese Pflicht abgelehnt und die Entscheidung über die Zertifizierung den einzelnen Mitgliedsinstituten überlassen.

Verantwortlich für diese unterschiedlichen Positionen der Verbände hinsichtlich der Zertifizierung von Markt- und Sozialforschungsinstituten sind in erster Linie die nationalen Unterschiede im professionellen Selbstverständnis sowie die spezifischen Situationen und Traditionen der Markt- und Sozialforschung in den einzelnen Ländern. Unabhängig von der Zertifizierung sind die nationalen Qualitätsnormen allerdings auch über die jeweiligen Satzungen der nationalen Verbände verbindlich.

In jedem Fall nachhaltiger ist der „Zertifizierungsdruck" auf die Markt- und Sozialforschungsinstitute, wenn er nicht von den Verbänden, sondern vom Markt ausgeht, d.h. wenn die privaten und öffentlichen Auftraggeber die Frage der Zertifizierung von Forschungsdienstleistern zu einem Kriterium für die Auftragsvergabe machen. Auch diesbezüglich waren und sind länderspezifische Unterschiede festzustellen.

Darüber hinaus ist zukünftig von der Politik ein allgemeiner „Zertifizierungsdruck" gegenüber Dienstleistungsunternehmen zu erwarten, der auch die Markt- und Sozialforschungsinstitute betreffen wird. Ausgangspunkt ist die „Richtlinie 2006/123/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 über Dienstleistungen im Binnenmarkt", deren Anforderungen von den Mitgliedsstaaten der Europäischen Gemeinschaft bis Ende des Jahres 2009 in nationales Recht umzusetzen sind. Diese so genannte „Dienstleistungsrichtlinie" schreibt in Artikel 26 Maßnahmen zur freiwilligen Sicherung der Qualität von Dienstleistungen vor. Dabei werden an erster Stelle die Zertifizierung von Tätig-

keiten oder die Bewertung durch unabhängige oder akkreditierte Einrichtungen genannt.

Die Markt- und Sozialforschung internationalisiert sich zunehmend. Die Überführung der vorliegenden nationalen Qualitätsnormen in eine internationale Norm für die Markt- und Sozialforschung ist die logische Folge. Bei der Entwicklung dieser internationalen Qualitätsnorm treffen sich die vorübergehend getrennten Wege der nationalen Verbände der Markt- und Sozialforschung wieder. Als weitere „Player" betreten in dieser Phase der Entwicklung auch die europäische (CEN – Comité Européen de Normalisation) und vor allem die internationale (ISO – Internationale Organisation für Normung) Normungsorganisation die Bühne.

5 DIN ISO 20252 „Markt-, Meinungs- und Sozialforschung"

Es ist vornehmlich der Initiative von EFAMRO – der European Federation of Associations of Market Research Organisations – zu verdanken, dass bei ISO ein Technisches Komitee eingerichtet wurde, dessen Aufgabe die Entwicklung einer internationalen Qualitätsnorm für Dienstleistungen der Markt-, Meinungs- und Sozialforschung war. Dieses Technische Komitee wurde im Jahr 2003 unter spanischer Sekretariatsführung gegründet und bereits drei Jahre später konnte die Norm ISO 20252 „Markt-, Meinungs- und Sozialforschung – Begriffe und Dienstleistungsanforderungen" veröffentlicht werden.

Voraussetzung für diesen schnellen Erfolg war die Tatsache, dass die Entwicklung der Norm als lupenreiner „bottom up"-Ansatz konzipiert war. Die beteiligten Länder konnten ihre Dokumente und Überlegungen gleichberechtigt in einen fairen und offenen Diskussionsprozess einbringen. Dadurch wurden die unterschiedlichen Situationen und Traditionen der Markt- und Sozialforschung in den einzelnen Ländern gleichermaßen berücksichtigt, was für die globale Akzeptanz einer internationalen Qualitätsnorm von entscheidender Bedeutung ist.

Die Norm DIN ISO 20252 „Markt-, Meinungs- und Sozialforschung – Begriffe und Dienstleistungsanforderungen" ist eine alle Schritte des Forschungsprozesses umfassende und die verschiedenen Forschungsansätze berücksichtigende Qualitätsnorm, die in ihrem Anwendungsbereich den folgenden Anspruch erhebt: „Diese internationale Norm legt die Begriffe und Definitionen sowie die Anforderungen an die Dienstleistungen fest, die an Organisationen und Personen gestellt werden, die Markt-, Meinungs- und Sozialforschung betreiben."

Die Norm DIN ISO 20252 enthält sowohl organisationsspezifische als auch prozess- und produktspezifische Anforderungen. Als Dienstleistungsnorm ist sie darüber hinaus nachfragespezifisch an den Bedürfnissen und den Interessen der Auftraggeber orientiert. Allerdings muss bei letztgenanntem auch berücksichtigt

werden, dass die Markt-, Meinungs- und Sozialforschung eine auf wissenschaftlichen Zielen und Methoden basierende Dienstleistung ist, deren einzelne methodische Anforderungen nicht von allen Auftraggebern immer vollständig nachvollzogen werden können.

Die Norm DIN ISO 20252 ist in die folgenden Kapitel unterteilt, in denen (Kapitel 3 bis 7) die qualitätsrelevanten methodischen Anforderungen an die Durchführung der einzelnen Schritte des Forschungsprozesses als zu erreichende Zielvorgaben im Detail formuliert sind:

1. Anwendungsbereich,
2. Begriffe,
3. Anforderungen an das Qualitätsmanagementsystem,
4. Handhabung der konzeptionellen Forschungsaspekte,
5. Datenerhebung,
6. Datenmanagement und Datenverarbeitung und
7. Berichterstattung.

In holzschnittartiger Vereinfachung kann man feststellen, dass das dritte Kapitel auf die organisationsspezifischen Anforderungen und das vierte Kapitel auf die prozessspezifischen Anforderungen fokussiert ist. Dagegen enthalten die Kapitel fünf bis sieben vorwiegend produkt- bzw. dienstleistungsspezifische Anforderungen.

6 Gegenwärtige und zukünftige Aufgaben

Nachdem die Norm DIN ISO 20252 „Markt-, Meinungs- und Sozialforschung“ im Jahr 2006 veröffentlicht wurde, muss sie jetzt in der Profession bekannt gemacht und ihre globale Akzeptanz weiter gefördert werden. Das ist vornehmlich die Aufgabe der die Markt- und Sozialforschung repräsentierenden nationalen und internationalen Verbände.

Darüber hinaus muss der Prozess der Konformitätsbewertung, d.h. die Zertifizierung, etabliert werden. Auch hierbei kommt den nationalen und internationalen Verbänden der Markt- und Sozialforschung eine wichtige Rolle zu. Vor allem die letztgenannten sollten es außerdem als ihre Aufgabe ansehen, langfristig zu einer weitgehenden Harmonisierung der Bedingungen der Konformitätsbewertung beizutragen, damit das Zertifikat „ISO 20252 geprüft“ eine von Ländergrenzen und Zertifizierungsstellen unabhängige Bedeutung für die wissenschaftliche Qualität der Markt- und Sozialforschung erhält.

Auch wenn die Norm DIN ISO 20252 als eine umfassende Norm für die Markt- und Sozialforschung konzipiert ist, war die Profession sich länderübergreifend darin einig, eine Zusatznorm für (Online-)Access Panels in der Markt-

und Sozialforschung zu entwickeln, weil in diesem Bereich die wissenschaftliche Qualität der Markt- und Sozialforschung zunehmend in Gefahr geriet. Deshalb wurde bei ISO innerhalb des bestehenden, für die Markt-, Meinungs- und Sozialforschung zuständigen Technischen Komitees eine unter deutscher Sekretariatsführung stehende Arbeitsgruppe eingerichtet, deren Aufgabe die Entwicklung dieser Zusatznorm ist. Die Arbeiten an dieser internationalen Norm für (Online-)Access Panels in der Markt- und Sozialforschung sind weit fortgeschritten und können voraussichtlich im Jahr 2008 abgeschlossen werden.

Da die wissenschaftlichen Methoden und Techniken der Markt- und Sozialforschung ständigen Weiterentwicklungen unterliegen, kann die Norm DIN ISO 20252 nicht „in Stein gemeißelt" sein. Sie muss vielmehr diese Entwicklungen berücksichtigen und deshalb in regelmäßigen Abständen überarbeitet und aktualisiert werden. Man kann davon ausgehen, dass dies in Zeitintervallen von drei bis fünf Jahren notwendig ist. Folglich wird vermutlich im Jahr 2010 erstmals eine Revision der Norm DIN ISO 20252 anstehen.

Qualitätsstandards bei Online-Access Panels

Olaf Hofmann

1 Einleitung

Die Entwicklung und Veröffentlichung der allgemeinen Marktforschungsnorm DIN/ISO 20252 im Jahr 2006 war ein Meilenstein in der Entwicklung internationaler Qualitätsstandards. An einem Punkt jedoch hatte diese Norm eine – durch die Entstehungsgeschichte bedingte - Lücke: Dem Thema Access Panels, insbesondere Online-Access Panels, wurde zu wenig Raum geschenkt. In Relation zur Bedeutung der Online-Access Panels für die Markt- und Sozialforschung und in Relation zur Normierungsnotwendigkeit, bedingt durch qualitativ zweifelhaft arbeitende forschungsfremde Anbieter, sind die in DIN/ISO 20252 enthaltenen Passagen über das Thema Online-Access Panels eindeutig zu kurz und zu unspezifisch.

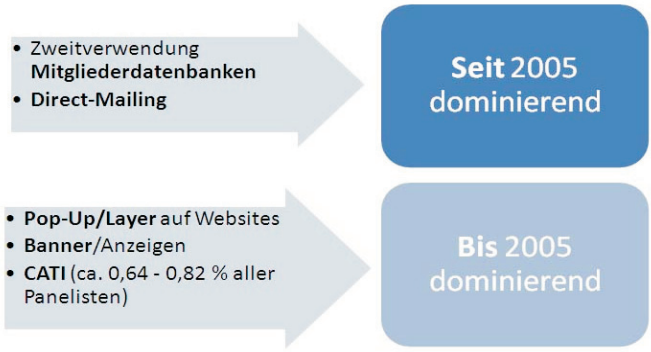
Im Jahr 2006 wurde daher entschieden, die allgemeine Marktforschungsnorm DIN/ISO 20252 durch eine spezifische Access Panel Norm zu ergänzen. Mit der Erstellung dieser neuen Norm ist zwischen Mai 2006 und ca. Juli 2008 eine sog. Working Group bei ISO betraut.

Da das Thema Qualität und Qualitätsstandards und Normierung im Zusammenhang mit Access Panels im Rahmen des Vortrages nicht erschöpfend behandelt werden kann, wird sich der Autor schwerpunktmäßig auf drei Bereiche konzentrieren (siehe Grafik 1): Rekrutierung, Stichprobenbildung und Befragungshäufigkeit. Der Grund hierfür ist, dass diese drei Bereiche a) stark qualitätsbeeinflussend sind und b) vergleichsweise gut seitens des Auftraggebers kontrolliert bzw. beeinflusst werden. Zudem sind sie in der Lage, die Vorteile, aber auch Limitierungen der angesprochenen, neuen Access Panel Norm beispielhaft zu verdeutlichen. Der Autor wird sich dabei ausschließlich dem kritischen Bereich der Online-Access Panels widmen. Offline-Access Panels (auch Mail-Panels genannt) werden hier also – anders als in der Norm selbst – nicht berücksichtigt.

Grafik 1: Themenbeschränkung im Rahmen des Vortrages

2 Rekrutierung der Befragten:

Der Rekrutierung der Befragten, genauer der Rekrutierungsquelle kommt bei der Bewertung der Voraussetzungen für Qualität bei Online-Access Panels eine entscheidende Bedeutung zu. Wie Grafik 2 zeigt, lassen sich in den letzten Jahren starke Änderungen in der Schwerpunktsetzung der Rekrutierungsquellen verzeichnen:



Grafik 2: Typische Methoden der Rekrutierung für KONSUMENTEN-Panels

Die Rekrutierung über Websites, also entweder Pop-Ups oder Layer-Technologie bzw. über Banner ist dabei so ineffizient geworden, dass sie praktisch keine Rolle mehr spielt. Teilnahmequoten bei Einsatz von Pop-ups liegen deutlich unter 1%, häufig unter 0,1%, d.h. es müssen in der Praxis ca. 1.000 Personen angesprochen werden, um einen Teilnehmer zu rekrutieren. Dies ist weder ökonomisch noch methodisch sinnvoll. Noch schlimmer sieht es derzeit bei z.B. Full Size Bannern aus. Auch das Thema CATI hat sich zumindest im Bereich Konsumenten/kommerzielle Marktforschung praktisch erledigt. Dies ist im Wesentlichen auf den Faktor hohe Kosten und geringe Effizienz zurückzuführen. Erfahrungswerte zeigen, dass maximal 5% von per CATI kontaktierten Personen auf Dauer in einem Online-Access Panel Fragebogen beantworten. Gleichwohl: Zahlreiche Studien deuten darauf hin, dass die Rekrutierung über CATI zu einer besseren Qualität führen kann, aber längst nicht muss.

Derzeit dominieren zwei Rekrutierungsquellen im Konsumentenbereich: 1. Zweitverwertung von Mitgliederdatenbanken und 2. Nutzung von Direct Mailing-Listen. Bei der Zweitverwertung von Mitgliederdatenbanken werden i.d.R. Kooperationen mit großen, reichweitenstarken Internetangeboten (z.B. Yahoo oder GMX) vereinbart, so dass deren Mitglieder in das Panel rekrutiert werden können. Bei der Beurteilung der Qualität von Mitgliederdatenbanken für ein Forschungspanel generell bzw. für ein konkretes Forschungsprojekt muss beachtet werden, welche Motive die Mitglieder der Datenbank hatten, dort Mitglied zu werden. Bei General Interest Websites ist dies häufig, jedoch längst nicht immer, unproblematisch. Bei Special Interest Websites hingegen muss man davon ausgehen, dass deren Mitglieder aus einem speziellen Interesse heraus Mitglied geworden sind. Dies kann für das Forschungsvorhaben große Nachteile haben.

Die unterschiedliche Rekrutierungsart kann einen massiven Einfluss auf Ergebnisse haben, insgesamt ist die Forschungslage jedoch nicht eindeutig:

- eine offline-Ansprache (z.B. CATI) spricht die leicht zu Hause erreichbare, „bodenständigere“ Bevölkerung an, eine online-Ansprache dagegen stärker die Heavy User des Internet,
- die Rekrutierungsart ist dabei nicht das entscheidende Kriterium, sondern das Ausmaß, zu dem bestimmte Panelistentypen durch die Rekrutierung vermittelt im Panel vorkommen, klärt die Varianz der stark unterschiedlichen Datenqualität auf.

Da die Forschungsergebnisse zur Frage der Rekrutierungsquellen weder vollständig noch eindeutig sind, musste die für die internationale Access Panel Norm verantwortliche Arbeitsgruppe eine Formulierung finden, die einerseits auf die möglichen Probleme durch die Rekrutierungsquelle hinweist, andererseits dem Kunden/Nutzer eines Access Panels die Möglichkeit an die Hand geben, das Panel hinsichtlich seiner Rekrutierungsquellen adäquat beurteilen zu können. Da der

derzeitige Entwurf des neuen Access Panel Standards noch nicht zur Veröffentlichung freigegeben ist, kann an dieser Stelle nur ein sinngemäßes Zitat veröffentlicht werden:

> *Der Panelanbieter muss die Art der Rekrutierungsquellen und ihre Größe in Relation zur gesamten Panelgröße transparent machen. Der Auftraggeber wird dadurch in die Lage versetzt, diese Informationen zu nutzen, um die Angemessenheit der Rekrutierungsquellen für das Forschungsvorhaben zu evaluieren.*

Dieser Auszug aus dem neuen ISO-Standard zu Access Panels zum Thema Rekrutierungsquellen macht zum einen deutlich, dass auch ein Qualitätsstandard kein Patentrezept sein kann zur Lösung des gesamten Problems der Rekrutierungsquellen; zum anderen macht er deutlich, wie wichtig es ist, dass sich die Nachfrager nach Paneldienstleistungen selber ein Bild machen von dem Anbieter, mit dem sie zusammenarbeiten.

3 Forschungsfremde Tätigkeiten

Da wie oben geschildert bei der Rekrutierung von Panelmitgliedern sehr häufig auf bestehende Datenbanken zurückgegriffen wird, kommt den Auswirkungen forschungsfremder Aktivitäten bei der Nutzung derselben Datenbanken eine erhebliche praktische Bedeutung zu. Dies bedeutet konkret, dass z.B. eine Mitgliederdatenbank einer Portalseite für verschiedene Nutzer offen steht; neben einem Access Panel ist dies sehr häufig auch ein Direktmarketing-Anbieter, der dieselbe Datenbank nutzt, um Massenmailings für den Vertrieb bestimmter Produkte zu versenden. Oft sind der Direktmarketing-Anbieter und der Panelanbieter nur zwei verschiedene Abteilungen derselben Firma.

Diese forschungsfremden Tätigkeiten haben bei vielen Fragestellungen erhebliche Effekte auf die mit Online-Access Panels ermittelten Ergebnisse. Ein Beispiel: Die Datenbankmitglieder erhalten ein Werbemailing zu XYZ-Notebooks. Ein Access Panel, das aus derselben Quelle rekrutiert wurde, wird danach zur Awareness von XYZ-Notebooks befragt. Die Ergebnisse werden massiv vom tatsächlichen Wert abweichen.

Da – leider – die Nutzung von Access Panel-Mitgliederdatenbanken auch für Marketing-Zwecke gerade im anglo-amerikanischen Bereich Usus ist, war es nicht möglich, eine Formulierung zu vereinbaren, die dieser Praxis einen Riegel vorgeschoben hätte. Daher wurde eine Formulierung gewählt, die auf die nationalen Datenschutzbestimmungen und Richtlinien abstellt:

> *Wenn Panelmitglieder von einer bereits existierenden Datenbank oder Kontaktliste rekrutiert werden, muss der Panelanbieter sichergehen, dass eine der-*

artige Kontaktierung mit den Datenschutzbestimmungen und den bestehenden Richtlinien übereinstimmt.

4 Aktive Rekrutierung und Auswahl versus Selbstselektion (passive Rekrutierung)

Die Unterscheidung in eine aktive Auswahl der Online-Access-Panelmitglieder versus (passive) Selbstselektion der Teilnehmer allein lässt keinen Rückschluss auf bestimmte Qualitätsmaße zu (u.a. Ausmaß Varianzaufklärung, externe Validierungsmaße etc.). Vielmehr existieren andere varianzaufklärende Variablen mit stärkerem Einfluss, wie z.B. die Dauer der Zugehörigkeit zum Panel bzw. das Ausmaß, zu dem eine materielle Teilnahmemotivation überwiegt.

Generell ist festzuhalten, dass die Stabilität des Antwortverhaltens bei aktiv rekrutierten Online-Access-Mitgliedern höher ist. Der Aufwand der Rekrutierung hat klar positive Auswirkung auf die Strukturtreue und Stabilität des Rücklaufs.

Zudem steigt das Ausmaß der Varianzaufklärung mit zunehmender Rekrutierungsqualität von passiv zu aktiv. Tendenziell ist es so, dass aktive Rekrutierung die Wahrscheinlichkeit, professionelle Befragungsteilnehmer (und die damit in Verbindung stehenden Probleme) zu generieren, verringert.

Die Forschungslage ist an diesem Punkt inkonsistent, so dass sich feststellen lässt, dass eine aktive Rekrutierung weder eine hinreichende noch eine notwendige Bedingung für eine gute Datenqualität ist.

Als Konsequenzen für Praxis ergibt sich, dass eine reine Fokussierung auf eine aktive Rekrutierung für eine hohe Ergebnisqualität nicht ausreicht, andere Qualitätskriterien wie Panelnutzung, Panelmanagement und –pflege sogar wichtiger sind.

Je höher die Anforderungen an die Differenzierungs-Schärfe der Ergebnisse ausfallen, desto höhere Anforderungen werden an die Rekrutierungsform des Panels (aktive Rekrutierung) gestellt.

Da auch hier die Forschungslage weder eindeutig noch hinreichend gut abgesichert ist, hat sich die Access Panel Working Group auf folgende Formulierung geeinigt:

Selbstselektion von Panelmitgliedern kann zu methodischen Problemen führen. Auf Nachfrage des Kunden müssen Panelanbieter transparent machen, wie sie dem Problem der Profitester bei der Rekrutierung begegnen. (Kapitel 4.3.2)

5 „Typische“ Panelpopulationen und die Implikationen in Bezug auf den Forschungsalltag

Klassifikationsversuche im Rahmen von Metaforschungsprojekten über Panelistenpopulationen haben gezeigt, dass sich Panelisten in disjunkte Gruppen einteilen lassen. Gruppen, die sich in unterschiedlichen Studien immer wieder zeigen, sind u.a. die „professional survey taker“, aber auch die loyalen Panelisten. Z.T. werden diese Gruppenzuordnungen sogar für die Stichprobenbildung herangezogen (z.B. möglichst viele „loyale“, möglichst wenig „professionelle“ Mitglieder).

Jedoch sind die Klassifikationssysteme uneinheitlich bzw. proprietär, hinzu kommt, dass die zahlenmäßige Stärke einer bestimmten Panelistengruppe sich von Panel zu Panel stark unterscheidet.

Die Metaforschung zur Panelistenmotivation zeigt, dass eine materielle Motivation bei der Mitgliedschaft zwar häufig vorhanden, aber i.d.R. nicht dominierend ist. Ein weiterer Befund ist, dass die Teilnahmemotivation nicht systematisch mit der Mitgliedschaftsdauer im Panel korreliert ist, Teilnehmer werden also nicht automatisch zu „professionellen“ Testern/Befragten.

Erstaunlicherweise und zu Beginn kontraintuitiv anmutend verhält es sich sogar so, dass Panelisten, die bei mehreren Panels Mitglied sind, a) schneller und b) „besser antworten“. „Besser antworten“ wird dabei operationalisiert über Maße wie Datendichte, Konsistenz des Antwortverhaltens, Plausibilitätsmaße und Umfang der Antworten auf offene Fragen.

Wichtig bei der Nutzung von Online-Access Panels ist es, festzuhalten, dass die breite Masse der Online-Access Panel-Mitglieder eine im Vergleich zur Allgemeinbevölkerung extrem hohe Internetnutzung aufweist. Diese liegt bei 3-5 Stunden pro Tag. Personen, die eine „normale“, d.h. durchschnittliche Internetnutzungsdauer aufweisen, wie z.B. etwas über 1h in Deutschland Anfang 2007, sind klar in der Minderheit und machen häufig nur 10-20% aller Panelisten eines Panels aus.

Dies ist ein ganz wesentlicher Grund, warum es fast immer fatal ist, Online-Access Panels im Bereich Mediennutzungsforschung oder Marketing-Effektivitätsstudien zu nutzen.

Weitere Merkmale sind extrem hohe Breitband-Nutzung (i.d.R. DSL).

Dass die typische Online-Access Panel-Population eher männlich, jung, hoch/höher gebildet ist, ist mittlerweile gut bekannt. Diesem Problem kann man aber in der Regel gut begegnen, indem Quotenstichproben gebildet werden.

6 Stichprobenziehung oder die Dominanz der Quote

Bei der Nutzung von Online-Access Panels wird in den weitaus meisten Fällen eine Quotenstichprobe gezogen (random within quota). Zufallsstichproben machen ist nur sehr selten sinnvoll, z.B. bei Metastudien zum Panel selbst im Rahmen methodologischer Studien.

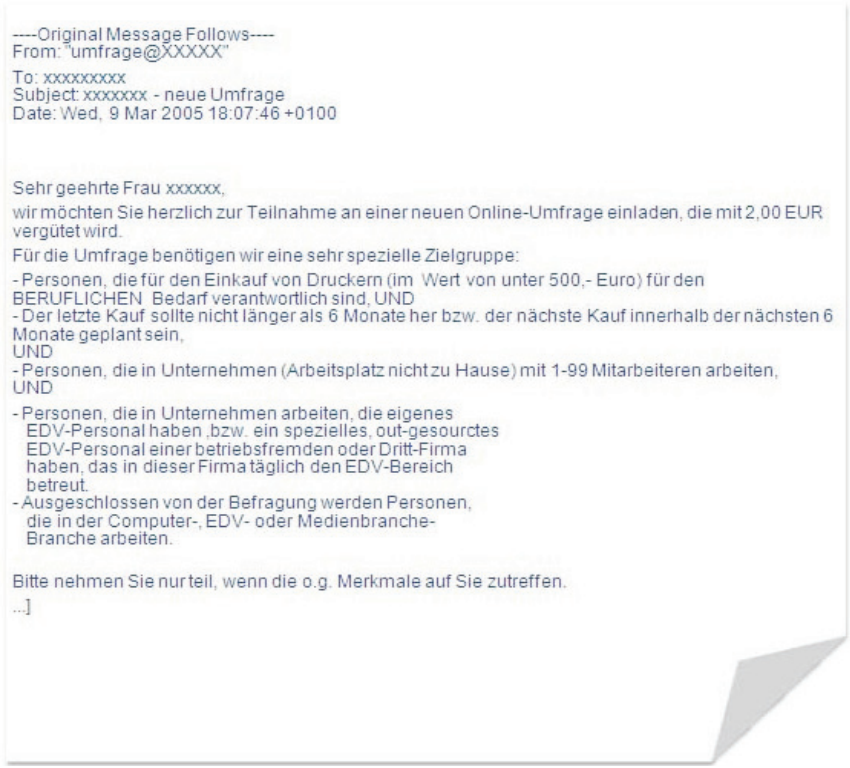
Praktisch wird die Zielverteilung bestimmter Merkmale in der Stichprobe in Entsprechung zu den bekannten Merkmalsausprägungen in der Grundgesamtheit modelliert. Dies ermöglicht in der Regel eine gute und angemessene Abbildung der Zielgruppe in der Stichprobe. Da es nicht notwendig und auch nicht möglich ist, einen verbindlichen Weg der Stichprobenbildung vorzuschreiben, hat sich die Working Group der Access Panel-Kommission entschieden, dass a) eine Kooperation mit dem Auftraggeber bezüglich der Stichprobenbildung unerlässlich ist und b) die Details der Stichprobenbildung dokumentiert werden müssen.

> *Der Anbieter des Panels sollte sich mit den Auftraggebern auf das Design und die Methode der Stichprobenziehung in einem Access Panel verständigen. Dies sollte en Detail dokumentiert werden. (4.6.1)*

7 E-Mail-Einladungen/Screenings

Da Zuverlässigkeit und Termintreue zentrale Säulen des Geschäfts bilden und es wichtig ist, bestimmte Samples rechtzeitig und vollständig zu liefern, ist es Usus, dass Online-Access Panel-Betreiber weitere Online-Access Panels beauftragen, bestimmte Zellen im Quotenplan zu schließen. Beobachtet wird dann eine Sub-Sub-Auftragsvergabe, weil auch der Subunternehmer es nicht ganz geschafft hat, die Vorgabe zu erfüllen. Der (End-)Kunde erfährt dies in der Regel nicht, schon gar nicht, wo zugekauft wurde.

Da zu beobachten ist, dass gerade Untersuchungen mit teilweise absurd niedrigen Inzidenzen (z.B. 0,2%, d.h. nur jeder 500. Befragte qualifiziert sich) in den Online-Bereich abwandern, ist leider die Konsequenz, dass einige Access Panel-Anbieter das Screening umgehen. Dies bedeutet, dass in der E-Mail-Einladung zu einer Befragung die gesuchte Stichprobe („Sehr geehrter *Herr* XY, heute suchen wir NN-Jacken-Käufer*innen* mit einem Haushaltsnettoeinkommen von über 4tsd. Euro im Monat, die berufstätig und zwischen 35 und 55 Jahren alt sind“) so exakt beschrieben wird, dass die anschließenden Screening-Fragen zur Prüfung, ob ein Access Panel-Mitglied zur Zielgruppe passt oder nicht, keine Bedeutung mehr haben. Die nächste Abbildung zeigt eine weitere – originale - E-Mail-Einladung, die demselben Zweck dient:



-----Original Message Follows----
From: "umfrage@XXXXX"

To: xxxxxxxxx
Subject: xxxxxxx - neue Umfrage
Date: Wed, 9 Mar 2005 18:07:46 +0100

Sehr geehrte Frau xxxxxx,

wir möchten Sie herzlich zur Teilnahme an einer neuen Online-Umfrage einladen, die mit 2,00 EUR vergütet wird.

Für die Umfrage benötigen wir eine sehr spezielle Zielgruppe:

- Personen, die für den Einkauf von Druckern (im Wert von unter 500,- Euro) für den BERUFLICHEN Bedarf verantwortlich sind, UND
- Der letzte Kauf sollte nicht länger als 6 Monate her bzw. der nächste Kauf innerhalb der nächsten 6 Monate geplant sein,
UND
- Personen, die in Unternehmen (Arbeitsplatz nicht zu Hause) mit 1-99 Mitarbeiteren arbeiten,
UND
- Personen, die in Unternehmen arbeiten, die eigenes EDV-Personal haben ,bzw. ein spezielles, out-gesourctes EDV-Personal einer betriebsfremden oder Dritt-Firma haben, das in dieser Firma täglich den EDV-Bereich betreut.
- Ausgeschlossen von der Befragung werden Personen, die in der Computer-, EDV- oder Medienbranche-Branche arbeiten.

Bitte nehmen Sie nur teil, wenn die o.g. Merkmale auf Sie zutreffen.

...]

Grafik 3: Beispiel für eine betrügerische Einladung zu einem Screening

8 Befragungshäufigkeit und Datenqualität bzw. Reaktanz

Der Punkt Befragungshäufigkeit ist ein vieldiskutiertes Thema, wenn es um Qualitätsstandards im Bereich Online-Access Panels geht. Auch hier gilt - wie in leider vielen anderen Qualitäts-Bereichen auch: Es fehlt derzeit die empirische Datenbasis, auf deren Grundlage mehr oder weniger exakt definiert werden könnte, ab welcher Befragungshäufigkeit so starke Reaktanzeffekte entstehen, dass weitere Befragungen die Ergebnisqualität der nächsten Forschungsprojekte unvertretbar stark beeinflussen würden.

Das Kernproblem ist, dass jede Befragung innerhalb eines Zeitraumes von weniger als ca. 6 Monaten Reaktanzeffekte aufweist. Nun ist es keinem Access Panel-Betreiber zuzumuten, dass er auf Nummer sicher geht und jeden Panelisten

maximal alle 6 Monate befragt. Zumal es Belege dafür gibt, dass der Panel Drop-out stark ansteigt, wenn ein Betreiber ein Panelmitglied im Durchschnitt sehr selten (z.B. 2-mal pro Jahr) befragt. Es ist also nicht nur ein Zielkonflikt zwischen Qualität (Vermeidung von Reaktanzeffekten durch extrem seltene Befragung) und Kosten (Möglichkeit der Amortisation von Panelinvestitionskosten durch regelmäßige Befragungen) zu unterstellen, sondern auch widerstreitende Qualitätsziele (niedriger Drop-out versus maximal niedrige Reaktanzeffekte). Daher können und sollen Kosten-Nutzen-Erwägungen durchaus eine Rolle spielen, d.h. eine häufigere Befragungsfrequenz und die damit in Verbindung stehenden Reaktanzeffeke werden – bis zu einem gewissen Maße - toleriert. Die Gretchenfrage ist jedoch: Bis zu welchem Maße?

Es existieren gerade in den USA durchaus große und stark genutzte Online-Access Panels, bei denen man als Mitglied stündlich (sic!) zu Befragungen eingeladen wird. Hier ist stark zu vermuten, dass nicht mehr der „naive", unvoreingenommene Verbraucher erreicht wird, weil der Verbraucher in vielen Bereichen durch den sehr häufigen Kontakt mit vielen neuen Produkten und Konzepten und den zugehörigen Fragen dazu eine systematisch andere Haltung entwickelt hat gegenüber z.B. einem neuen Shampoo-Konzept. Wie stark aber die Effekte sind und ob nicht vielleicht 2-3 Befragungen pro Woche diese mutmaßlich starken Reaktanzeffekte nicht/signifikant weniger hervorrufen, vermag derzeit niemand zu sagen.

Somit blieb der Online-Access Panel Working Group nichts anderes übrig, als auch hier dafür Sorge zu tragen, dass a) der Panelbetreiber die Befragungshäufigkeit dokumentieren muss (inkl. der Art der Projekte) und b) der Kunde die Aufgabe hat zu entscheiden, ob das Online-Access Panel für das in Frage kommende Projekt – zum gegebenen Zeitpunkt - geeignet ist oder nicht, bzw. ob genügend Panelisten ausgeschlossen werden können, bei denen aufgrund einer bestimmten jüngst zurückliegenden Befragungshistorie zu vermuten ist, dass Reaktanzeffekte in einem unerwünschten Umfange eintreten werden.

Entsprechend hat die Working Group formuliert:

> *Der Anbieter eines Access Panels soll Mittel einsetzen, mit denen es möglich ist zu verfolgen, welches Panelmitglied an welcher Umfrage teilgenommen hat. Diese Informationen können dann dafür genutzt werden, bestimmte Panelmitglieder auszuschließen. (4.6.2)*

9 Panelgröße

Die Panelgröße wird von Online-Access-Panelbetreibern oft unter Marketing- und Vertriebsgesichtspunkten bemüht, wenn es darum geht, das eigene Online-

Access Panel zu vermarkten. Oft wird unterstellt, dass nur „große“ Panels „repräsentative“ Samples zu generieren vermögen. Fakt ist, dass es keine direkte Beeinflussung/Beziehung zwischen der Repräsentativität eines Samples und der zu Grunde liegenden Größe gibt, schon gar nicht eine stetige/lineare Beziehung. Maximal lässt sich formulieren, dass ein für eine repräsentative Stichprobe zu kleines Panel nicht ausreichend dimensioniert ist und daher aus praktischen Gründen ungeeignet sein kann. Keinesfalls jedoch, und darauf kommt es an, darf man formulieren, dass ein großes Panel (z.B. 200tsd. Mitglieder in Deutschland) – automatisch - repräsentative bzw. „repräsentativere“ Samples liefert. Die Konsequenz für die Access Panel Working Group bestand darin zu formulieren:

> *Die Größe eines Panels ist für sich genommen kein Qualitätskriterium. Andere Faktoren wie Panelmanagement und Rekrutierung sind entscheidend für die Stichprobenqualität.*

Auf einen Qualitätsaspekt in Zusammenhang mit der Panelgröße muss jedoch hingewiesen werden: Bei der Nutzung sehr kleiner Panels (z.B. weniger als 10.000 Mitglieder pro Land) UND niedrigen Inzidenzen ist es oft notwendig, dass das genutzte Panel mindestens ein weiteres Panel zur Unterstützung hinzuzieht. Hier passiert es oft, dass die hinzugezogenen Panels nicht mehr der direkten Kontrolle des Auftraggebers unterliegen und die Stichprobe zumindest teilweise in einer Art Black Box gebildet wird. Damit in Verbindung stehend muss das Problem des sog. Paneloverlap angesprochen werden. Es existieren verschiedene Belege dafür, dass 20% aller Access Panel-Mitglieder eines Landes ca. 80% aller Access-Panel Online-Befragungen ausfüllen. Erstens ist eine bestimmte Gruppe sehr aktiv beim Ausfüllen (hohe Teilnahmequote), zweitens sind diese Mitglieder oft in mehreren Panels aktiv (im Durchschnitt 2-3 Mitgliedschaften) und drittens werden diese Mitglieder durch den oben angesprochenen Panelzukauf und durch die Mehrfachmitgliedschaft häufig mehrfach zu derselben Befragung eingeladen.

Erstaunlicherweise scheinen die Vertriebsbemühungen vieler Panelbetreiber hinsichtlich überzogener Angaben zur Panelgröße Früchte zu tragen mit dem Effekt, dass sich Panelanbieter gegenseitig übertreffen wollen mit stark übertriebenen Panelgrößen. So besaßen die Niederlande, wenn man die Mitglieder aller Panels Ende 2005 addierte und um eine empirisch ermittelte Überschneidungskonstante (Mehrfachmitgliedschaft) diskontierte, ca. 2 Millionen Panelmitglieder. Holland hatte zum gegebenen Zeitpunkt 16 Millionen Einwohner, darunter 12,5 Millionen, die in Haushalten mit Onlineanschluss lebten bzw. einen solchen z.B. am Arbeitsplatz regelmäßig nutzen konnten. Unterstellt man weiter, dass auch in Holland, wie z.B. in Deutschland, junge Männer in Access Panels deutlich überrepräsentiert sind, so ergibt sich, dass ca. jeder 4. bis 5. junge Holländer Mitglied in einem Online-Access Panel gewesen sein muss. Dies ist sicher kein 100% abgesicherter Beweis, dass die Angaben zur Panelgröße stark übertrieben sind, aber

ein Indiz dafür, dass die faktischen Mitgliedszahlen oft nur einen Bruchteil der angegebenen Größe ausmachen können.

10 Fazit

Bei Access Panels existieren sehr viele die Qualität der Stichprobe und der daraus ermittelten Daten beeinflussende Faktoren, die entweder in dieser Form oder in dieser Konstellation neu sind.

In vielen die Stichprobenqualität beeinflussenden Bereichen gibt es derzeit – angesichts sehr weniger methodologischer Grundlagenstudien - kein grundsätzlich Richtig oder Falsch, sondern eher ein „angemessenes" versus „nicht angemessenes" Panel für das konkrete Projekt.

Trotz aller aus der Sache heraus notwendigen Limitierungen der neuen Access Panel Standards helfen sie dabei, einen großen Schritt in Richtung Transparenz, Vergleichbarkeit und vor allem Potential zur Qualitätserhöhung zu gehen. Wichtig dabei ist jedoch, dass die Access Panel Standards per se keine Garantie für Qualität darstellen, dergestalt, dass man sich beispielsweise beruhigt zurücklehnen kann, sollte ein Access Panel-Anbieter beispielsweise zertifiziert sein.

Vielmehr muss jeder Kunde/Nutzer von Access Panels grundsätzlich hinterfragen, mitdenken, antizipieren und querchecken. Gerade hierfür aber sind die Access Panel Standards sehr gut geeignet: Sie sollten – aus Kundensicht - als Sammlung potenzieller Faktoren verstanden werden, die die Qualität eines über Access Panels generierten Samples beeinflussen und von Kunden aktiv mitgestaltet werden sollten.

Literatur

Gadeib, Andrea (2005): Online-Stichproben - Anspruch und Wirklichkeit. In: planung & analyse 1/2005. S. 87-92.

Gräf, Lorenz (2003): Internet Access Panels in der Praxis. In: Theobald, Axel/ Dreyer, Marcus/Starsetzki, Thomas (Hrsg.): Online-Marktforschung. Beiträge aus Wissenschaft und Praxis. Wiesbaden: Gabler. S. 115-132.

Hoppe, Michael/Lamp, Rainer (2001): Die Qualität von Online Panel - Ein Methodentest. In: planung & analyse 03/2001. S. 46-51.

Kunath, Jörg (2006): Vergleichsstudie Online-Stichproben: Macht „aktiv" oder „passiv" einen Unterschied? In: planung & analyse 5/2006. S. 63ff.

Persiel, Stefan (2005): Risiken durch Profitester in der Online-Forschung. In: planung & analyse 1/2005. S. 98-101.

Rodenhausen, Thomas/Kaufmann, Gwen/Hartmann, Adriane (2007): Der Reiz des Neuen: Teilnahmemotivation von Online-Panelisten. In: Research & Results 01/2007. S. 30f.

Vonk, Ted/van Ossenbruggen, Robert/Willems, Pieter (2006): The Effects of panel recruitment and management on research results - a study across 19 online panels. ESOMAR Panel Research. Barcelona.

Ex-ante-Qualitätskontrolle im European Social Survey

Silke I. Keil, Katja Neller

1 Einleitung: Ziele und Design des European Social Survey

Das wichtigste inhaltliche Ziel des European Social Survey (ESS) besteht darin, die Einstellungen, Wertorientierungen und Verhaltensmuster der Bevölkerungen in den europäischen Staaten zu beschreiben und zu erklären sowie Trends zu ermitteln. Um die Entwicklungen in längerfristiger Perspektive analysieren und entsprechende Theorien prüfen zu können, ist der ESS als Zeitreihe und somit als kontinuierliches Projekt angelegt. Die Erhebungen sollen möglichst regelmäßig in Zweijahresabständen stattfinden. Dabei stehen nicht die Interessen einzelner Forscher oder Forschergruppen im Vordergrund. Vielmehr soll den Sozialwissenschaften als Kollektivgut eine Datenbasis für Analysen des politischen und sozialen Wandels sowohl im Länder- als auch im Zeitvergleich geliefert werden.

Das Projekt beinhaltet repräsentative Bevölkerungsumfragen in allen beteiligten Ländern. Dabei wird ein einheitlicher Fragebogen über verschiedene Probleme des politischen und gesellschaftlichen Zusammenlebens verwendet, der durch länderspezifische Fragen sowie wechselnde Schwerpunktthemen (z.B. Lebenszufriedenheit und persönliches Wohlbefinden, Immigration oder „Citizenship, Involvement, Democracy“) ergänzt wird. Die einzelnen Länderstudien werden mit möglichst hohen methodischen Standards der empirischen Umfrageforschung durchgeführt. Die erste Welle der Befragungen fand 2002/2003 statt. Mittlerweile wurden in vielen Ländern bereits die Erhebungen für die dritte Welle abgeschlossen. Es stehen Daten aus weit über 20 Ländern zur Verfügung. Diese sind für alle Interessierten über die Website http://ess.nsd.uib.no frei zugänglich. Die DFG hat die deutsche Teilstudie des ESS mittlerweile in ihr Langfristförderprogramm aufgenommen. Damit ist die Teilnahme Deutschlands am ESS zunächst bis 2015 gesichert. Nicht nur hier zeigt sich der Erfolg des ESS. 2005 wurde das Projekt von der Europäischen Kommission mit dem René-Descartes-Preis für “excellence in scientific research” ausgezeichnet. Erstmals wurde dieser Preis einem europäischen Forscherteam in der Disziplin Sozialwissenschaften verliehen. Die Jury stellte in ihrer Begründung sowohl die gesellschaftliche Relevanz als auch die methodische Bedeutsamkeit dieses europäischen Umfrageprojektes in den Vordergrund. So sei das Besondere des ESS nicht nur sein Inhalt, sondern

auch die strengen methodischen Mindeststandards für interkulturell vergleichende Studien.

Im Rahmen dieses Beitrags werden die wichtigsten Instrumente der Ex-ante-Qualitätskontrolle im ESS näher vorgestellt. Zudem wird über einige Erfahrungen des Koordinationsteams der deutschen Teilstudie des ESS mit diesen Instrumenten berichtet.

2 Qualitätssicherung und Qualitätskontrolle als zentrale Themen des European Social Survey

Der ESS stellt alles in allem ein Modellprojekt für die europäische Sozialforschung dar, bei dem es um die Weiterentwicklung sowohl theoretischer als auch methodischer Aspekte geht. Ein wichtiger Teilbereich hiervon sind Fragen der *Qualitätssicherung* bzw. *Qualitätskontrolle*. Diese Themen spielen im ESS in allen drei wesentlichen Planungsabschnitten der Projektwellen eine zentrale Rolle: während der Periode der Vorbereitung, während der Feldphase, und auch nach Abschluss der Befragungen (vgl. Abbildung 1). Die methodischen Standards der bisher vorliegenden großangelegten international vergleichenden Studien sind sehr unterschiedlich und in nahezu keinem Fall so rigoros wie die im Rahmen des ESS definierten Qualitätskriterien.

Im Folgenden konzentrieren wir uns auf die *Ex-ante-Qualitätssicherungsinstrumente des ESS*, d.h. auf Maßnahmen vor Beginn der Feldphase, die dazu dienen sollen, die hohen methodischen Standards des ESS möglichst in allen Teilnehmerländern gleichermaßen zu erfüllen. Die zentrale Steuerungsinstanz zur Organisation der entsprechenden Prozesse im Rahmen des ESS ist das sogenannte Central Coordinating Team (CCT, vgl. Abbildung 2). Das CCT umfasst Vertreter von sechs Institutionen aus fünf verschiedenen Ländern. Durch die vom CCT und den Expertengremien des ESS ausgeübten Kontroll- und Feedbackfunktionen konnten insbesondere Länder mit einer kürzeren Tradition der international vergleichenden empirischen Sozialforschung bzw. mit bislang vergleichsweise niedrigen methodischen Standards große methodische Fortschritte erzielen.

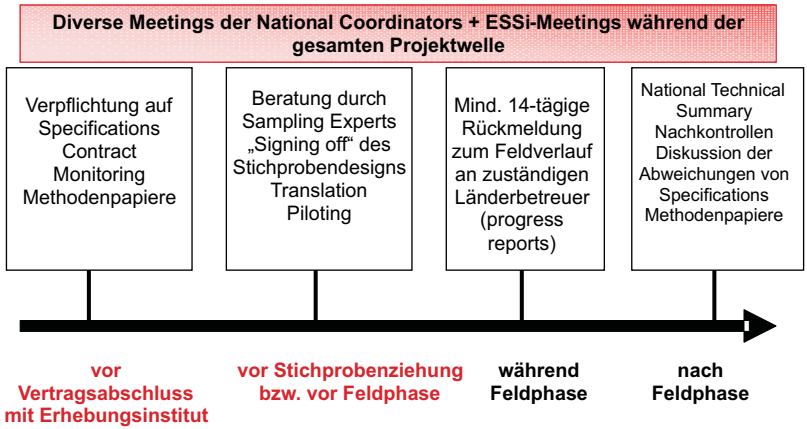


Abb. 1: Qualitätskontrolle und Qualitätssicherung in allen Projektphasen des ESS
Quelle: Eigene Darstellung. ESSi: European Social Survey Infrastructure Project.

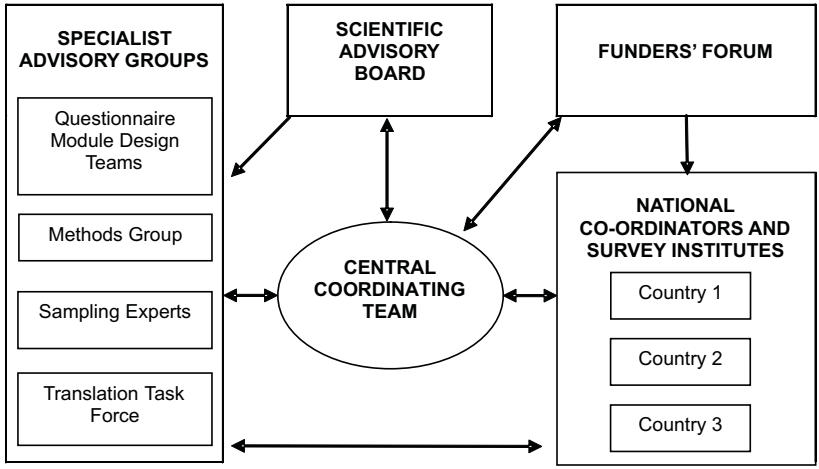


Abb. 2: Organisationsstruktur des European Social Survey
Quelle: European Social Survey, Round 1, 2002-2003: Technical Report, Chapter I: Introduction (www.europeansocialsurvey.org).

3 Die Ex-ante-Qualitätssicherungsinstrumente des European Social Survey im Überblick

3.1 Die Specification for Participating Countries

In dem Dokument „*Specification for Participating Countries*" sind die für die Gesamtstudie definierten Vorgaben und Teilnahmebedingungen für alle Teilnehmerländer des ESS zusammengefasst (vgl. www.europeansocialsurvey.org, Rubrik Project Specification). Mit diesen Spezifikationen werden vom CCT die Standards für alle beteiligten Länder gesetzt, um eine möglichst hohe Qualität des internationalen „Endproduktes" zu erreichen. Die Konzeption dieses Papiers liegt vor allem in der Zuständigkeit der internationalen Projektleitung (Roger Jowell und Team), die an der City University in Großbritannien lokalisiert ist.

Diese detaillierten, umfangreichen und kontinuierlich aktualisierten Vorgaben sind einzigartig in der Umfrageforschung. Sämtliche zu erledigenden Arbeiten, die in den einzelnen Projektphasen und in den verschiedenen Bereichen anfallen, werden durch dieses Dokument geregelt. Mit jeder neuen Erhebung wird eine aktualisierte Version der Spezifikationen erstellt und an die teilnehmenden Länder übermittelt. In diese Aktualisierungen fließen die im Rahmen der Studie gesammelten Erfahrungen ein. Diese sind das Resultat eines ständig andauernden Kommunikationsprozesses des Central Coordinating Team mit den einzelnen Ländern bzw. Ländervertretern.

Die Spezifikationen sind ein ex-ante-qualitätssicherndes Instrument, da sich alle Länder vorab vertraglich auf die Einhaltung der in den Spezifikationen schriftlich fixierten Vorgaben verpflichten müssen. Verweigert sich ein Land dieser Verpflichtung, kann es nicht am ESS teilnehmen. Diese Verpflichtung auf die Einhaltung der zentralen Richtlinien und Qualitätskriterien des ESS gilt auch für die beteiligten Umfrageinstitute. Eine Verletzung der Richtlinien zieht Sanktionsmaßnahmen seitens des CCT nach sich; eine Umgehung ist kaum oder nur in begründeten Ausnahmefällen möglich. Die internationale Leitung fordert in allen Phasen des Projekts detaillierte Berichte von den in den Ländern zuständigen Koordinatoren, die wiederum eingehend geprüft werden. Die aufwändige Dokumentationspflicht und die damit verbundene große Transparenz sichert die Einhaltung der Mindeststandards.

Die Spezifikationen beziehen sich auf alle Phasen des Projekts: auf die Vorbereitungs-, Durchführungs-, Datenprüfungs-, Datenanalyse-, Datenauswertungs- und Schlussphase. Die Richtlinien spezifizieren ferner die organisatorische Infrastruktur auf wissenschaftlicher Seite und die Auswahl eines Umfrageinstituts sowie die Benennung eines National Coordinators (NC). Die Aufgaben und Kompetenzen des NC als zentraler wichtiger Bestandteil der im Rahmen des ESS geschaffenen sozialwissenschaftlichen Infrastruktur sind genau definiert. In

Deutschland gibt es neben dem NC ein nationales Team. Neben Prof. Jan van Deth aus Mannheim als NC sind Prof. Oscar W. Gabriel aus Stuttgart, Prof. Heiner Meulemann aus Köln und Prof. Edeltraud Roller aus Mainz Mitglieder dieses Teams. Die Geschäftsführung der deutschen Teilstudie des ESS ist die Aufgabe einer wissenschaftlichen Mitarbeiterin (bis September 2006: Dr. Katja Neller; seit Oktober 2006: Dr. Silke I. Keil).

Der größte Teil der Spezifikationen bezieht sich auf die Vorbereitungs- und Durchführungsphase. So ist beispielsweise die strikte Einhaltung bestimmter Vorgaben im Hinblick auf die Stichprobenziehung ein Qualitätskriterium des ESS. Auch für den Prozess der Übersetzung des englischen Originalfragebogens in eine deutsche Fassung gibt es genaue Vorgaben. Das Gleiche gilt für die Vertragsverhandlungen und den Pretest. Diese Phasen des Projekts finden vor der Durchführung der eigentlichen Erhebung statt, können demnach explizit als Ex-Ante-Instrumente bezeichnet werden. Daher werden diese im Folgenden eingehender beleuchtet. Qualitätssichernde Maßnahmen, die während oder nach der Erhebung stattfinden, werden daher nur kurz skizziert. Zu nennen ist dabei zum einen das Event-Monitoring. Jedes Land muss dokumentieren, welche wichtigen politischen, wirtschaftlichen und gesellschaftlichen Ereignisse vor und während der Feldzeit stattgefunden haben. Es sind vor allem solche Ereignisse zu erwähnen, die Bezug zu den einzelnen Fragen haben. Dazu müssen wichtige tagesaktuelle Medien ausgewertet werden.

Zum zweiten ist die Phase der Feldarbeit geregelt: Die Feldzeit muss in jedem Land mindestens einen Monat betragen, wobei die Feldphase in den Monaten September bis Dezember stattzufinden hat. De facto ist der Erhebungszeitraum in allen beteiligten Ländern wesentlich länger, so auch in Deutschland mit rund drei Monaten. Daneben existiert eine Vielzahl weiterer von der Projektleitung vorgegebener Bedingungen, von denen nur einige exemplarisch dargelegt werden. So muss das Hauptinterview face-to-face stattfinden; mindestens vier persönliche Kontaktversuche müssen unternommen werden, wobei mindestens ein Versuch am Wochenende und ein Versuch am Abend stattfinden muss; die Kontaktverläufe müssen genau dokumentiert werden; adäquate Hilfsmittel zur Steigerung der Ausschöpfungsquote sollen eingesetzt werden etc.

Eine noch umfangreichere Vorgabenliste hat das Central Coordinating Team zum Thema Datenaufbereitung erstellt. Ein umfangreiches „Dataprotocol“ enthält detaillierte Angaben zur Benennung und Reihenfolge der einzelnen Variablen. Wenn alle Länder die angeforderten Datensätze und Dokumente bis zu einem festgesetzten Zeitpunkt geliefert haben, werden die Datensätze einer eingehenden Prüfung durch das Norwegian Social Science Data Archive (NSD) unterzogen. Das NSD garantiert die schnelle und zuverlässige Aufbereitung des internationalen Datensatzes nach höchsten Kontrollstandards und bietet auf der

eigens für den ESS konzipierten Homepage neben der Möglichkeit zum Download der Daten auch Zugang zu allen zentralen Studienmaterialien.

3.2 Das Sampling

Im Bereich des Samplings (Stichprobenziehung) werden sowohl durch die Vorgabe von methodischen Papieren mit Hinweisen zur korrekten Stichprobenziehung im Rahmen des ESS als auch durch die Arbeit des sogenannten Sampling Panels (Leitung: ZUMA, Mannheim) die Qualitätsansprüche des ESS umgesetzt.

Generell wurde für den ESS das Kriterium der Repräsentativität für Personen über 15 in Privathaushalten, ungeachtet ihrer Nationalität, Staatsbürgerschaft oder Sprache, formuliert. Die Stichprobenziehung soll in allen Teilnehmerländern mittels eines strikt randomisierten Wahrscheinlichkeitsverfahrens (probability random sampling) erfolgen. Die relativen Auswahlwahrscheinlichkeiten jedes Stichprobenfalles müssen bekannt sein und zudem in einer gesonderten Datei festgehalten werden. Auf keiner Auswahlstufe ist nach den Richtlinien des ESS ein Quotenverfahren oder eine Substitution von Individuen (weder ‚Verweigerungen' noch ‚kein Kontakt') zugelassen.

Darüber hinaus müssen von allen Teilnehmerländern bei der Stichprobenkonstruktion sogenannte Designeffekte berücksichtigt werden. Diese Designeffekte umfassen das Clustern in den Sample Points, den zu erwartenden Anteil stichprobenneutraler und nicht stichprobenneutraler Ausfälle sowie das ggf. bei der Stichprobenziehung vorgesehene Oversampling für bestimmte Gruppen. Die ESS-Vorgabe für die minimale effektive Stichprobengröße, die nach Diskontierung von Designeffekten in den Teilnehmerländern erreicht werden soll, liegt bei 1500 Fällen (bzw. bei 800 Fällen in Ländern mit weniger als zwei Millionen Einwohnern). Auf der Basis dieses Zielwerts muss in jedem Teilnehmerland vorab und mit Unterstützung der Experten des Sampling Panel die adäquate Größe der Bruttostichprobe bestimmt werden.

In Deutschland wurden diese Vorgaben mit Hilfe einer Einwohnermeldeamtsstichprobe und einer zweistufigen Stichprobenziehung umgesetzt (Auswahl von Gemeinden, Auswahl von Adressen aus den jeweiligen Einwohnermeldeamtsregistern). Die Auswahlschritte waren vorab spezifiziert. Da West- und Ostdeutschland nach wie vor als unterschiedliche Gesellschaften gelten können, war die Ziehung zweier unabhängiger, äquivalenter Stichproben in West- und Ostdeutschland erforderlich.

Die im jeweiligen Land gewählte Methode der Stichprobenziehung muss nach den Vorgaben des CCT vorab genauestens beschrieben und mit dem zuständigen Experten (das Sampling Experts Team des CCT teilt sich die Zuständigkeiten für die Teilnehmerländer dabei auf) diskutiert werden. Das Sampling Panel, eines der Expertengremien des ESS, übernimmt dabei sowohl Beratungs- und Unterstüt-

zungs- als auch Kontrollfunktionen. Ohne eine Freigabe des landesspezifischen Stichprobendesigns durch einen Vertreter des Sampling Panels („signing off“) können die Stichprobenziehungen und damit auch die Erhebungen in den einzelnen Teilnehmerländern nicht gestartet werden.

Aus Sicht des deutschen Koordinationsteams des ESS war die Unterstützung im Hinblick auf die Berücksichtigung von Designeffekten, insbesondere durch Sabine Häder von ZUMA, sehr hilfreich. Problematisch gestaltete sich jedoch in bisher allen Projektwellen das Timing des „signing-off“ der deutschen Stichprobenkonstruktion. Die damit verbundenen Diskussionen zogen sich regelmäßig bis weit in das Frühjahr des Erhebungsjahres, was mit dem internen deutschen Projektzeitplan kollidierte, da das Erhebungsinstitut für die deutsche Teilstudie des ESS mehrere Monate für die Ziehung der Einwohnermeldeamtsstichprobe benötigt.

3.3 Das Contract Monitoring

Im Rahmen des sogenannten *Contract Monitoring*, das in der Zuständigkeit von ZUMA und dem niederländischen Sociaal en Cultureel Planbureau (SCP) liegt, wird in der Phase der Vertrags- bzw. Angebotsverhandlungen kontrolliert, inwiefern die mit den Erhebungsinstituten ausgehandelte Umsetzung des ESS-Studiendesigns den Vorgaben des CCT entspricht.

Die Basis für diesen Prozess bilden Fragebögen zu den Inhalten des Angebots bzw. des Vertrags mit dem Erhebungsinstitut sowie (soweit dies angesichts der Ressourcen der nationalen Koordinationsteams des ESS möglich ist) eine englische Kurzfassung des Angebots bzw. des Vertrags mit dem Erhebungsinstitut. Der Fragebogen, der zu den Inhalten des Angebots bzw. des Vertrags mit dem ausgewählten Erhebungsinstitut ausgefüllt werden muss, umfasst unter anderem die folgenden Elemente: Art der Vertrags- bzw. Auftragsgestaltung mit dem Erhebungsinstitut, Erhebungsmodus, geplante Erhebungssprachen, geplante nationalspezifische Ergänzungen, Art des Pretests (Fallzahl, Zeitraum, Teilnehmerauswahl), Interviewer (Auswahl, Beschäftigungsstatus, Bezahlung, Schulungen), Kontaktierungsstrategien für Zielpersonen, Art der Kontaktprotokolle, Kontaktbrief, Begleitmaterial wie Infobroschüren, Incentives, Konvertierungsstrategien für Verweigerer, sonstige Maßnahmen zur Erhöhung der Ausschöpfungsquote, angestrebte Ausschöpfungs- und Non-Contact-Quote, angestrebte Zahl von Interviews, geplanter Feldzeitraum, Art und Häufigkeit der Berichterstattung während der Feldzeit, Nachkontrollen für realisierte Interviews/Verweigerer/Non-Contacts, Erhebungskosten sowie eine Projektion des Feldverlaufs (prognostizierte Zahl von Interviews pro Feldwoche).

Sofern sich im Rahmen der Abstimmung der nationalen ESS-Teams mit den zentralen Koordinierungsinstanzen des ESS Probleme bezüglich der mit den Er-

hebungsinstituten getroffenen Absprachen offenbaren, sind Anpassungen bzw. Nachverhandlungen erforderlich. Diese müssen vor der Stichprobenziehung bzw. vor Beginn der Feldphase in den jeweiligen Ländern umgesetzt werden. Solange die zentralen Anforderungen des ESS, so wie in den „Specifications" verbindlich formuliert (vgl. 3.1), sich nicht bzw. nicht in gewünschtem Maße in den Studiendesigns wiederfinden, ist kein weiterer Schritt im Rahmen der nationalen Teilstudien möglich. Gegebenenfalls kann sich so auch der geplante Feldstart beträchtlich verzögern.

In Deutschland wurden mit diesem Instrument der Qualitätskontrolle bisher neben den klaren Vorteilen einer eindeutigen Richtlinienkontrolle auch teilweise problematische Erfahrungen gemacht. Ähnlich wie beim Sampling-Prozess liegt auch hier die Schwierigkeit in den unterschiedlichen Zeitplänen des zentralen Steuerungsteams des ESS und des nationalen deutschen Koordinationsteams. Vom deutschen Koordinationsteam des ESS muss das Angebot des Erhebungsinstituts, das später Gegenstand des Contract Monitoring ist, regelmäßig bereits VOR dem Vorliegen der Projektspezifikationen für die nächste Projektwelle angefordert werden, um rechtzeitig Fördermittel der Deutschen Forschungsgemeinschaft zur Verfügung zu haben. Dies führt jedoch dazu, dass das Angebot des deutschen Erhebungsinstituts in einigen Details nicht den später formulierten Vorgaben entspricht und entsprechend in aufwändigen Nachverhandlungen angepasst werden muss. Generell ist der Effekt des Contract Monitoring jedoch sehr positiv zu bewerten, da somit bereits im Vorfeld größere Abweichungen der Teilnehmerländer von den Vorgaben bzw. methodischen Standards des ESS verhindert werden können.

3.4 Das Piloting

Im Rahmen des sogenannten *Piloting* erfolgt in Form von Pilotstudien in (i.d.R.) zwei Ländern ein großangelegter Pretest des Erhebungsinstrumentes für die jeweilige Welle des ESS. Die hierfür zuständigen Teams aus Belgien (Katholieke Universiteit Leuven) und den Niederlanden (Universiteit van Amsterdam, School of Communications Research) achten im Rahmen dieses Pretests vor allem auf die Besonderheiten der in den jeweiligen Wellen des ESS neu eingebrachten Wechselmodule des Fragebogens. Die Ergebnisse des Pretests gehen in den Prozess der weiteren Überarbeitung bzw. Finalisierung des englischen Masterfragebogens ein, der den Teilnehmerländern einige Monate vor dem geplanten Feldstart vom CCT für die Übersetzungsarbeit zur Verfügung gestellt wird.

Ein umfassender Test des Erhebungsinstrumentes für die einzelnen ESS-Wellen ist vor allem angesichts des beträchtlichen Umfangs an jeweils neuen Fragen sehr sinnvoll und hat in vielen Fällen schon zu deutlichen Verbesserungen des geplanten Fragenprogramms geführt. Dennoch bietet das Piloting als Instrument

der Ex-ante-Qualitätskontrolle im ESS einige Ansatzpunkte für Kritik. Generell problematisch, wenn auch aus Gründen der Zeitplanung nachvollziehbar, ist die Tatsache, dass eine sehr frühe Fassung des Erhebungsinstrumentes, die zum Teil sehr stark vom späteren Masterfragebogen abweicht, in den Pilotstudien getestet wird. Probleme der Fragebogenendfassung können auf diese Weise per se nicht erkannt werden.

Dies könnte durch die Erfahrungen der für jedes Teilnehmerland des ESS vorgeschriebenen nationalen Pretests ausgeglichen werden, die auf den Übersetzungen des finalen Fragebogens basieren und ein weiteres Element der ESS-Qualitätssicherung darstellen. Leider zeigt die bisherige Erfahrung jedoch, dass Rückmeldungen von Erkenntnissen, die in diesen spezifischen Pretests gewonnen werden, oder Problemen, die dort von den nationalen Teams konstatiert werden, nur in den seltensten Fällen zu Änderungen des Masterfragebogens führen. Darüber hinaus werden die Ergebnisse der Pilotstudien vom CCT kaum allgemein kommuniziert oder diskutiert. Hier könnte ein Plus an Transparenz das Gesamtergebnis sicher noch weiter optimieren.

3.5 Translation

Für den Prozess der *Übersetzung* (*translation*) wurde im Rahmen des ESS ein mehrstufiges Verfahren zur Ex-ante-Qualitätssicherung entwickelt. Zunächst werden die Wechselmodule des Fragebogens von den sogenannten Questionnaire Module Design Teams entworfen, die im Rahmen eines internationalen Wettbewerbsverfahrens in jeder ESS-Welle neu ausgewählt werden. Erste Entwürfe des Fragebogens werden dann innerhalb des CCT und weiterer ESS-Institutionen diskutiert. Im Anschluss werden diese Entwurfsfassungen mehrfach von den nationalen Koordinatoren des ESS kommentiert sowie bei internationalen Projekttreffen diskutiert. Parallel dazu erfolgt ein Pretest in ausgewählten Ländern (siehe Kapitel 3.4). Sobald der englische Masterfragebogen finalisiert wurde, beginnt der Übersetzungsprozess in den Teilnehmerländern. Dieser Prozess wird von der sogenannten Translation Task Force des ESS (Leitung: ZUMA) durch Papiere mit allgemeinen Empfehlungen (z.B. zur Auswahl geeigneter Übersetzer) sowie spezielle Beratungsangebote bei Übersetzungsproblemen unterstützt.

Die Struktur und der Prozess der Übersetzung in den teilnehmenden Ländern muss nach einem von der Projektleitung vorgegebenen Schema ablaufen, dem so genannten TRAPD-Verfahren (*T*ranslation, *R*eview; *A*djucation, *P*retesting, *D*ocumentation). Zunächst werden die Rollen der an der Übersetzung Beteiligten (translators, reviewer, adjudicator) definiert und an das CCT bzw. das zuständige Translation Team rückgemeldet. Die Übersetzung wird in einem mehrstufigen Prozess unter Einbeziehung mehrerer Übersetzer, die wiederum bestimmte Kompetenzen und Erfahrungen mitbringen müssen, durchgeführt.

In Deutschland werden die Übersetzungsrollen unter den Mitgliedern des nationalen Koordinationsteams verteilt. Darüber hinaus wird das Erhebungsinstitut in den Übersetzungsprozess eingebunden, um dem Anspruch des Austausches über unterschiedliche Interpretationen und Erfahrungen mit Frageformaten noch besser gerecht zu werden. Zur Prüfung des Erhebungsinstrumentes wird dann ein nationaler Pretest durchgeführt. Abschließend wird der Prozess der Übersetzung nach den Vorgaben des CCT dokumentiert.

Diese Phase ist sehr kommunikationsintensiv, denn während des Übersetzungsprozesses werden viele Probleme offenkundig. Nach Vorgabe des CCT soll sich jedes Land bei der Übersetzung möglichst an dem Wortlaut des Masterfragebogens orientieren. Dies ist jedoch in einigen Fällen auf Grund sprachlicher und kultureller Aspekte nicht möglich. Taucht ein solches Problem auf, sind die Verantwortlichen in den Ländern angehalten, mit dem CCT Rücksprache zu halten und gemeinsam nach einer Lösung zu suchen. Wird während des Übersetzungsprozesses ein Problem identifiziert, das sich auf alle Länder bezieht, dann wird dieses und die entsprechende Lösung von der Projektleitung an die NCs der Länder kommuniziert (sogenannte *Alert-Mails*). Diese regelmäßigen Benachrichtigungen sind ein weiteres wichtiges Instrument zur Sicherung und Optimierung der Qualität des Übersetzungsprozesses.

Die Vorgabe einer möglichst wortgetreuen Übersetzung ist vor dem Hintergrund der Vergleichbarkeit der Ergebnisse eine wichtige Bedingung. Gleichwohl kann dies zu Übersetzungen führen, die nicht immer im Einklang mit den Vorstellungen der nationalen NCs stehen. Dies soll beispielhaft an einer Formulierung aus dem ESS 3 (2006/2007) dokumentiert werden. In der dritten Welle werden eine Reihe von Fragen aus dem Bereich Gewalt/Terror formuliert, so unter anderem (Großschreibung nicht im Original, hier zur Hervorhebung): „If a MAN is suspected of planning a terrorist attack in (country), the police should have the power to keep him in prison until they are satisfied he was not involved". Die vom CCT autorisierte deutsche Übersetzung lautet folgendermaßen: „Wenn ein MANN verdächtigt wird, einen terroristischen Anschlag in Deutschland zu planen, sollte die Polizei das Recht haben, ihn zu inhaftieren, bis sie davon überzeugt ist, dass er nichts damit zu tun hat." Bei der Übersetzung war das deutsche Team zunächst davon ausgegangen, dass der englische Begriff „man" stellvertretend für das deutsche neutrale „man", also „jemand" steht. Entsprechend wurde das CCT um Bestätigung dieser Annahme gebeten. Die Projektleitung wies jedoch explizit darauf hin, dass der „Mann" intendiert sei. Daraufhin formulierte das deutsche Team einige kritische Stellungnahmen gegenüber dem CCT. Die internationale Projektleitung bestand jedoch auf ihrer Haltung.

Eine weitere Richtlinie bei der Übersetzung bezieht sich auf Länder, die über eine gemeinsame Sprache verfügen. So muss sich Deutschland im Hinblick auf seine Frageformulierungen mit allen anderen deutschsprachigen Ländern

(Schweiz, Österreich und Luxemburg) in gemeinsamen Meetings abstimmen. Diese dürfen jedoch erst dann erfolgen, wenn jedes Land bereits eine endgültige Version vorliegen hat. Nach Abschluss der Abstimmung müssen die Ergebnisse umfassend gegenüber der Projektleitung dokumentiert werden. Ziel ist die Homogenisierung der Fragebögen und eine Optimierung der Übersetzungsqualität. In der Praxis haben sich die Abstimmungsprozesse als positiv erwiesen, trotz des sehr zeitaufwändigen Verfahrens.

4 Fazit und Ausblick

Was bleibt abschließend festzuhalten?

Die Umsetzung der hohen Qualitätsansprüche des ESS ist in den meisten Teilnehmerländern gelungen. Die strikten Qualitätskontrollen und -standards dieses internationalen Surveyprojekts haben generell die Aufmerksamkeit und den Blick für methodische Probleme, insbesondere in international vergleichenden Studien, geschärft. Darüber hinaus haben vor allem Länder, in denen die moderne Umfrageforschung bisher nicht auf eine lange Tradition zurückblicken kann, im Kontext des ESS sehr schnell sehr große Fortschritte in ihren methodischen Standards erreichen können.

In Zusammenarbeit mit den zentralen ESS-Institutionen, den nationalen Erhebungsinstituten und im Austausch der Teilnehmerländer untereinander konnten einige interessante neue Ansätze, beispielsweise zur Erhöhung von Ausschöpfungsquoten, diskutiert und erarbeitet werden.

Generell gilt, dass die ESS-Qualitätsstandards von allen Beteiligten, also sowohl von den internationalen Steuerungs- und Beratungsgremien als auch den nationalen Koordinationsteams und den Erhebungsinstituten, ein sehr hohes Maß an offener und intensiver Kommunikation, Flexibilität und Transparenz erfordern. Hier mussten sich alle Akteure im Rahmen des ESS im Hinblick auf sonst übliche Verhaltensweisen, Aufgabenverteilungen oder methodische Traditionen deutlich „bewegen".

Die zahlreichen Qualitätskontroll- und –sicherungsinstrumente haben jedoch auch ihren Preis. Damit verbunden ist ein sehr hoher Aufwand, sowohl materiell als auch immateriell, der auf Dauer vermutlich nicht in der eigentlich gewünschten Art und Weise von allen bisherigen Teilnehmerländern des ESS geleistet werden kann. Bereits jetzt zeigt sich, dass nicht alle Länder an allen Wellen des ESS teilnehmen können oder wollen.

Umso wichtiger ist es, die Bemühungen zur weiteren Optimierung der im Rahmen des ESS fixierten Qualitätskriterien nicht nur fortzusetzen, sondern durch Methodenexperimente auch Ansätze zu entwickeln, wie vergleichbare Qualitätsstandards möglicherweise kostengünstiger oder weniger aufwändig erreicht wer-

den können. So hat das CCT im Rahmen der dritten Welle zusätzlich zur Standarderhebung eine Telefonstudie in fünf ausgewählten Ländern durchgeführt, darunter Deutschland (andere beteiligte Länder: französischsprachige Schweiz, Polen, Ungarn, Zypern). Ziel der Pilotstudie war die Analyse von Designeffekten für Mixed-Mode-Erhebungen (Telefon und face-to-face) und damit die Ermittlung von preiswerteren Alternativen zur face-to-face-Befragung. Der für die dritte Welle vorliegende Fragebogen wurde vom CCT für eine Telefonstudie modifiziert. Es wurden drei verschiedene Fragebogenvarianten unterschiedlicher Zeitdauer (60 Minuten, 45 Minuten und 2 x 30 Minuten) entwickelt, um die optimale Länge für die Durchführung von Telefonumfragen zu testen. Die Ergebnisse dieser Pilotstudie stehen noch aus.

Eines seiner grundlegenden Ziele, die Bereitstellung eines intensiv genutzten sozialwissenschaftlichen Kollektivgutes für Forschung und Lehre, hat der ESS bereits heute in vollem Umfang erreicht. So verzeichnete die vom Norwegischen Datenarchiv NSD konzipierte und betriebene ESS-Projekthomepage Anfang Februar 2007 über 12.323 registrierte Nutzer. Deutschland stellt davon insgesamt über zehn Prozent. Damit belegt Deutschland den Spitzenplatz in der internationalen Nutzerstatistik (vgl. Tabelle 1).

Tabelle 1: Top 10 der Länder mit den meisten registrierten Nutzern der ESS-Homepage

Land	Anzahl 2007	Anzahl 2005	Prozent der Nutzer insgesamt (Stand 2007)
Deutschland	1260	515	10,2
Großbritannien	1059	450	8,6
Norwegen	900	386	7,3
Spanien	837	251	6,8
Slowenien	758	-	6,2
USA	691	260	5,6
Belgien	640	-	5,2
Dänemark	552	247	4,5
Frankreich	471	-	3,8
Polen	466	-	3,8

Quelle: interne Nutzerstatistik des NSD, Februar 2007 und Februar 2005.

Weitere Projektinformationen:

Allgemeine Informationen: http://www.europeansocialsurvey.org

Datendownload, Onlinecodebook, Dokumentation: http://ess.nsd.uib.no

Deutsche Teilstudie des ESS: http://www.europeansocialsurvey.de

Qualitätsanforderungen der Auftraggeber an die Umfrageforschung aus Sicht der Medien- bzw. Verlagsmarktforschung

Michael Pusler

Überblick

Die folgenden Ausführungen stellen das Betätigungsfeld des (Print-) Medienforschers und hierin die wichtigsten zum Einsatz kommenden Untersuchungstypen mit den zugleich höchsten Qualitätsanforderungen vor. Im Mittelpunkt der Betrachtung stehen dabei Serviceinstrumente der Medien- bzw. Verlagsmarktforschung für die Mediaplanung (primär Printmedien) sowie die Konsortialforschung in Forschungsgremien.

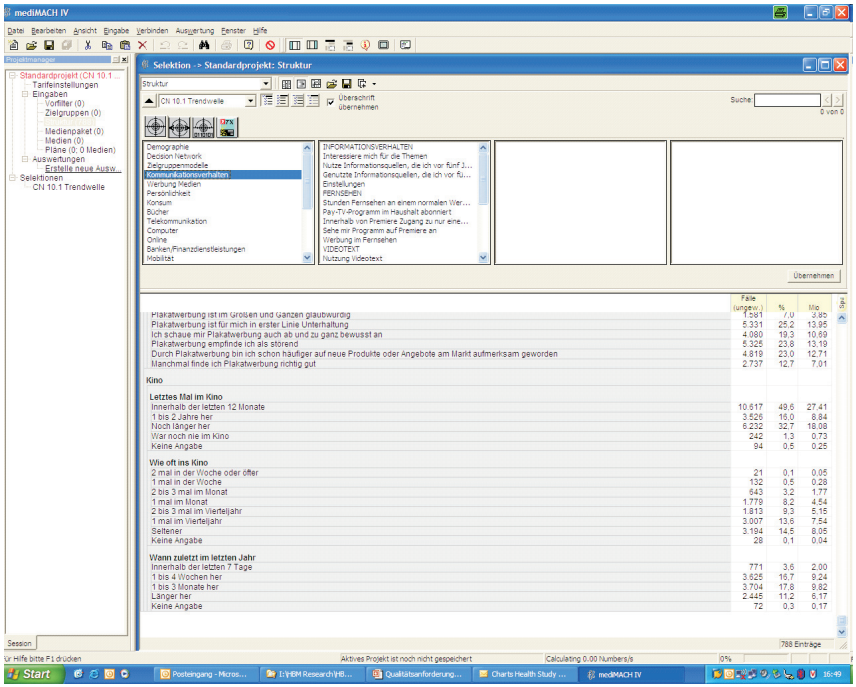
Zentrales Qualitätskriterium dabei ist neben der Konstruktion valider Fragebögen die Stichprobengüte, ergänzt um spezielle Anforderungen an die Werbeträgerforschung, insbesondere zur mediengerechten Abbildung von Werbeträgerleistung. Diese Kriterien betreffen für die Konsortialforschung in erster Line die Media-Analyse (MA) der Arbeitsgemeinschaft Media-Analyse e.V. (ag.ma). Sie sind für die nationale Mediaplanung in Deutschland – im Konsens der verschiedenen Marktpartner - über das *ZAW-Rahmenschema* geregelt. Eine europäische DIN-Norm zu Standards von Werbeträgeranalysen für Printmedien ist in Arbeit und für 2009 geplant.

Welche Studientypen werden in Großverlagen (Publikumszeitschriften) durchgeführt?

Der an erster Stelle zu nennende Studientyp in der Verlagsmarktforschung (nach Bedeutung und Kosten), und zugleich der mit den höchsten Qualitätsanforderungen, ist die *Markt-Media-Studie*. Hierbei werden für Zwecke der strategischen und operativen Mediaplanung repräsentativ gewonnene Leserschaftsdaten mit Konsumentendaten single source ermittelt.

Die hohen Qualitätsansprüche ergeben sich primär aus der Tatsache, dass mit den Daten unterschiedliche Nutzer (Verlage, Agenturen) auf der Grundlage von Zählsoftware intensiv arbeiten. Zumeist werden hiermit deskriptive Analysen auf Grundlage von Kreuztabellierungen durchgeführt. Fehler in Form von Inplausi-

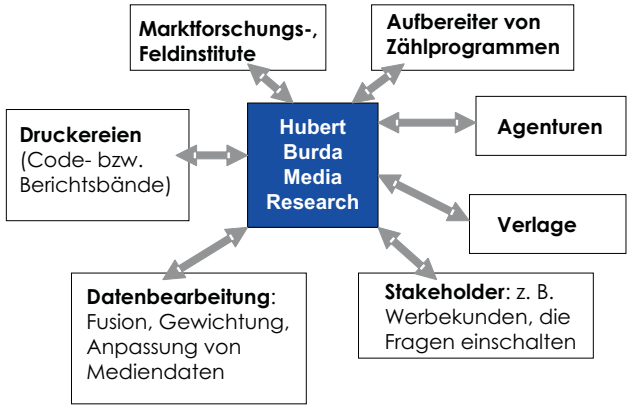
bilitäten können also leicht auffallen und müssen daher unbedingt vermieden werden.



Arbeitsgrundlage: PC-Zählsoftware

Zugleich ist dieser Studientyp der mit Abstand komplexeste, weshalb hier die Einhaltung hoher Qualitätsstandards eine ganz besonders wichtige Rolle einnimmt.

Neben der Berücksichtigung der Ansprüche der Marktpartner (Agenturen, Verlage, Stakeholder) gilt es dabei, die Arbeit der Marktforschungs- bzw. Feldinstitute zu steuern und zu kontrollieren, ebenso diejenige der EDV-Provider zur Erbringung von Zählprogrammen oder aber der Druckereien zur Erbringung gedruckter Berichtsbände.



Hohe Komplexität der Untersuchungsabläufe - Bsp. Markt-Media-Studien -

Der Vollständigkeit halber sei auch kurz erwähnt, mit was sich die Medienforschung neben den Markt-Media-Studien sonst noch beschäftigt. Weitere häufig verwendete Untersuchungstypen bzw. –instrumente zur Leserschaftsforschung sowie Redaktionsmarktforschung sind insbesondere

- der Copytest (Heftinformationen) – nach Quotenvorgabe gemäß Leserstruktur werden Heftbestandteile und Anzeigen (Posttest) im Hinblick auf Ihre Rezeptionsleistung geprüft;
- redaktionelle Umfragen – repräsentative Befragungen, zumeist zu aktuellen Themen:
- online-basierte Untersuchungen – z. B. Strukturanalysen zu Mediennutzern (Print- und Online);
- postalische Befragungen, z. B. zur Akzeptanz bei Markteintritt neuer Titel, publizistischen Sonderaktionen (eingeschränkte bis gar keine demografische Aussagekraft).

Deren Qualitätsanforderungen z. B. im Hinblick auf die Stichprobengüte erfahren aus Kosten-Nutzen-Erwägungen gewisse zulässige Einschränkungen und werden hier nicht näher behandelt.

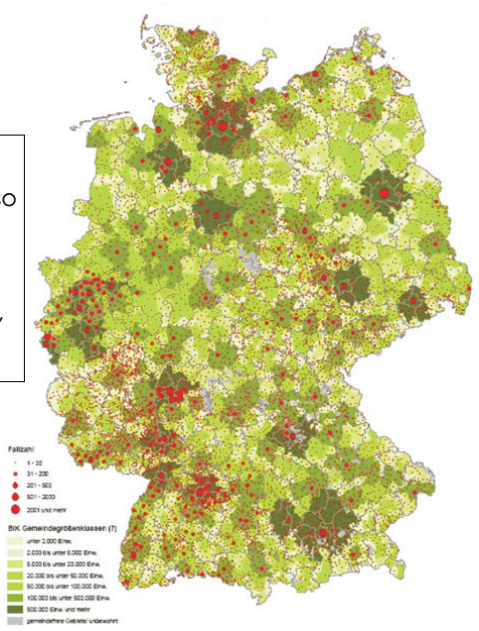
Markt-Media-Studien: hohe Qualitätsanforderungen

Bei Hubert Burda Media werden mit der Typologie der Wünsche (TdW; Hrsg. Burda Community Network) sowie Communication Networks (CN; FOCUS Magazin Verlag) zwei große Markt-Media-Studien durchgeführt (mit Fallzahlen von

jährlich 10.000 und mehr Befragten), die sich hohen Qualitätsmaßstäben verpflichtet fühlen. Das gilt selbstverständlich auch für die Verbraucheranalyse (VA; herausgegeben von den Verlagen Axel Springer und Heinrich Bauer), der nach Fallzahlen größten deutschen Markt-Media-Studie (in 2006 ca. 31.000 Fälle über einen Zeitraum von zweieinhalb Jahren).

Die in den Markt-Media-Studien erhobenen Reichweiten der Printmedien werden aus Gründen der Vergleichbarkeit an die „Leitwährung" Media Analyse (Arbeitsgemeinschaft Media-Analyse, kurz ag.ma) angepasst. Ungern möchte man bei Anpassungsprozessen die Rohdaten zu sehr strapazieren. Dies erfordert für die Stichprobenanlage ebenso wie bei der MA-Stichprobe hohe Sorgfalt. Konkret die Festlegung und Einhaltung verbindlicher Standards für die Stichprobenziehung und die Bevölkerungsgewichtung (Grundlage: Mikrozensus), um anschließend saubere Hochrechnungen zu erhalten. Schließlich möchte man bei einem sehr differenzierten Zellenplan zu hohe Personengewichte und damit mögliche Inplausibilitäten der Daten vermeiden. Eine gemäß der Bevölkerungsverteilung proportionale Stichprobenanlage ist daher anzustreben.

Die Stichprobenanlage sollte von den Auswahleinheiten her so gestaltet sein, das jede Person der Grundgesamtheit eine zumindest annähernd **gleiche und berechenbare Chance** hat, in die Stichprobe zu gelangen.



Stichprobengüte

Aus diesem (und weiteren) Gründen befolgen entsprechende Studien die Anforderungen des ZAW-Rahmenschemas für Werbeträgeranalysen (Zentralausschuss der deutschen Werbewirtschaft e. V.). Insbesondere die Wichtigkeit einer hohen Ausschöpfung unter strenger Kontrolle neutraler Ausfälle bzw. der Ausfallgründe wird hier zu Recht immer wieder betont.

Markt-Media-Studien sind zentrale Serviceinstrumente der Verlagsmarktforschung. Aufgrund der sehr hohen Kosten für Erhebung und Aufbereitung sollen sie möglichst vielen Ansprüchen bzw. Zielsetzungen gerecht werden. Da ist in erster Linie die Vielzahl an Marktinformationen (Besitz, Verwendung, Anschaffungsplanung) von der Marken- bis zur Produktebene zu nennen, die neben den Medieninformationen den Kernnutzen der Untersuchung ausmachen. Des Weiteren interessieren spezielle Zielgruppeninformationen wie z. B. das Konsum- und Mediennutzungsverhalten von Entscheidern in der Wirtschaft und der öffentlichen Verwaltung („Decision Networks" in „Communication Networks").

Studien im Vergleich: Welche Studie für welche Märkte/ZG

	CN	TdW	VA	AWA	ACTA
Medien					
Print	★★★★	★★★★	★★★★	★★★★	★★
TV	★★★	★★★	★★	★★★	★
Online	★★★	★★★	★★	★★	★★★★
Märkte					
Bücher	★★★	★	★	★★	★
Telekommunikation/ PC/ EDV	★★★★	★★	★★	★★	★★★★
Geldanlagen/ Versicherungen	★★★	★★★	★★	★★	★
Mode/ Uhren/ Kosmetik	★★	★★★	★★★★	★★	
Nahrung/ Getränke/ Gesundheit	★	★★★	★★★★	★★	
Wohnen/ Haushalt	★	★★★★	★★★★	★★★	★
Pkw/ Reisen	★★★★	★★★	★★★	★★★	★
Zielgruppen					
Konsumentenzielgruppen	★	★★★★	★★★	★★★★	★★★
Kommunikations-/Entscheidertypen	★★★★				★★★

★= geringer Umfang
★★★★= großer Umfang an Frageninhalten (eigene Einstufung des Autors)

(Studien: CN = Communication Networks (FOCUS/Burda); TdW = Typologie der Wünsche (Burda); VA = Verbraucher Analyse (Springer/Bauer); AWA = Allensbacher Werbeträgeranalyse (IfD Allensbach); ACTA = Allensbacher Computer & Telekommunikations-Analyse (IfD Allensbach).

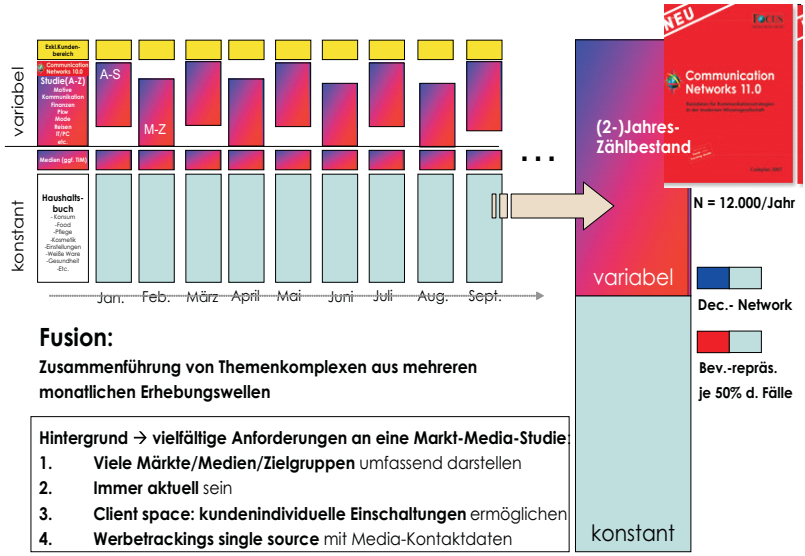
Markt-Media-Studien sollten möglichst aktuell sein und keine zeitlichen Lücken aufweisen, um nicht saisonale Veränderungen z. B. der Konsum- oder Mediennutzungsgewohnheiten zu unterschlagen. Hier hilft eine ganzjährige Erhebung. Um zeitnah mit Teilergebnissen arbeiten zu können, muss die Gesamtstudie „portionierbar" gemacht werden. D. h., es müssen kürzere Zeiträume einzeln darstellbar gemacht werden.

Als zeitlich kleinste Darstellungseinheit haben wir bei der Konstruktion von „Communication Networks" Monatsabschnitte gewählt. Da mit der Studie aber auch eine ausgabenbezogene Nutzung von Publikumszeitschriften ermittelt wird (Wochen- bis Monatstitel), macht dies zudem (zumindest nahezu) strukturgleiche Wochenstichproben erforderlich. Demnächst zieht die „Typologie der Wünsche" mit diesem Untersuchungsaufbau nach.

Diese Monatswellen dienen dann auch in der Kundenakquisition als eine Option, um individuelle Befragungswünsche realisieren zu können. Der Mehrwert hierbei ist, dass die eingeschalteten Fragen mit der ganzen Fülle der übrigen Untersuchungsinhalte zu Märkten und Verbrauchereinstellungen kombiniert und dargestellt werden können. Hierin liegt der große Unterschied z. B. zu Omnibus-Einschaltungen, die lediglich sozidemografische Informationen neben der eingeschalteten Untersuchungsfrage bereitstellen. Nicht zuletzt aus diesem Grund zieht demnächst die „Typologie der Wünsche" mit diesem Untersuchungsaufbau nach.

Somit können mit der konzipierten Untersuchungsanlage auch Werbetrackings „single source" durchgeführt werden. Denn neben den standardmäßig erhobenen und MA-angepassten Printmedien-Kontaktdaten bietet der flexible Befragungsplatz Möglichkeiten, für unsere Marktpartner wichtige Erfolgskontrollen zu Werbe- bzw. Kommunikationsmaßnahmen durchzuführen.

Schließlich machen komplexe Untersuchungen wie „Communication Networks" aufgrund der Fülle an Informationen, die aus Gründen des dann sehr hohen Befragungsdrucks gar nicht von ein und derselben befragten Person erbracht werden können, Split-Ergänzungen mittels Fusion erforderlich. Hierzu bedienen wir uns wiederum der Erfahrungswerte der Media-Analyse, die bereits seit 2003/04 aufwändige Fusionen zur Ergänzung fehlender titelbezogener Befragungsinformationen auf Befragtenebene einsetzt.



Markt-Media-Studie: Anforderungen und Methodik

Qualitätskriterien: Anforderungskatalog der Forschungskonsortien

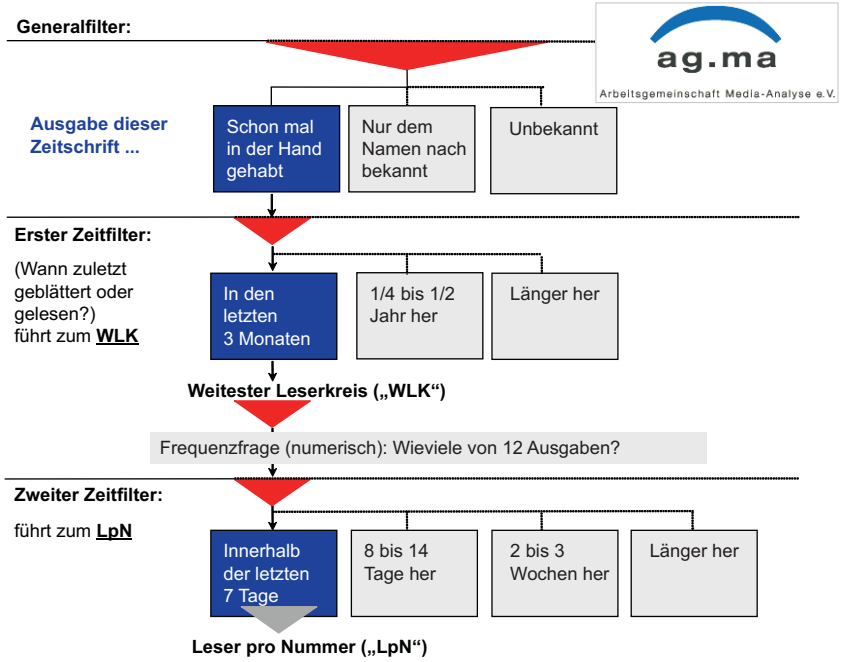
Die Medienforschung ist sehr stark durch Forschungskonsortien geprägt. Schließlich wird Forschung großteils für die „Nutzer" aus den Mediaagenturen betrieben. Die erwarten eine Transparenz der einzelnen Medien und ihrer Wirkungen zur Leistungsbeurteilung. Zugleich möchte sich hierbei, aus für den Leser sicher nachvollziehbaren Gründen, keine Mediengattung unter Wert präsentieren.

Welche Forschungskonsortien sind hier für Medienforscher relevant? Zunächst die bereits genannte ag.ma, die verschiedene Mediengruppen vereint und für die Printmedien (Publikumszeitschriften und Tageszeitungen) von zentraler Bedeutung ist. Für die ag.ma-Kommissionen stellen die Mediengruppen Print, TV, Online, Plakat, Radio Methodenexperten aus einzelnen Häusern der Mitglieder ab. Es gibt daneben noch die „Leseranalyse Entscheidungsträger in Wirtschaft und Verwaltung" (LAE), die nach dem Vorbild der ag.ma konstituiert ist, allerdings neben den Nutzern nur die in der Zielgruppe unter dem Gesichtspunkt institutioneller Anschaffungsplanung relevanten Print-Werbeträger vereint.

Darüber hinaus sind die Fernsehsender unter dem Dach der „Arbeitsgemeinschaft Fernsehforschung" (AGF) vereint, die Online-Anbieter haben sich zur „Arbeitsgemeinschaft Online-Forschung" (AGOF) zusammengeschlossen.

Was gilt es zur validen und zugleich mediengerechten Abbildung der Werbeträger- und Werbemittelleistung nun zu beachten? Die zentralen Qualitätsanforderungen an das Erhebungsinstrument selbst (für Print: zumeist an den Fragebogen) zur Abbildung von Werbeträgerreichweite und Werbemittelkontaktchance sind hierbei nicht neu und werden deshalb hier nur kurz gestreift. Sie unterliegen durch die Marktpartner einer kontinuierlichen, durchaus kritischen Prüfung. Entsprechend führt dies auch ständig zu Weiterentwicklungen. Grundsätzlich gilt: direktes Erfragen komplexer Sachverhalte wie z. B. der Erinnerung des Anteils gesehener Werbung in einer durchschnittlichen Ausgabe des FOCUS verbietet sich zumeist, denn es würde zu unsinnigen Ergebnissen führen.

Stattdessen sind modellgestützte, mehrstufige Erhebungs- und daran anschließende Verrechnungsprozeduren erforderlich. Ein Beispiel hierfür ist der Leser pro Ausgabe (LPA), die „Währung" der Printmedien-Reichweiten in der Media-Analyse (MA). Dessen Bestimmung erfolgt auf Basis eines empirischen Messwertes zum Lesen eines Titels im letzten Erscheinungsintervall (Leser pro Nummer; LPN).



Mehrstufiges Erhebungsmodell der MA für Wochentitel

Da für prognostische Zwecke aber eine durchschnittliche Ausgabe auch ex ante kalkulierbar sein soll, werden Wahrscheinlichkeiten erforderlich. Die Bildung dieser Wahrscheinlichkeiten erfolgt dann per Segmentation auf Grundlage der Frequenzangaben im weitesten Leserkreis (WLK) für den Werbeträgerkontakt, ergänzt um eine anschließende Faktorbildung für den durchschnittlichen Werbemittelkontakt (Leser pro Werbung führende Seite, LPWS) auf Basis des Werbeträgerkontakts. Hierzu sind bzw. waren im Vorfeld umfangreiche Methodenuntersuchungen erforderlich, die gemeinsam – durch Auftraggeber und Institute – konzipiert und evaluiert werden.

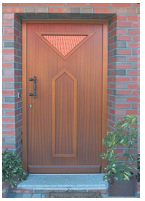
Die Anforderungen insbesondere für die medienvergleichende Konsortialforschung, aber auch an Markt-Media-Untersuchungen werden national durch das bereits erwähnte ZAW-Rahmenschema für Werbeträgeranalysen sowie Nutzerstrukturanalysen geregelt. Dieses beschreibt das zu empfehlende, teilweise sogar ein verbindliches Vorgehen sowie einzuhaltende Standards für die mediengattungsvergleichende Werbeträgerforschung. Geregelt werden so z. B. die Art der Medienabfrage, die Stichprobenbildung bis hin zur Modellbildung für den Werbeträgerkontakt oder das prinzipielle Vorgehen bei Datenfusionen. Grundgedanke dahinter ist die bereits erwähnte anzustrebende Mediengerechtigkeit, d. h. keine Mediengruppe soll im Intermediavergleich Nachteile erleiden.

Daneben beteiligen sich die Vertreter der Forschungskonsortien bzw. der Medien (neben den berufsständischen Vertretern einzelner Marktforschungsverbände) am europäischen CEN-Normungsprozess für Print-Werbeträgeranalysen für hohe Qualitätsnormen. Eine EU-weite Norm ist momentan in Vorbereitung und soll 2009 wirksam werden.

Es erscheint zwar fast trivial, der Hinweis darauf ist aber aufgrund eigener Erfahrungen in der Praxis dennoch berechtigt und wichtig: Sowohl im Bereich der Konsortialforschung, ganz besonders aber auch bei exklusiven Forschungsprojekten für einzelne Auftraggeber ist es wichtig, vom Institut, das mit der Feldforschung betraut ist, Rückmeldung über das Funktionieren der gewählten Operationalisierung für die Forschungsfragestellungen zu erhalten. Im Falle eines Nicht-Funktionierens sind die Auftraggeber neben dem Hinweis hierauf selbstverständlich auch für Verbesserungsvorschläge dankbar. Schließlich besteht für den Medienforscher im Verlag zumeist nicht die Möglichkeit, im zeitkritischen, oft kurzlebigen ad hoc-Geschäft die gewählte Operationalisierung nach allen „Regeln der Kunst“ (den in der empirischen Sozialforschung gebräuchlichen Gütekriterien) ausgiebig zu prüfen.

„Hürden der Befragung“: Qualitätsaspekte im Hinblick auf Stichprobenausfälle bzw. Ausfallgründe

Neben einer möglichst hohen Ausschöpfung ist eine Kontrolle, möglichst ein Konstanthalten, neutraler Ausfälle unabdingbar. Die Forderung nach hoher Sorgfalt bei der Kontrolle von Ausfällen ist und bleibt zentral. Mögliche Folgen einer sicher nicht leichter werdenden Befragungspraxis, systematische Ausfälle als neutrale zu deklarieren, weil z. B. Ausschöpfungsanforderungen nicht wirtschaftlich zu realisieren sind, sind aus Sicht der Auftraggeber nicht akzeptabel. Eine wichtige Forderung besteht daher im Eingrenzen der Anzahl neutraler Ausfälle auf eine realistische Größenordnung innerhalb einer definierten Feldzeit. Die wiederum darf natürlich nicht zu knapp bemessen sein, um die geforderte Ausschöpfungsrate auch erreichen zu können.



- im Haushalt niemand angetroffen
- Haushalt verweigert jede Auskunft
- Zielperson trotz mehrfacher Besuche nicht angetroffen
- Zielperson vorübergehend krank
- Zielperson verweigert das Interview

- angegebene Straße/ Haus-Nr. nicht auffindbar
- Wohnung/Untermieterwohnung unbewohnt/ Geschäft, Büro
- im Haushalt lebt keine Person der vorgegebenen Zielgruppe (z.B. Ausländerhaushalt)

➔ Neutral: Keine Relevanz für die Güte der Stichprobe

Der Anteil an den Ausfällen sollte sich aber in einem möglichst kleinen Rahmen halten!

➔ Systematisch: Relevanz für die Güte der Stichprobe

Auch eine aufwendige Stichprobe wie bei der Media-Analyse erreicht keine 100%ige Ausschöpfung (ca. 80%).

„Hürden der Befragung“: qualitätsneutrale und systematische Ausfälle

Das Ziel ist die Sicherung einer hohen Stichprobengüte über die Zeit, so dass Veränderungen der Variablen zum Untersuchungsgegenstand wirkliche inhaltliche Entwicklungen aufzeigen und eben beispielsweise nicht auf eine schwankende Stichprobenqualität zurückgeführt werden müssen. Hier stehen die Print-Medienforscher in der Pflicht gegenüber ihren Marktpartnern.

Selbstverständlich soll an dieser Stelle nicht unerwähnt bleiben, dass die Ausschöpfung nicht die einzige „Hürde" ist, die erfolgreich genommen werden muss. Zu weiteren Güteindikatoren an die Stichprobenausschöpfung sei auf die einschlägige Fachliteratur sowie auf das Informationsangebot des statistischen Bundesamtes zu Methodenfragen verwiesen.

Die Zukunft der Umfrageforschung: Herausforderungen durch (passive) technische Messverfahren

Im Vergleich mit Verfahren auf Basis technischer Messungen machen Medienforscher momentan eine wachsende Kritik an der Validität des Befragungsansatzes für Zwecke der Werbeträgerforschung aus. Der Hauptvorwurf an die Vertreter der Printmedien lautet: Personen, die mittels Vorlage von Titelkarten mit dem Zeitschriftenlogo nach ihrem Leseverhalten gefragt werden, seien doch wohl kaum in der Lage, ihr Leseverhalten bis zu einem Jahr zurückliegend zu erinnern (was so behauptet sowohl zu kurz gegriffen als auch falsch ist, aufgrund der Fülle zu betrachtender Facetten aber den Rahmen dieser Ausarbeitung sprengen würde). Eine technische Messung sei da (vermeintlich) doch wohl valider.

Technische Verfahren messen, im Gegensatz zur aktiven Befragung, passiv. In Deutschland sind sie z. B. für die Bestimmung der TV-Quoten und neuerdings für Online-Medien im Gebrauch, werden darüber hinaus aktuell aber auch für Plakate (GPS-Messung) intensiv diskutiert und zunächst versuchsweise realisiert. Aufgrund von Erfahrungen mit diesen neuesten Messsystemen in Nachbarländern (z. B. der Mediawatch in der Schweiz, die neben der Radio- auch die Zeitschriftennutzung messen können soll) erfahren sie momentan in Deutschland eine große Beachtung. Die ist wohl aber in nicht geringem Maß mit dem „Charme des technisch Machbaren" zu erklären. Deren hohe Wertschätzung halten nicht wenige Medienforscher momentan allerdings aufgrund methodischer Schwächen in der Untersuchungsanlage für unberechtigt.

Es würde an dieser Stelle zu weit führen, näher auf die einzelnen Messsysteme einzugehen. Jedes für sich wäre schon eine eigene Betrachtung wert. Die folgende grundsätzliche Bewertung konzentriert sich daher, wie bereits die vorangegangene, speziell auf die Auswahlgrundlage und Stichprobenziehung.

Was auf den ersten Blick sehr überzeugend aussehen mag, verrät bei näherem Hinsehen seine Schwächen. Sich nur auf das technisch Machbare zu fixieren, geht nämlich nicht selten zu Lasten der Qualität, insbesondere bei den realisierten Stichproben: In der Regel müssen nicht für jeden gleichermaßen leicht zu bedienende Instrumente rund um die Uhr aktiviert werden, um z. B. das Lesen von Zeitschriften lückenlos abzubilden (wie in der Studie „Mediascan" der Agentur MEDIACOM in Österreich, bei der Personen mit einem Barcode-Leser in Form

eines Handscanners ausgestattet, jede Nutzung von Zeitschriften und Zeitungen dokumentieren sollten).

Dieses und weitere Verfahren erfordern folglich sehr viel Aufwand in der Auswahl und Instruktion von Untersuchungsteilnehmern. Nicht jede Zielperson ist hierzu bereit bzw. in der Lage. Die Güte dieser „Repräsentativstichproben" ist dabei vor dem Hintergrund geforderter gleicher Auswahlchancen von Personen für Hochrechnungszwecke zur Abbildung der Grundgesamtheit kritisch zu beurteilen.

Der berechtigten Forderung der Marktpartner an Stichproben bei Mediauntersuchungen, gleiche und berechenbare Auswahlchancen für alle Personen einer Grundgesamtheit zu gewährleisten, trägt die Printmedienforschung daher seit Jahren Rechnung, sieht darin gar eine zentrale Grundlage valider und reliabeler Leserschaftsforschung.

Fazit

Der komplette Forschungsprozess, von der Fragebogenkonstruktion über die Untersuchungsanlage, Stichprobenziehung und Gewichtung unterliegt insbesondere zur Erstellung von Markt-Media-Studien (aufgrund vielfältiger Auswertungs- bzw. Analysemöglichkeiten) hohen methodischen Anforderungen. Dabei ist wiederum die Stichprobengüte essentiell für die Printmedienforschung, um in den Untersuchungen weiterhin den eigenen hohen Qualitätsansprüchen und den diesbezüglichen Forderungen der Marktpartner Rechnung zu tragen.

Eine strenge Qualitätskontrolle durch die Auftraggeber bleibt daher weiterhin wichtig, vielleicht heute sogar in einem noch stärkeren Ausmaß als in der Vergangenheit. Denn die Glaubwürdigkeit und damit die Relevanz der Ergebnisse der Medienforschung, insbesondere für die Werbeträgerforschung, steht oder fällt in Zukunft in Abhängigkeit von deren Güte.

Bewertung von Datenqualität in der amtlichen Statistik

Ulrike Timm, Andrea Kron

1 Einleitung

Voraussetzung für die Bereitstellung qualitativ hochwertiger Statistiken ist die Möglichkeit, die Qualität der Daten messen bzw. bewerten zu können. Auch der Verhaltenskodex für europäische Statistiken (European Statistics Code of Practice) hebt die Wichtigkeit der Qualitätsbewertung sowohl in allgemeiner Hinsicht als auch im Hinblick auf einzelne Qualitätskriterien hervor. Dies betrifft grundsätzlich alle Stadien der Datenproduktion: Datenerhebung, -aufbereitung, -analyse sowie Verbreitung der Statistiken.

In den vergangenen Jahren haben verschiedene Projekte auf europäischer Ebene im Bereich der Qualitätsbewertung von Statistik stattgefunden. Zu nennen sind hier beispielsweise die Arbeiten der Leadership Expert Group (LEG) on Quality sowie darauf aufbauend verschiedene nationale Projekte zu speziellen Aspekten wie Selbstbewertung, Auditierung, Nutzerbefragungen etc. Ein einheitliches System der Qualitätsbewertung statistischer Daten im Europäischen Statistischen System gibt es derzeit allerdings noch nicht. Eines dieser europäischen Projekte hat sich daher speziell diesem Thema gewidmet. Unter Mitarbeit von Kolleginnen und Kollegen aus Schweden, Norwegen, Ungarn, Portugal sowie dem Statistischen Landesamt Berlin-Brandenburg wurde unter Leitung des Statistischen Bundesamtes das Handbuch zur Bewertung von Datenqualität („Handbook on Data Quality Assessment Methods and Tools“) entwickelt.

Das Handbuch zielt auf eine systematische Implementierung von Qualitätsbewertungsprozessen in das Europäische Statistische System ab. Schwerpunkt des Handbuches und dieses Konferenzbeitrages bilden die derzeit gebräuchlichen Qualitätsbewertungsmethoden und -instrumente. Darüber hinaus werden Empfehlungen gegeben, wie diese Instrumente und Methoden in den jeweiligen Stadien der Datenproduktion einzusetzen und gegebenenfalls miteinander zu kombinieren sind. Übergeordnetes Ziel bleibt dabei, diese Methoden in möglichst harmonisierter Form im gesamten Europäischen Statistischen System einzuführen und damit dem Anspruch des Verhaltenskodex Rechnung zu tragen.

2 Bewertung von Datenqualität im Rahmen eines umfassenden Qualitätsmanagements

Die Qualitätsbemühungen in statistischen Ämtern können an verschiedenen Aspekten ansetzen. Neben organisatorischen Bemühungen, die das Management-System sowie Fragen der Führung und Unterstützung umfassen, können der Produktionsprozess, die Statistikprodukte selbst, die Anforderungen der Nutzer sowie in beschränktem Ausmaß auch das organisationale Umfeld berücksichtigt werden.

Das Handbuch fokussiert auf das Zusammenspiel von Produktionsprozessen, Statistikprodukten sowie Nutzeranforderungen, greift diesbezüglich formulierte Ziele auf, wie z. B. die Grundsätze des Code of Practice, und stellt Methoden und Instrumente zu deren Umsetzung vor.

Diese Methoden zur Qualitätsbewertung der statistischen Produkte lassen sich in drei Gruppen einteilen. Zum einen solche Methoden, die der eigentlichen Qualitätsmessung und -dokumentation dienen. Dazu zählen Qualitätsberichte und Qualitätsindikatoren, die Messung von Prozessvariablen sowie Nutzerbefragungen. Zum zweiten solche Methoden, die über die Messung hinaus die Qualität der statistischen Produkte bewerten, wie beispielsweise Selbstbewertung und Auditierung. Darüber hinaus gibt es Verfahren, in deren Rahmen die statistischen Produkte im Hinblick auf die Außenwirkung ausgezeichnet werden – entweder durch die Statistischen Ämter selbst (Labelling) oder durch externe Organisationen (Zertifizierung).

Das Handbuch stellt nicht nur die erwähnten Verfahren vor, sondern führt auch Erfahrungen auf, die statistische Ämter bisher mit diesen Methoden gesammelt haben, und gibt Umsetzungsempfehlungen. Außerdem verfügt das Handbuch über einen umfangreichen Anhang, der unter anderem vielfältige Verfahrensbeispiele aus nationalen Statistischen Ämtern enthält.

3 Qualitätsmessung und -dokumentation

3.1 Qualitätsberichte und Qualitätsindikatoren

Die Qualität statistischer Daten sollte in regelmäßigen Abständen in Berichten offengelegt werden. Im Rahmen von Gemeinschaftsstatistiken sind die Mitgliedstaaten in der Regel verpflichtet, standardisierte Qualitätsberichte an Eurostat zu liefern. Qualitätsberichte verfolgen zunächst das Ziel, die Nutzer der Daten über die Qualität der Statistiken zu informieren. Die Nutzer profitieren insofern von Qualitätsberichten, als diese ihnen bei der Einschätzung der Aussagekraft der Daten helfen. Aber auch für die Datenproduzenten kann dies von Interesse sein, da

darauf aufbauend mögliche Verbesserungen mit Blick auf jene Faktoren vorgenommen werden können, die die Qualität statistischer Daten beeinflussen. Qualitätsberichte können in Umfang und Inhalt eine sehr unterschiedliche Komplexität aufweisen. Grundbestandteil eines jeden Qualitätsberichtes sollte zunächst eine Methodenbeschreibung der zugrundeliegenden Statistik sein. Diese Beschreibung kann dann je nach gewünschtem Detaillierungsgrad ergänzt werden um qualitative Bewertungen der unterschiedlichen Qualitätskriterien oder quantitative Kennzahlen (Qualitätsindikatoren).

Sowohl Qualitätsberichte als auch Qualitätsindikatoren orientieren sich an den im „Europäische Statistiken-Verhaltenskodex" genannten Qualitätskriterien.

Tabelle 1: Qualitätsindikatoren (Standard des Europäischen Statistischen Systems)

Kriterium	Indikator
Relevanz	R1. Index der Nutzerzufriedenheit
	R2. Quote der verfügbaren Statistiken
Genauigkeit und Zuverlässigkeit	A1. Variationskoeffizient
	A2. Unit response rate (ungewichtet/gewichtet)
	A3. Item response rate (ungewichtet/gewichtet)
	A4. Imputationsquote
	A5. Quote der Übererfassung und Falschklassifikation
	A6. Quote der geografischen Untererfassung
	A7. Durchschnittliche Zahl der Revisionen
Aktualität und Pünktlichkeit	T1. Termintreue und pünktliche Veröffentlichung
	T2. Zeitspanne zwischen Referenzzeitraum und vorläufiger Veröffentlichung
	T3. Zeitspanne zwischen Referenzzeitraum und endgültiger Veröffentlichung
Zugänglichkeit und Klarheit	AC1. Anzahl der verkauften oder anderweitig verbreiteten Publikationen
	AC2. Anzahl der Zugriffe auf Datenbanken
	AC3. Quote der Vollständigkeit von Metadaten
Vergleichbarkeit	C1. Länge der vergleichbaren Zeitreihen
	C2. Anzahl der vergleichbaren Zeitreihen
	C3. Abweichungsquote von europäischen Vorgaben
	C4. Asymmetrien in Flussstatistiken
Kohärenz	CH1. Quote der Statistiken die die Anforderungen der wichtigsten Sekundärnutzung erfüllen

Damit diese gemessen werden können und dabei außerdem sowohl international als auch zwischen verschiedenen Statistiken eine einheitliche und somit vergleichbare Vorgehensweise gewählt wird, existieren im Europäischen Statistischen System zu jedem Kriterium festgelegte Indikatoren (Tabelle 1).

Die Tabelle zeigt die Vielzahl der möglichen Qualitätsindikatoren. Nicht für jede Statistik ist jeder der genannten Indikatoren sinnvoll. In Kombination mit einem ausführlichen Qualitätsbericht kann es ausreichend sein, nur einige wenige Indikatoren aufzuführen. Qualitätsindikatoren können darüber hinaus auch unabhängig von einem Qualitätsbericht als Kriterium der Qualitätsbewertung verwendet werden.

3.2 Messung von Prozessvariablen

Der Prozess der Datengewinnung und -aufbereitung wirkt sich immer auch auf die Produktqualität aus. Deshalb dürfen die Prozesse und Arbeitsabläufe der Datenerhebung bei der Bewertung der Qualität statistischer Daten nicht vernachlässigt werden. Obwohl einige der Prozessvariablen bereits in einigen Statistischen Ämtern erhoben werden, fehlt häufig eine systematische Herangehensweise.

Diesbezüglich ist es zunächst notwendig, zu Beginn die Schlüsselprozesse zu identifizieren, d. h. jene Elemente, die auf die entscheidenden Produkteigenschaften einwirken. Dies können Prozesse der Datenerhebung – wie z. B. die Antwortrate –, der Feldarbeit – wie z. B. die Interviewdauer – oder Ähnliches sein. Optimal wäre beispielsweise, Prozesse statistikübergreifend möglichst zu standardisieren, und zwar sowohl die Prozesse selbst als auch deren Dokumentation.

Um gezielt Analysen und anschließende Verbesserungen durchführen zu können, ist es zweckmäßig, den Prozess stabil zu halten, d. h. eine geringe Schwankungsbreite in der Ausprägung der Prozessvariablen zu gewährleisten. Ein Beispiel wäre eine relativ konstante Antwortrate bei regelmäßigen Durchführungen der gleichen Statistik im Zeitverlauf. Bei der Überprüfung der Prozesse sollten folgende Prozesskomponenten – gemäß dem „Europäische Statistiken-Verhaltenskodex“ – beachtet werden:

- solide Methodik,
- geeignete statistische Verfahren,
- Vermeidung einer übermäßigen Belastung der Auskunftgebenden und
- Wirtschaftlichkeit.

Bei der Einführung eines solchen Verfahrens ist es empfehlenswert, sehr vorsichtig, schrittweise und selektiv vorzugehen, da die Messung von Prozessvariablen eine schwierige und aufwändige Methode ist.

3.3 Nutzerbefragungen

Die Qualität statistischer Produkte ist nicht zuletzt im Hinblick auf die Bedürfnisse der Nutzer ein zentrales Ziel der amtlichen Statistik. Auf Grund dessen ist deren Befragung ein weiterer wichtiger Bestandteil, wenn man die Qualität statistischer Daten beurteilen möchte. Dabei ist es ratsam, nach Nutzergruppen differenzierte Befragungen durchzuführen. Zu unterscheiden sind dabei:

- allgemeine („traditionelle“) Nutzerbefragungen, die an bekannte Nutzer der Produkte und Dienste des Statistischen Amtes/Institutes gerichtet sind,
- Image-Studien, die mit unbekannten Nutzern/der Öffentlichkeit durchgeführt werden, sowie
- Zielgruppen- bzw. produktspezifische Nutzerbefragungen, zu welchen unter anderem Interviews mit Hauptnutzern, Fragebögen in Printpublikationen sowie Fragebögen für Empfänger von Pressemitteilungen zählen.

Im Rahmen dieser Studien stellt die Datenqualität eine latente Variable dar, die über manifeste Indikatoren zur Zufriedenheit mit einzelnen Bestandteilen erfasst wird. Der Befragte kann beispielsweise gebeten werden, die Zufriedenheit mit einer der Komponenten mittels einer 5er-Skala von „gar nicht zufrieden“ bis „sehr zufrieden“ einzuschätzen. Dabei empfiehlt es sich, zu beachten, dass die Einschätzung eines Aspekts durch den Befragten immer im Verhältnis zu den verschiedenen anderen Merkmalen erfolgt, da es zum einen keinen objektiven Maßstab gibt und es sich zum anderen um Wahrnehmungen handelt. Außerdem sind die meisten Nutzer keine Qualitätsexperten, so dass sie andere Schwerpunkte setzen (z. B. Zugänglichkeit oder Pünktlichkeit eher beurteilen können als Zuverlässigkeit).

Auf dem Weg zur Einführung von Nutzerbefragungen wird neben der Kategorisierung der verschiedenen Nutzer empfohlen, die richtigen Ansprechpartner innerhalb von Organisationen ausfindig zu machen und Pretests der nutzerspezifischen Fragebögen durchzuführen. Wichtig erscheint außerdem, dass die Ergebnisse der Nutzerbefragung tatsächlich zur Qualitätsverbesserung eingesetzt und erst danach erneut Nutzererhebungen durchgeführt werden.

4 Selbstbewertung und Auditierung

Die Ergebnisse der vorangegangenen Verfahren stellen die Basis für die nun folgenden Methoden der Qualitätsbeurteilung dar. Bei beiden Vorgehensweisen geht es darum, den festgestellten Ist-Zustand an einem objektiven Standard zu messen, um Möglichkeiten der Qualitätsverbesserungen aufzudecken. Um das Potenzial dieser Verfahren voll auszuschöpfen, ist es empfehlenswert, sie miteinander zu kombinieren.

4.1 Selbstbewertung

Dem Verfahren der Selbstbewertung liegt folgende Definition der European Foundation for Quality Management (EFQM) zu Grunde : „Self-assessment is a comprehensive, systematic and regular review of an organisation's activities and results referenced against a model/framework."

Je nachdem, aus welchem Blickwinkel heraus man die Selbstbewertung durchführt, variiert deren Schwerpunkt: Er kann zum einen auf der Prozessqualität, zum anderen auf der Produktqualität liegen. Andererseits ist es auch möglich, institutionelle Aspekte zu evaluieren. Das umfassendste der möglichen Selbstbewertungsverfahren umfasst das gesamte Qualitätsmanagementsystem (EFQM-Bewertung).

Möchte man eine Beurteilung des statistischen Prozesses und der daraus hervorgehenden Produkte vornehmen, kann dies unter Zuhilfenahme der „European Self-Assessment Checklist" erfolgen, die im Rahmen des „Development of a Self Assessment Programme" (DESAP) entwickelt wurde. Die Checkliste ist für die Produzenten von Statistiken entwickelt worden. Sie bringt die Verantwortlichen für die jeweilige Statistik dazu, sich mit den angewandten Produktionsprozessen und -methoden auseinanderzusetzen, diese zu bewerten und so möglicherweise Schwachstellen aufzudecken. Darüber hinaus wird empfohlen, die Ergebnisse dieser Evaluierung innerhalb der Gruppe zu diskutieren, in einem Bericht festzuhalten und als Ausgangspunkt für Verbesserungen zu wählen.

Vorteilhaft am Verfahren der Selbstbewertung ist, dass es mit geringen Kosten einhergeht, leicht umzusetzen ist und den Blick der Mitarbeiter für die qualitätsrelevanten Aspekte schärft. Um wirkliche Verbesserungen erzielen zu können, ist es zweckmäßig, Selbstbewertungen in regelmäßigen Abständen zu wiederholen. Allerdings empfiehlt es sich, die Selbstbewertungen einer externen Validierung durch unabhängige Beobachter zu unterziehen.

4.2 Auditierung

Im Rahmen der Auditierung findet eine Qualitätsbewertung durch Personen statt, die nicht in den Prozess der Erhebung und Bearbeitung der statistischen Daten involviert sind. Diese Personen können entweder aus der eigenen Organisation stammen (interne Auditierung) oder aus einer unabhängigen Institution, die mit der Überprüfung beauftragt wurde (externe Auditierung). Eine besondere Form der externen Auditierung sind Peer Reviews, in deren Rahmen die Gutachter aus anderen statistischen Ämtern stammen, die im Hinblick auf die zu überprüfenden Aspekte bereits Erfahrungen gesammelt haben. Darüber hinaus besteht auch bei diesem Schritt eine große Bandbreite hinsichtlich der Fokussierung: Es können sowohl einzelne Aspekte wie die Datenqualität oder der Prozess bewertet werden, eine Überprüfung des gesamten Systems ist ebenfalls denkbar.

Auch wenn der Erfolg von Auditierungen an recht umfangreiche Bedingungen geknüpft ist (z. B. Ergebnisse der Qualitätsmessung liegen vor; qualifizierte und anerkannte Prüfer; Einbindung der zu bewertenden Teams), stellen sie eine leistungsfähige Methode zur Bewertung und möglichen Verbesserung der Qualität in der gesamten Organisation dar.

5 Labelling und Zertifizierung

Im Rahmen von Labelling und Zertifizierung geht es um die Kommunikation der Ergebnisse der Qualitätsbewertung nach innen und außen. Dazu werden Gütezeichen entweder von der Organisation selbst oder von externen Gutachtern vergeben, wenn die untersuchten Aspekte bestimmte Qualitätsanforderungen erfüllen.

5.1 Labelling

Der Grundgedanke des Labellings ist es, den verschiedenen Nutzergruppen auf einen Blick zu signalisieren, ob die statistischen Produkte den ausgewiesenen Qualitätskriterien genügen. Die Verwendung von Gütezeichen („Labels“) soll zugleich die Transparenz als auch das Vertrauen in die Daten steigern. Ein entscheidender Vorteil von Labels für externe Nutzer ist, das sie auch für Laien sofort verständlich sind. Neben diesen externen Wirkungen wird erwartet, dass es die Mitarbeiter Statistischer Ämter motiviert, wenn es darum geht, dass die in ihren Arbeitsbereich fallenden Produkte mit einem Label für hohe Qualität ausgezeichnet werden.

Darüber hinaus unterscheiden sich die Etikettierungen bezüglich des Anforderungsgrades an das auszuzeichnende Objekt: Das Statistische Amt Neuseelands beispielsweise zeichnet nur Produkte aus, die besonders hoch gesteckte Kriterien erfüllen (so genannte „Tier 1 statistics“). Im Gegensatz dazu werden in Schweden Produkte etikettiert, die ein bestimmtes Minimum an Qualität erfüllen. Sie werden als „Official Statistics“ gekennzeichnet.

Hinsichtlich der Verwendung von Qualitäts-Labels ist allerdings Folgendes zu berücksichtigen: Zwar ermöglicht das Labelling, die Ergebnisse der Qualitätsbewertung klar und auch für Laien verständlich zusammenzufassen, der Informationsgehalt eines Labels ist jedoch als recht niedrig einzustufen, da über die Informationsverdichtung die Detailschärfe verloren geht.

5.2 Zertifizierung nach ISO 20252

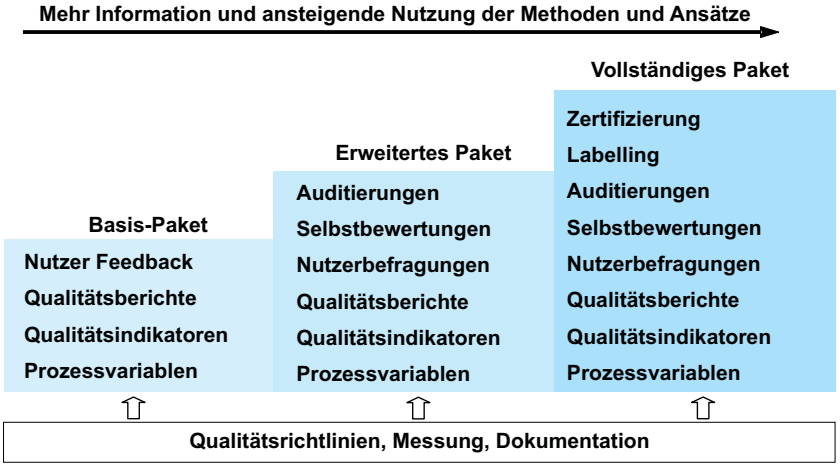
Eine Zertifizierung, beispielsweise nach ISO 20252, stellt durch die Prüfung anhand externer Kriterien der ISO-Norm eine „strenge“ Variante des Labelling dar. Zu weiteren Informationen zu DIN ISO 20252 sei auf den Tagungsbeitrag Herrn

Dr. Mühlbauers „Entstehung einer Norm: DIN ISO 20252 Markt-, Meinungs- und Sozialforschung“ in dieser Publikation verwiesen.

Bei der Anwendung der Norm sollten generell zunächst die bestehenden Qualitätssysteme mit der ISO 20252 verglichen werden, um darauf aufbauend den notwendigen Veränderungsbedarf im Qualitätssystem feststellen zu können und die entsprechenden Veränderungen vorzunehmen. Dazu zählen je nach Ausprägung des Qualitätssystems die Anpassung der Dokumentationsprozesse, Veränderungen im Arbeitsablauf sowie interne Auditierungen, um die korrekte Umsetzung der Qualitätsstandards zu überprüfen. Erst im folgenden Schritt sollte eine externe Organisation mit der Qualitätsbewertung und abschließenden Zertifizierung beauftragt werden.

6 Auf dem Weg zu einer Umsetzungsstrategie

Angesichts der Vielzahl der vorgestellten Methoden und Instrumente stellt sich die Frage, auf welche Weise diese am besten zu implementieren sind. Dies gilt insbesondere für solche statistischen Organisationen, in denen bislang ein System zur Bewertung der Datenqualität fehlt. Generell empfiehlt sich eine schrittweise Umsetzung der Methoden. Am Ende dieser Entwicklung sollte eine Kombination der Methoden der Qualitätsmessung und -bewertung stehen, in einer Weise, dass diese sich gegenseitig ergänzen und unterstützen. Da es nicht sinnvoll sein kann, alle Methoden gleichzeitig einzuführen, empfiehlt es sich, aufeinander folgende Pakete (siehe Grafik 1) zu implementieren, die ein jeweils unterschiedlich hohes Niveau der Qualitätsbewertung erzielen.



Grafik 1: Stufen der Implementierung einer umfassenden Qualitätsbewertung

Bei der Realisierung der „Pakete“ sollte auf jenen Bestandteilen aufgebaut werden, die bereits existieren. Das bedeutet, zunächst elementare Methoden der Qualitätsmessung und -dokumentation umzusetzen. So sind beispielsweise mit Blick auf das Nutzer-Feedback Online-Befragungen denkbar, bezüglich der Qualitätsberichte ist eine Beschränkung auf Kernelemente möglich. Ratsam ist die vorläufige Beschränkung auf diese Verfahren deshalb, weil deren solide Ausführung die Basis für weitergehende Methoden ist.

Darauf aufbauend ist es empfehlenswert, Komplexität und Detaillierungsgrad der bereits eingeführten Verfahren kontinuierlich zu steigern sowie weitere Methoden hinzuzunehmen. Denkbar ist, dass das vollständige Paket zunächst nur auf einige Statistikprodukte angewandt wird.

7 Schlussbemerkung

Im Rahmen dieses Beitrags wurden Methoden zur Qualitätsmessung, -dokumentation und -bewertung vorgestellt, wie sie das auf europäischer Ebene entwickelte „Handbook on Data Quality Assessment Methods and Tools“ enthält. Dieses Handbuch, das primär darauf abzielt, den nationalen statistischen Instituten eine Orientierung bei der Umsetzung ihrer Qualitätsziele zu geben, steht ab Mitte September 2007 allen Interessierten auf der Homepage von Eurostat unter dem Stichwort „Qualität“ zur Verfügung.

Damit die Qualitätsanstrengungen erfolgreich sind, ist es unabdingbar, dass sich die gesamte Organisation zur Relevanz von Qualität bekennt. Dieser Gedanke sollte nicht zuletzt auch von der Leitungsebene getragen werden. Im Idealfall kommuniziert sie die entsprechenden Ziele und Notwendigkeiten und unterstützt die Umsetzung der erforderlichen Verfahren. Wie dieser Beitrag und insbesondere das Kapitel über eine Umsetzungsstrategie verdeutlichen, ist die Beurteilung der Datenqualität ein langer Prozess. Neben Impulsen zu dessen Anstoß erscheint es angebracht, darauf zu achten, dass dieser auch am Laufen gehalten wird. Deshalb empfiehlt es sich, für die Qualitätsbewertung umfassende Ressourcen bereitzustellen und die dabei erzielten Erfolge extern wie intern zu kommunizieren.

Der Verhaltenskodex des Europäischen Statistischen Systems

Martina Hahn[1]

Abstract

Ist schon auf nationaler Ebene die Qualität von statistischen Daten ein schwer zu fassendes Konzept, wie viel schwieriger wird es erst für den Nutzer, die Qualität europäischer Statistiken einzuschätzen, werden doch der Mehrdimensionalität und Relativität des Qualitätsbegriffes weitere Parameter hinzugefügt. Mit der Annahme des Verhaltenskodex für Produzenten der Europäischen Statistiken (Kommission 2005) hat sich das Europäische Statistische System (ESS) einen gemeinsamen Qualitätsrahmen gesteckt, der sowohl einen systemweiten Ansatz im Sinne einer konsistenten und ganzheitlichen Qualitätssicherung fördert als auch den Nutzern der Europäischen Statistik eine wichtige Orientierungshilfe bei der Beurteilung der Datenqualität und Vertrauenswürdigkeit der Statistikproduzenten sein kann.

Dieses Papier geht kurz auf den Inhalt des Verhaltenskodex ein und beschreibt den vom Europäischen Statistischen System bisher eingeschlagenen Weg bei seiner Umsetzung. Dazu werden einige Entwicklungen des ESS zur Qualitätssicherung auf Produkt- u. Prozessebene kurz umrissen.

Ein Absatz über Ergebnisse der von den Nationalen Statistikämtern und Eurostat durchgeführten Selbsteinschätzungen bzgl. der Übereinstimmung ihrer Praktiken mit dem Verhaltenskodex und der sog. „Peer Reviews", die gegenwärtig ESS-weit durchgeführt werden, versucht eine erste Bestandsaufnahme. Weitere Schritte der im Wesentlichen auf Selbstregulierung der beteiligten Institutionen setzenden Umsetzungsstrategie sehen die Definition von Maßnahmenplänen und die Einbeziehung aller an der Produktion der Europäischen Statistiken beteiligten Stellen vor.

Abschließend werden Optionen für die Schaffung eines Qualitätsgütesiegels für europäische Aggregate diskutiert.

1 Die in diesem Papier vertretene Meinung ist die der Autorin und nicht notwendigerweise auch die der Europäischen Kommission.

1 Der Verhaltenskodex des ESS

Das Europäische Statistische System hat sich im Februar 2005 einen Verhaltenskodex gegeben, der von den Leitern der Nationalen Statistischen Ämter (NStÄ) der Mitgliedstaaten einstimmig angenommen, von der Europäischen Kommission im Mai 2005 in einer Empfehlung verkündet (Kommission 2005) und vom Rat der Wirtschafts- und Finanzminister ausdrücklich begrüßt wurde. Der Verhaltenskodex knüpft unmittelbar an die berufsethischen Grundsätze amtlicher Statistik der Vereinten Nationen (VN 1994) an, basiert auf der ESS-Qualitätserklärung (Eurostat 2001) und konkretisiert diese in 15 Grundsätzen sowie einem Indikatorenset zu jedem Grundsatz, das als Bezugspunkt zur Überprüfung und Anwendung des Kodex dient. Der Verhaltenskodex stellt einen umfassenden Qualitätsrahmen dar, der sich in 3 Teile gliedert, den institutionellen Rahmen, statistische Prozesse und statistische Produkte.

Der ESS-Qualitätsbegriff war die längste Zeit auf die Ergebnisqualität fokussiert, bevor die zugrundeliegenden Statistikprozesse als Voraussetzung und Garant der Produktqualität in das Blickfeld der Qualitätssicherung rückten. Der Verhaltenskodex greift diese Entwicklung auf, indem er den ‘klassischen’ Qualitätsdimensionen Relevanz, Genauigkeit und Zuverlässigkeit, Aktualität, Kohärenz, Vergleichbarkeit und Zugänglichkeit noch eine Reihe von Grundsätzen voranstellt, die auf statistische Prozesse abzielen. Dazu zählen eine solide Methodik und die

Wahl geeigneter Verfahren sowie die Vermeidung einer übermäßigen Belastung der Auskunftgebenden und Wirtschaftlichkeit.

Mit der Maßgabe, Prozesse nach Möglichkeit zu automatisieren und ihre Variabilität zu verringern sowie Verwaltungsdaten zu nutzen, werden die in den Produktqualitätsdimensionen inhärenten Optimierungsprobleme explizit gemacht und auf die Realisierung von Kostensparpotentialen ausgerichtet. Insofern dokumentiert der Verhaltenskodex, dass sich das ESS auch der bereits seit einiger Zeit auf politischer Ebene geführten Belastungsdiskussion aktiv stellt.

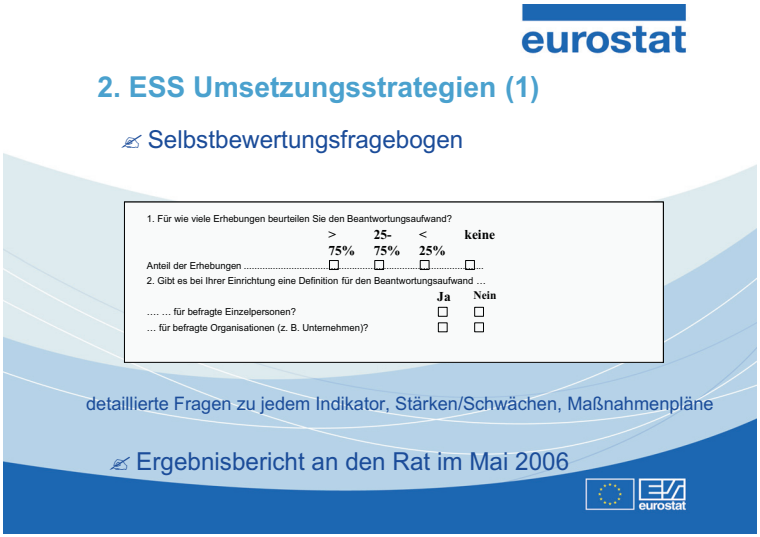
Weitere 6 Grundsätze des Verhaltenskodex, die sich auf den institutionellen Rahmen der Statistikerstellung und -verbreitung beziehen, vervollständigen den ganzheitlichen Qualitätsansatz. Darunter fallen die statistische Geheimhaltung, die Ressourcenausstattung der statistischen Stellen und eine Verpflichtung zur Qualität sowie Anforderungen an das gesetzliche Rahmenwerk und die fachliche Unabhängigkeit. Besonders hervorzuheben in diesem Zusammenhang ist das dem Kodex zugrunde liegende, weit gefasste Unabhängigkeitskonzept, das sowohl die Statistikproduktion als auch deren Verbreitung umfasst und insoweit über internationale Anforderungen bzw. nationale Praktiken in einigen Ländern hinausgeht (Eurostat 2005a).

Auch bezüglich der Qualitätsverpflichtung, der statistischen Methoden und Verfahren und der Produktqualitätsdimensionen setzt der Verhaltenskodex im weltweiten Vergleich hohe und so ambitionierte Standards, dass sie von vielen, wenn nicht allen ESS-Partnern bestenfalls angestrebt, gegenwärtig aber noch nicht vollständig umgesetzt sind.

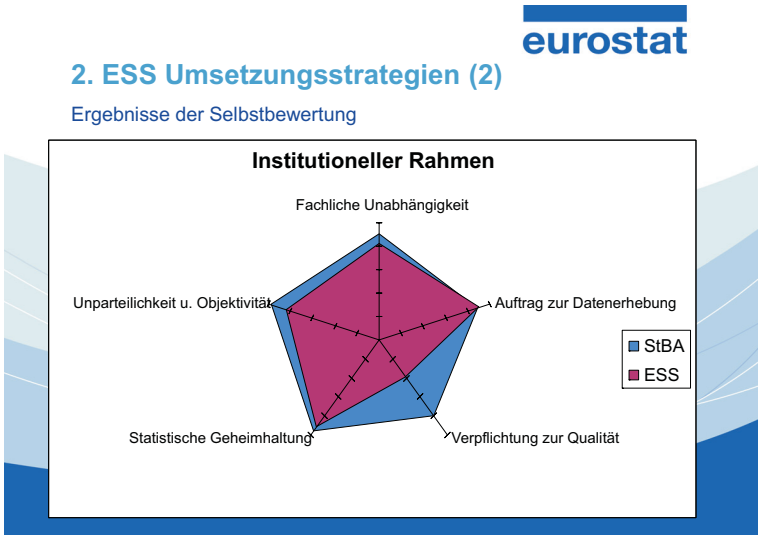
2 Die ESS-Umsetzungsstrategien

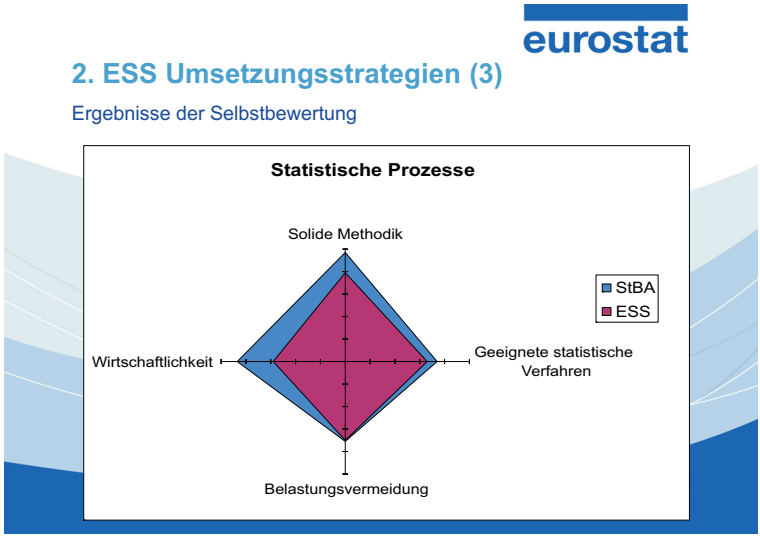
Umsetzung in den Nationalen Statistischen Ämtern

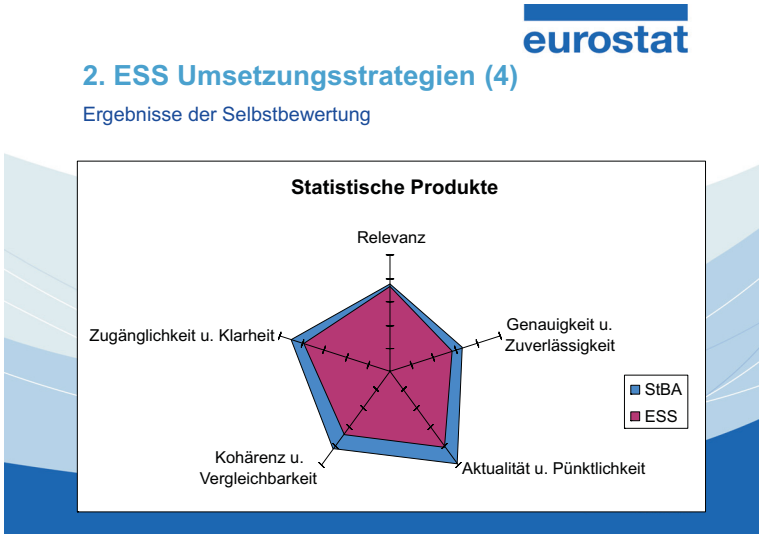
Eine erste Bestandsaufnahme, die Ende 2005 systemweit in Form einer Selbstbewertung der NStÄ anhand eines einheitlichen detaillierten Fragebogens (Eurostat 2005b) durchgeführt wurde, gibt einen ersten Überblick über den Umsetzungsstand des Verhaltenskodex.



Ihre Ergebnisse wurden in einem Bericht an den Rat zusammengefasst und veröffentlicht (Eurostat 2006a). Aus ihnen lässt sich ein Stärke-Schwäche-Profil ablesen, dass dem ESS bzw. seinen Hauptakteuren, den NStÄ, tragfähige Strukturen und Herangehensweisen in den klassischen Kernbereichen der amtlichen Statistik bescheinigt (Statistische Geheimhaltung, gesetzlicher Auftrag, solide Methodik, Statistikverbreitung und in weiten Teilen Unabhängigkeit).







Gleichzeitig wird deutlich, dass verstärkte Anstrengungen nötig sind, um moderne Konzepte und Verfahren der Qualitätssicherung und -bewertung, der Statistikproduktion einschließlich der Nutzung von Verwaltungsdaten systemweit umzu-

setzen, um auch insgesamt ein höheres Maß an Ergebnisqualität zu erreichen. Dabei werden z. T. erhebliche Unterschiede zwischen einzelnen Statistischen Ämtern und Statistikbereichen ausgemacht.

Entsprechend folgt die Umsetzungsstrategie des Verhaltenskodex im Wesentlichen einem selbstregulierenden Ansatz, der den einzelnen Akteuren im System ihren eigenen Weg und Zeitplan bis zur vollständigen Einhaltung des Kodex ermöglicht. Eurostat nimmt hierbei zunächst die Rolle eines Katalysators ein, der gemeinsam mit der Task Force „Implementierung des Verhaltenskodex", in der 10 Mitgliedstaaten vertreten sind, darunter auch Deutschland, gemeinsame Lösungen zur Unterstützung und Vereinfachung der nationalen Umsetzungsstrategien entwickelt.

Gleichzeitig kann das System auf eine Reihe gemeinsam entwickelter Standards, Werkzeuge und Handbücher zurückgreifen, die von Eurostat im Hinblick auf eine Beschleunigung des Technologietransfers innerhalb des Systems und Ausnutzung von Synergien subventioniert werden. Sie werden alle auf der Eurostat-Internetseite in einem eigens dazu entwickelten Teil „Verfahren des ESS"[2] zugänglich gemacht. Sie gehen z. T. auf Initiativen der Leadership Group Qualität zurück und werden im Rahmen der Umsetzung des Verhaltenskodex weiterentwickelt. Auch die zweijährigen Qualitätskonferenzen, die jeweils von einem Nationalen Statistikamt in Zusammenarbeit mit Eurostat organisiert werden, geben einen Überblick und ermöglichen auch den Austausch mit Wissenschaft und Praxis über Europa hinaus.

Weitere Stimuli zur Umsetzung des Verhaltenskodex gehen von den ESS-Arbeitsgruppen und Task Forces aus, die sich mit sämtlichen Qualitätsaspekten individueller Statistikbereiche, beginnend mit ihrer Harmonisierung, befassen. Sie werden ergänzt durch bereichsübergreifende Initiativen, wie z. B. die sog. FROCH Group, in der sich Nutzer und Statistikproduzenten gemeinsam mit Qualitätsverbesserungen einer Reihe zentraler Wirtschaftsindikatoren befassten und wichtige Impulse z. B. zu einer ESS-weiten einheitlichen Revisionspolitik gaben.

Im Juni 2007 hat Eurostat einen Musterrechtsakt für Gesetze zur Qualitätsberichterstattung vorgelegt, der Qualitätsberichte verpflichtend vorschreibt und einheitliche Qualitätsindikatoren vorsieht (Eurostat 2007). Auf deren Grundlage soll ein sog. ESS-Qualitätsbarometer Auskunft über die systemweite Entwicklung der Qualität in den wichtigsten Statistikbereichen ermöglichen.

Im Rahmen sog. ESS-Netze schließen sich Nationale Statistikämter in Zusammenarbeit mit Eurostat zusammen, um ihre Expertise in ausgewählten Themenbereichen zu bündeln und Lösungen zu entwickeln, die ESS-weit umgesetzt werden können. Beispielhaft sei hier das ESS-Netz zur Eilschätzung genannt.

Ohne auf die genannten Qualitätsinitiativen im Detail einzugehen, lässt sich ein genereller Trend zur Prozessorientierung (Hahn, 2005) und auch -harmoni-

2 http://ec.europa.eu/eurostat/quality

sierung ausmachen, und zwar sowohl innerhalb einzelner Statistikämter (z. B. im Rahmen von umfassend angelegten Umorganisationen zur Standardisierung der Statistikproduktion oder systematischen Qualitätsscreenings, sog. Statistikaudits sämtlicher Statistikbereiche) als auch auf der Ebene des Europäischen Statistischen Systems. Die damit verbundenen erwarteten Effizienzgewinne und Qualitätsverbesserungen werden zumindest auf Systemebene verstärkt durch ihre Ausrichtung an dem jeweils neuesten Stand des Wissens und dem Benchmarkingcharakter, mit dem dieser Wissenstransfer organisiert wird. Insofern scheint gegenwärtig ein Gleichgewicht zwischen Selbstregulierung der Mitgliedstaaten und Handlungsbedarf auf europäischer Ebene zur Umsetzung des Verhaltenskodex gelungen.

Systemweite Umsetzung des Verhaltenskodex

Angesichts der Bedeutung, die nationale statistische Stellen außerhalb des Statistikamtes in einigen Mitgliedstaaten für die Europäische Statistik haben, richtet sich der Verhaltenskodex ausdrücklich an alle Produzenten der Europäischen Statistik. Im Sinne des selbstregulierenden Ansatzes und im Einklang mit dem Grundsatz der Subsidiarität geht Eurostat davon aus, dass die erforderlichen Umsetzungsschritte von den NStÄ koordiniert werden, die dazu eine Strategie und einen Fahrplan festlegen. Dabei wird empfohlen, sich an der Bedeutung der jeweiligen statistischen Stellen für die Europäische Statistik zu orientieren. Ergebnisse einer schriftlichen Umfrage (Eurostat 2006b) machen deutlich, dass sich die Strategie der meisten NStÄ zur Koordinierung der Umsetzung des Verhaltenskodex durch die anderen nationalen statistischen Stellen an ihren eigenen Umsetzungsschritten orientiert: Einer Selbstbewertung der Systempartner zur Umsetzung des Kodex folgt die Definition von Maßnahmenplänen im Rahmen einer neu zu schaffenden bzw. verstärkten inter-institutionellen Koordinierungsinfrastruktur, die in der Regel gemeinsame Gremien zur Umsetzung des Verhaltenskodex sowie Informationsveranstaltungen umfasst. Der Aufwand dazu wird von vielen Statistikämtern sowie den beteiligten Stellen als hoch, aber gleichzeitig lohnend empfunden im Hinblick auf eine Stärkung der internen Koordinierung und damit des gesamten Statistiksystems. Auch schätzen die beteiligten Akteure den damit einhergehenden strukturierten Dialog, in dessen Rahmen ein Austausch über Werkzeuge und Methoden für die als erforderlich angesehenen Qualitätsverbesserungen stattfinden kann.

3 Überwachung der Umsetzung des Verhaltenskodex

Um dem mit der Annahme des Verhaltenskodex verbundenen politischen Mandat der Wirtschafts- und Finanzminister gerecht zu werden, muss eine auf Selbstverpflichtung und -regulierung basierende Umsetzungsstrategie mit Elementen der Überwachung und Transparenz kombiniert werden. Zu diesem Zweck werden gegenwärtig ESS-weit sog. Peer Reviews in allen Nationalen Statistikämtern und Eurostat durchgeführt. Sie beschränken sich im Wesentlichen auf den institutionellen Rahmen der Statistikerstellung, d. h. auf die Grundsätze 1 (Fachliche Unabhängigkeit), 2 (Auftrag zur Datenerhebung), 3 (Angemessene Ressourcen), 4 (Verpflichtung zur Qualität), 5 (Statistische Geheimhaltung), 6 (Unparteilichkeit und Objektivität) und 15 (Zugänglichkeit und Klarheit) des Verhaltenskodex. Alle Peer Reviews werden anhand einer einheitlichen Methodik (Hahn, 2006a u. Eurostat, 2006c) durchgeführt und resultieren in detaillierten standardisierten Länderberichten, die auf der Eurostat-Internetseite veröffentlicht werden. Der Peer Review für das Statistische Bundesamt ist für November 2007 vorgesehen.

Zur Einschätzung der Umsetzungserfolge im nationalen Statistiksystem, das heißt durch die auf nationaler Ebene an der Erstellung europäischer Statistiken beteiligten Stellen, gehen die Peer Reviews auch auf die Koordinierungsrolle des NStA ein, der ein eigenes Kapitel im Bericht gewidmet ist. Aufgrund der Unterschiede zwischen einzelnen nationalen Statistiksystemen hat dieser Teil eher deskriptiven Charakter; bewertende Elemente beruhen auf Gesprächen mit den wichtigsten Systempartnern und den Ergebnissen des Peer Reviews für diejenigen Grundsätze des Verhaltenskodex, für deren Umsetzung ein systemweiter Ansatz besonders vordringlich erscheint (z. B. Auftrag zur Datenerhebung und Verbreitung). Eine Bewertung des Umsetzungsstandes des Kodex in den nationalen Stellen außerhalb des Statistikamtes ist jedoch nicht Gegenstand des Reviews.

Die bis einschließlich Mai 2007 durchgeführten Peer Reviews scheinen die Ergebnisse der Selbsteinschätzungen der Statistikämter weitgehend zu bestätigen (Hahn, 2007), wenn auch deutlich wird, dass in einigen NStÄ eine vollständige Umsetzung der für den institutionellen Rahmen beschriebenen Grundsätze Änderungen der nationalen Statistikgesetze (Grundsatz 1 u. 2) voraussetzt. Auch sind Verbesserungen notwendig im Hinblick auf die Nutzung von Verwaltungsdaten (Grundsatz 2), einen vereinfachten Zugang zu Mikrodaten für Forschungszwecke (Grundsatz 5) und mehr Transparenz im Zusammenhang mit Vorabzugang zu statistischen Daten (Grundsatz 6), wenngleich der Umsetzungsstand der genannten Grundsätze insgesamt als hoch bezeichnet werden kann. Größere Anstrengungen verlangen die Grundsätze 4 und 15 und z. T. auch 3, obwohl die Situation hier von Statistikamt zu Statistikamt sehr unterschiedlich ist. Im Zusammenhang mit dem Grundsatz 4 (Verpflichtung zu Qualität) werden einige vorbildliche Praktiken

ausgemacht, die im Rahmen eines Eurostat-Seminars am 5. Juni 2007 in Luxemburg vorgestellt wurden.

Die Peer Reviews erlauben eine Validierung und Weiterentwicklung der Selbsteinschätzungen der Statistikämter, z. B. durch eine im Rahmen des Peer Review-Prozesses von Eurostat erstellte Übersicht, die dem Statistikamt seine Positionierung im Vergleich zum ESS-Durchschnitt ermöglicht. Gleichzeitig tragen sie selbst auch zur Implementierung des Verhaltenskodex bei. Wesentliche Elemente dazu stellen der partnerschaftliche Diskussionsprozess, die Identifizierung vorbildlicher Lösungen und Empfehlungen des Peer Review Teams zur Ergänzung und Priorisierung der von dem Statistikamt im Vorfeld erarbeiteten Liste von Maßnahmenplänen zur Umsetzung sämtlicher Grundsätze des Verhaltenskodex dar.

An diese Maßnahmenpläne wird Eurostat anknüpfen, um die Umsetzung des Verhaltenskodex auch über die Peer Reviews hinaus zu überwachen. Sie bilden zusammen mit den Peer Review-Berichten eine wichtige Informationsbasis für den 2008 an den Rat und das Europäische Parlament zu übermittelnden Bericht der Kommission zum Umsetzungsstand des Kodex. Ein weiteres wichtiges Element stellt ein Sachstandsbericht der Nationalen Statistikämter über die Umsetzung des Verhaltenskodex durch die anderen nationalen statistischen Stellen dar, dessen Übermittlung bis Anfang 2008 vereinbart wurde.

Die Überwachung der Umsetzung des Verhaltenskodex wird ergänzt durch die Schaffung eines Beratungsgremiums zur Europäischen Statistischen Regierungs-

führung, das u. a. jährlich über den Umsetzungsstand durch Eurostat berichten und Anstöße zur Förderung der Umsetzung des Kodex im ESS als Ganzes geben soll (Kommission, 2006).

4 Ein Gütesiegel für die Umsetzung des Verhaltenskodex

Die Umsetzung des Verhaltenskodex im ESS wird weder für alle Statistikämter noch für alle Statistikbereiche gleichzeitig erfolgen. Vielmehr darf davon ausgegangen werden, dass ein Kern der insgesamt 15 Grundsätze relativ zügig systemweit umgesetzt wird oder bereits ist, während die Umsetzung anderer Grundsätze oder einzelner Indikatoren eher maßgeschneiderten Fahrplänen der Statistikämter folgt, die für die einzelnen Statistikbereiche durchaus unterschiedlich ausfallen können.

Dabei gibt es auch einige (wenige) Bereiche, in denen der Verhaltenskodex zu keinem Zeitpunkt vollständig umgesetzt werden kann aufgrund des spezifischen institutionellen Rahmens, innerhalb dessen sie erstellt werden. Bei ihnen handelt es sich eher um sog. statistische Dienste als um amtliche Statistik.

In anderen Statistikbereichen ist eine vollständige Umsetzung des Kodex theoretisch zwar denkbar, wird aber aufgrund von Kosten-Nutzen-Abwägungen, die zusätzliche Investitionen in Qualitätsverbesserungen als nicht verhältnismäßig erscheinen lassen, möglicherweise gar nicht angestrebt.

Die Unterscheidung zwischen diesen Bereichen ist für den Nutzer keineswegs evident, kann aber von zentraler Bedeutung im Hinblick auf die Beurteilung der Zweckmäßigkeit der Verwendung einer Statistik werden. Mit der Entwicklung eines Gütesiegels für Statistikaggregate versucht Eurostat ihre Qualität für die Nutzer auf europäischer Ebene transparenter zu machen. Der Verhaltenskodex soll dabei als Bezug dienen, der gleichzeitig auch erstmals eine Unterscheidung zwischen amtlicher und nicht-amtlicher Europäischer Statistik ermöglicht.

Der für ein Gütesiegel erforderliche Bewertungsprozess soll die Schritte zur Umsetzung des Verhaltenskodex integrieren und durch systematische Überprüfungen von Maßnahmen zur Prozess- und Produktqualitätssicherung auf europäischer Ebene ergänzen. Im Zusammenhang mit den sog. Qualitätsprofilen, deren Erstellung einem dem Gütesiegel ähnlichen Prozess zur Integration und Zusammenfassung von relevanten Informationen bis hin zu einem nutzer-orientierten Qualitätsüberblick für europäische Schlüsselindikatoren folgt, wurde deutlich, dass dieser Bewertungsprozess mindestens genauso bedeutsam wie die von einem Gütesiegel ausgehende Botschaft ist: „Europäische Statistik erstellt im Einklang mit dem Verhaltenskodex" (Hahn, 2004).

Wie Erfahrungen Nationaler Statistikämter mit ähnlichen Strategien gezeigt haben (Hahn, 2006b), bietet ein an der Umsetzung des Verhaltenskodex orientiertes Gütesiegel auch einen Rahmen zu einer verstärkten Kooperation der an der Erstellung von Europäischen Statistiken beteiligten Stellen auf europäischer Ebene. Dies scheint vorteilhaft insbesondere im Hinblick auf eine Stärkung der Koordinierungsrolle Eurostats gegenüber anderen Kommissionsdienststellen.

Schließlich gibt ein Gütesiegel und die damit verbundene Öffentlichkeitswirkung auch Anreize für kontinuierliche Qualitätsverbesserungen und leistet damit selbst einen Beitrag zur Umsetzung des Verhaltenskodex.

Referenzen

Europäische Kommission (2005): Empfehlung der Kommission zur Unabhängigkeit, Integrität und Rechenschaftspflicht der statistischen Stellen der Mitgliedstaaten und der Gemeinschaft, KOM (2005) 217 endg. vom 25. Mai 2005.

Europäische Kommission (2006): Vorschlag für einen Beschluss des Europäischen Parlaments und des Rates zur Einsetzung eines Europäischen Beratungsgremiums für die Statistische Governance, KOM (2006) 599 endg. vom 18. Oktober 2006.

Eurostat (2001): Quality Declaration of the European Statistical System. online unter: http://ec.europa.eu/eurostat/quality

Eurostat (2005a): Elements of the national statistical institutes' and Eurostat's institutional set-up, October 2005. online unter: http://ec.europa.eu/eurostat/quality

Eurostat (2005b): Europäische Statistiken: Verhaltenskodex, Fragebogen zur Selbstbewertung. online unter: http://ec.europa.eu/eurostat/quality

Eurostat (2006a): Report on the results of the first self-assessments carried out by the statistical authorities of the European Statistical System against the principles and indicators of the European Statistics Code of Practice, May 2006, online unter: http://ec.europa.eu/eurostat/quality

Eurostat (2006b): Extension of the European Statistics Code of Practice to other relevant national producers of statistics, October 2006. Eurostat internal working paper.

Eurostat (2006c): Eurostat Methodik der Peer Reviews zum Verhaltenskodex, Leitlinien für Nationale Statistische Ämter u. Leitlinien für Peers, 2006, online unter: http://ec.europa.eu/eurostat/quality

Eurostat (2007): Eurostat guidelines for data quality reporting in Community statistical legislation, June 2007. Eurostat internal working paper.

Hahn, M. (2004): Quality Profile for Structural Indicators, Proceedings of the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004, CD-ROM. online unter: http://q2004.destatis.de (ISBN 3-8246-0733-6).

Hahn, M., und Lindén, H. (2005): Quality Measures for enhanced Quality of Statistics in the European Statistical System, Proceedings of Statistics Canada Symposium 2005, Canada.

Hahn, M. (2006a): Code of Practice Peer Reviews: First steps in defining a European Statistical System wide approach. Conference on Data Quality for International Organizations, Newport (UK), 27-28 April 2006.

Hahn, M., und Willeke, C. (2006b): Labelling of Statistical Information: Some introductory Considerations. Paper presented at the Conference on Data Quality for International Organisations Newport, UK 27-28 April 2006.

Hahn, M. und Eidukynaite, S. (2007): The European Statistical System peer review methodology and emerging results. ISI Conference August 2007 – im Erscheinen.

United Nations Statistical Commission (1994): Fundamental Principles of Official Statistics. online unter: www.un.org.

Qualitätskontrolle im Telefoninterview

Christina Buchwald, Katja Lukanow

1 Einleitung

Die rasch wachsende Bedeutung computergestützter Telefonbefragungen sowohl in der Wissenschaft wie bei kommerziellen Befragungen ist unumstritten. Diese Entwicklung resultiert vor allem daraus, dass telefonische Interviews gegenüber mündlichen Interviews deutliche forschungsökonomische Vorteile haben und in inhaltlich-methodischer Hinsicht – vor allem durch entsprechende Filterführung – deutlich mehr Leistungen erlauben als schriftliche Befragungen.

Neben solchen Aspekten sollten alle Erhebungstechniken sicherstellen, dass die Forschung valide ist. In diesem Zusammenhang ist die Befragungssituation von wesentlicher Bedeutung, die sich jedoch bei der telefonischen Befragung neu gestaltet. Die Befragungssituation, in welcher der Interviewer eine zentrale Rolle spielt, wirft neue Fragen und Probleme auf. Dementsprechend besteht Bedarf an der Aufdeckung von Kriterien und Einflussfaktoren der Interviewqualität und an der Bestimmung von Instrumenten der Qualitätskontrolle und Qualitätsverbesserung. Gemessen an der sehr traditionsreichen methodologischen Forschung zur Qualität mündlicher Interviews und der umfangreichen aus ihr entstandenen Literatur gibt es bisher zu CATI-Befragungen[1] nur wenige einschlägige Arbeiten.

Im CATI-Labor des Zentrums für Sozialforschung Halle e.V. (zsh) an der Martin-Luther-Universität Halle-Wittenberg, welches seit 2001 in Betriebsführung des zsh ist, konnten bereits über 60.000 telefonische Interviews durchgeführt werden. Ein Großteil der Befragungen wurde von Universitäten, verschiedenen Bundesinstituten und vom Sonderforschungsbereich 580 in Auftrag gegeben. Die Frage nach der Qualität der Interviews stand aufgrund der anspruchsvollen Auftraggeber und auch Befragungsthemen stets im Vordergrund. Aus den methodischen Forschungsergebnissen zu diesen CATI-Befragungen sollen im Folgenden vier wesentliche Faktoren, die sich auf den Interviewer und den Befragten als Akteure der Kommunikationssituation beziehen, diskutiert werden und wichtige Qualitätsmaßstäbe darstellen:

- die Befragungssituation im Telefoninterview,
- der Interviewer als Einflussfaktor,
- die Ausschöpfungsquote der Befragung,
- die Beurteilung des Interviews nach Meinung des Befragten.

1 Computer Assisted Telephone Interviewing

Da die Befragungssituation für die Qualität des Interviews eine zentrale Bedeutung hat, ist vor dem Hintergrund der Qualitätssicherung zunächst in Kapitel 2.1 die Analyse dieser Kommunikationssituation zwischen Befragtem und Interviewer notwendig. Weil der Interviewer hier eine zentrale Rolle spielt, werden die Effekte, welche von ihm ausgehen, genauer untersucht (Kapitel 2.2). Daraus sollen Erkenntnisse für die Schulung und Rekrutierung der Interviewer gewonnen werden (Kapitel 3).

Darüber hinaus war und ist das CATI-Labor des zsh bemüht, Antworten auf die Frage nach Instrumenten der Qualitätskontrolle zu finden. Dazu werden nach Abschluss eines Interviews einige zusätzliche Fragen an den Befragten und an den Interviewer gestellt, die für eine Klärung der Qualitätssicherung geeignet erscheinen (Kapitel 5). Neben den Ergebnissen dieser Nachfragen gilt vor allem auch die Ausschöpfungsquote (Kapitel 4) als ein wichtiger Qualitätsmaßstab für Umfragen.

In diesem Beitrag zur Qualität von Telefoninterviews werden Vorteile, aber auch Schwierigkeiten, die methodisch bei CATI-Befragungen auftreten können, diskutiert.

2 Das computergestützte Telefoninterview

Das Telefoninterview hat sich heute als Umfrageinstrument etabliert. Vor nicht so langer Zeit wurde diese Methode der Datenerhebung noch als „quick and dirty" abgewertet. Lange war diese neue Erhebungsmethode methodisch noch wenig fundiert und kam für die „seriöse" Forschung nicht in Frage. Ein wesentlicher Einwand gegen den Einsatz des Telefoninterviews zur Datenerhebung war die Repräsentativität von Stichproben, solange die Telefonversorgung in den privaten Haushalten als nicht ausreichend gelten konnte. In den USA konnte man bereits in den 70er Jahren von einer Vollversorgung der Privathaushalte mit Telefonen sprechen, was den Einsatz dieser Befragungsform schon viel früher möglich machte. Die zögerliche Entwicklung in Deutschland ist, neben dem ehemals schlechten Ansehen telefonischer Umfragen, auch auf die bis in die 80er Jahre hinein mangelnde Abdeckung der Privathaushalte mit Telefonanschlüssen zurückzuführen. Heute kann man in Ost- wie auch in Westdeutschland von einer Vollversorgung sprechen, die den Einsatz von Telefonumfragen als sinnvoll erscheinen lässt (vgl. Statistisches Bundesamt 2005, S. 139).

Telefonische Befragungen werden heute fast ausschließlich mit CATI-Systemen durchgeführt. Das telefonische Interview zeichnete sich erst durch die Verknüpfung von Telefon und Computer aus und ist heute vor allem ein computergestütztes Telefoninterview.

Aus der Anwendung von CATI resultiert eine Reihe forschungsökonomischer Vorteile[2], aber auch viele Vorteile für die Interviewer selbst. Bei der Durchführung von face-to-face Interviews kommen eine Fülle von Aufgaben auf den Interviewer zu. Er muss sich im Fragebogen und dessen Filterführungen zurechtfinden, anstehende Fragen wörtlich rezitieren, Antworten richtig protokollieren, bei offenen Fragen sinngemäß mitschreiben und natürlich auch auf sein Interviewverhalten achten, um von ihm ausgehende Effekte auszuschalten. Bei der Bewältigung dieser Aufgaben kann dem Interviewer eine Vielzahl von Fehlern unterlaufen. Es besteht die Gefahr, dass er sich in der Filterführung des Fragebogens „verirrt", unabsichtlich Fragen auslässt oder doppelt stellt, Anweisungen zur Rotation der Antwortvorgaben nicht befolgt oder Intervieweranweisungen missachtet (vgl. Fuchs 1994). Im computergestützten Interview kann die Arbeit des Interviewers erheblich erleichtert werden (vgl. Buchwald, 2006, S. 7ff.).

Die Anwendung von CATI garantiert, dass die Fragen in der richtigen Reihenfolge auf dem Bildschirm erscheinen. Auch die Filterführung wird dem Interviewer durch das CATI-System abgenommen, indem jedem Befragten nur die Fragen gestellt werden, die aufgrund seiner vorhergehenden Angaben zutreffend sind. Zudem ermöglicht CATI die Rotation von Statements und Fragekomplexen und gewährleistet, dass keine Fragen oder Rangreihen (beim Ranking) ausgelassen oder doppelt gestellt werden können. „Der computergestützte Fragebogen stellt sicher, dass jedem Befragten ein konsistentes und vollständiges Interview präsentiert wird (sofern das Programm stimmt)" (Fuchs, 1994, S. 40) .

CATI erleichtert nicht nur dem Interviewer die Durchführung der Befragung, sondern liefert darüber hinaus eine höhere Datenqualität als andere Erhebungsmethoden. Der Erhebungs- und Auswertungsaufwand wird durch den Einsatz von CATI erheblich verringert und somit daraus folgende Fehlerquellen minimiert.

Die Dateneingabe wird im Verlauf des Interviews vom Interviewer selbst vorgenommen, da er die bereits codierten Antworten in den Computer eingibt. Dabei werden die Eingaben einer Reihe von Kontrollmechanismen unterzogen. Nur gültige Codes werden akzeptiert und demzufolge fehlerhafte Eingaben häufig erkannt und korrigiert. Aufgrund der Dateneingabe während der Interviewsituation entfällt die Phase der Codierung und der Datenübertragung. Die vorliegenden Rohdaten können sofort, ohne großen Aufwand in ein spezielles Auswertungsprogramm wie zum Beispiel SPSS übertragen werden. Das CATI-System beugt somit Fehlern vor, welche bei dieser Übertragung möglich sind. Ebenfalls ist aufgrund der sofortigen Dateneingabe durch die Interviewer eine Zwischenauswertung jederzeit und ohne großen Aufwand realisierbar. Auf diesem Weg erfährt der

2 Geographisch weitgestreute Stichproben können kostengünstiger realisiert werden; Personen, die für persönliche Interviews kaum kontaktiert werden können, sind leichter zu erreichen; die Erhebungszeit einer Umfrage wird erheblich verkürzt usw. Eine detaillierte Aufzählung solcher Vorteile erläutert Fuchs (1994, S. 34ff.).

Forscher nicht nur etwas über die bisherige Ausschöpfung, sondern kann zudem Prioritäten setzten. Wird zum Beispiel ersichtlich, dass es sich bei den bisher befragten Personen fast ausschließlich um Frauen oder um jüngere Personen handelt, kann er Maßnahmen ergreifen, die zu einer ausgewogenen Verteilung von Frauen und Männern oder jungen und alten Personen führen.

Auf der Grundlage dieser Vielzahl technischer Möglichkeiten sind mittels CATI auch sehr komplexe Umfragen mit einer komplizierten Filterstruktur durchführbar, ohne den Interviewer zu überfordern. Während ihm durch die technischen Hilfen viel Arbeit abgenommen wird, stellt die spezielle Kommunikationssituation bei telefonischen Umfragen ganz neue Herausforderungen an ihn, was im folgenden Abschnitt erläutert werden soll.

2.1 Die Befragungssituation im Telefoninterview

Eine Befragung, ob face-to-face oder telefonisch, stellt eine spezifische Situation dar. Es handelt sich um eine soziale Situation, die durch eine streng formale Kommunikationsstruktur gekennzeichnet ist.

Die Kommunikation in einer Befragung läuft wie folgt ab:

- Der Forscher formuliert eine Erhebungsfrage.
- Der Interviewer liest dem Befragten die vom Forscher vorformulierte Frage wortgemäß vor.
- Der Befragte hört diese Frage und versucht die Bedeutung zu entschlüsseln.
- Der Befragte findet gedanklich eine Antwort.
- Er formuliert seine Antwort und ordnet sie in die gegebenen Antwortkategorien ein.
- Der Interviewer hört und verarbeitet die Antwort und protokolliert diese.
- Der Forscher interpretiert die Antwort.

Der Interviewer ist während dieser Kommunikation nur verfügt, Fragen zu stellen und möglichst keine weiteren Erläuterungen zu geben. Vom Befragten wird verlangt, kurz und präzise ohne Abschweifungen zu antworten. Zudem handelt es sich bei einer standardisierten Befragung um „(...) eine stark reglementierte und asymmetrische Kommunikation zwischen zwei Personen“ (Fuchs, 1994, S. 22). Von einer asymmetrischen Kommunikation muss gesprochen werden, weil dem Interviewer der Fragebogen und die Befragungssituation bereits vorher bekannt sind, der Befragte sich allerdings in einer fremden und ungewohnten Situation befindet.

Bei dem Versuch, diese formale und asymmetrische Kommunikationsstruktur in eine soziale Gesprächssituation einzubinden, kommt dem Interviewer eine Reihe von Aufgaben zu. Er hat den Auftrag, den Befragten in seine Rolle als Interviewpartner einzuführen sowie einen guten Rapport herzustellen und aufrecht-

zuerhalten. Da er selbst aber auch ein Teil der strukturierten Kommunikation ist, werden unterschiedliche Ansprüche an sein Verhalten gestellt.

Das bedeutet, dass der Interviewer sich zum einen veranlasst fühlen sollte, ein positives Gesprächsklima herzustellen, um somit den Befragten in seine Rolle einzuführen sowie seine Antwortbereitschaft und Motivation zu unterstützen. Zum anderen wird von ihm verlangt, sich weitgehend neutral zu verhalten. Der Interviewer muss demnach die vorgezeichnete Kommunikation abarbeiten und darüber hinaus die soziale Situation organisieren. „Der Befragte darf nicht den Eindruck gewinnen, es handele sich um eine Prüfungssituation“ (Friedrichs, 1990a, S. 216).

Eine Besonderheit der telefonischen Befragung ist die Beschränkung auf den akustischen Kommunikationskanal. In einer Kommunikation zwischen Menschen, in der man sich persönlich von Angesicht zu Angesicht gegenübersteht, werden Informationen nicht nur über Worte ausgetauscht, sondern auch über nicht-verbale Kommunikation, also Gesten, Mimik und die Art des Auftretens. Während der Kommunikation über das Telefon findet dieser Austausch ausschließlich über die Stimme statt. Aufgrund des Wegfallens aller äußerlichen Stimuli müssen sich Interviewer und Befragter mehr als in einer face-to-face Befragung aufeinander konzentrieren. Im Telefoninterview entsteht zwar eine „Intimität der Gesprächssituation“, aber gleichzeitig auch eine „Distanz zur (...) relativ anonymen Person des Interviewers“ (Hormuth/Brückner 1985, S. 540). Einerseits könnte sich der Befragte durch die räumliche und optische Distanz freier fühlen und demnach unbeschwerter und ehrlicher antworten, andererseits wird die Interaktion über das Telefon dadurch belastet, dass Befragte die telefonische Variante anonymer und somit weniger angenehm empfinden. Da die Kommunikationssituation anonymer ist als in einem persönlichen Interview, ist es über das Telefon um einiges schwieriger, einen angemessenen Konversationston zu finden.

Bei der computergestützten telefonischen Befragung kommt hinzu, dass der Interviewer nicht nur mit dem Befragten interagieren muss, sondern auch mit dem Computer, wie in Abbildung 1 erkennbar ist.

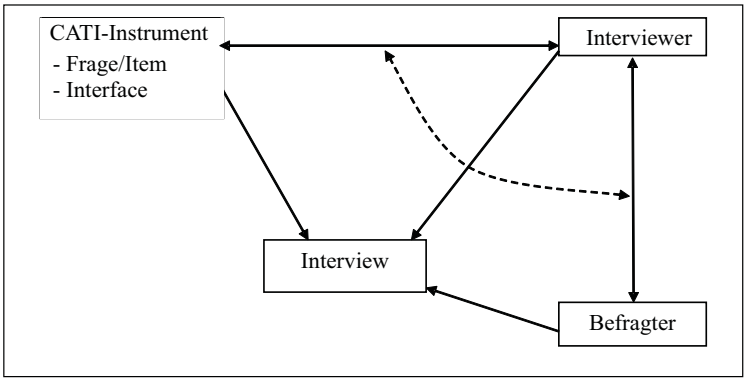


Abb. 1: Interviewer-Befragten-Interaktion im computergestützten Interview
Quelle: Fuchs (2000), S. 73

Der Computer stellt einen zusätzlichen „Akteur" in der Erhebungssituation dar, da er wesentlich an der Durchführung des Interviews beteiligt ist. Der Computer präsentiert Fragen und Antworten, speichert die eingegebenen Daten und gibt dem Interviewer Anweisungen. Dabei beeinflusst das Interface des Computers (z. B. Benutzeroberfläche, Anweisungen, Möglichkeiten der Dateneingabe) das Verhalten der beteiligten Akteure. „D. h. selbst bei gleichen Frageformulierungen und identischen Antwortvorgaben wirken unterschiedliche Bildschirmdesigns differenziert auf das Verhalten des Interviewers dem Befragten gegenüber und damit möglicherweise auf die Datenqualität" (Fuchs 2000, S. 72). Die Gestaltung des Bildschirms kann den Kommunikationsablauf beeinflussen, indem sie es dem Interviewer erschwert oder erleichtert, diesen Ablauf ohne Störungen zu befolgen.

Fehler, die bei der Bildschirmgestaltung auftreten können, sind unübersichtliche Bildschirmdesigns, unklare Trennungen zwischen den Fragen, Antwortkategorien und Intervieweranweisungen, notwendiges Scrollen oder die Verteilung wichtiger Informationen auf mehreren Bildschirmen (ebenda, S. 74).

Der Interviewer sollte eine Benutzeroberfläche vorfinden, die es ihm ermöglicht, Fragen optimal zu erfassen und wörtlich vorzulesen sowie die Intervieweranweisungen (z. B. durch farbliches Abheben) wahrzunehmen. Sind diese Voraussetzungen gegeben, kann der Interviewer seine Aufmerksamkeit auf den Kommunikationsablauf und somit auch auf den Befragten richten. Besteht diese Voraussetzung nicht, gilt seine Konzentration zunehmend dem Fragebogen. Das Bildschirmdesign beeinflusst demnach die Administration des Fragebogens durch den Interviewer und somit das Befragtenverhalten.

Die Interaktion mit dem Befragten und dem Computer sowie die Beschränkung auf den akustischen Kommunikationskanal stellen eine Reihe von Anforde-

rungen an den Interviewer. Gleichzeitig soll gewährleistet werden, dass sich jeder Befragte – unabhängig von der Person, welche die Befragung durchführt – in einer gleich gestalteten Interviewsituation befindet.

Aus dieser Vielzahl von Ansprüchen an den Interviewer resultiert eine Reihe von Fehlerquellen, die es zu untersuchen gilt.

2.2 Der Interviewer als Einflussfaktor – Interviewereffekte im telefonischen Interview

„Das – standardisierte – Interview ist ein sozialwissenschaftliches Instrument, und wie bei allen Instrumenten wird vorausgesetzt, dass es gleichartig misst, ungeachtet der Person, die es anwendet" (Friedrichs 1990b, S. 414) .

Diese Forderung entspricht jedoch einem Ideal. Antwortverzerrungen im Interview können viele unterschiedliche Ursachen haben. Der Fragebogen (Frageformulierung, Antwortkategorien, Fragereiheneffekte), der Befragte (soziale Erwünschtheit, Response-Set, Meinungslosigkeit), die Interviewsituation (Anwesenheit Dritter), aber auch der Interviewer selbst sind Fehlerquellen für systematische Antwortverzerrungen. Unerwünschte Einflüsse auf die Ergebnisse, die von den Interviewern ausgehen, werden als Interviewereinflüsse oder Interviewereffekte bezeichnet (vgl. Reinecke 1991, S. 35). Vom Interviewer wird verlangt, seinen Einfluss auf die Antworten des Befragten möglichst auszuschalten.

In zahlreichen Untersuchungen bezüglich persönlicher Interviews wurde analysiert, dass der Interviewer durch Geschlecht, Alter, Statur, Kleidung, aber auch durch Mimik, Gestik, Blickkontakt, Paralinguistik sowie sein Verhalten auf den Befragten wirken kann. Das Modell von Cannell/Kahn (1968) (siehe Abbildung 2) zeigt deutlich, wie sich die Merkmale, Einstellungen, Erwartungen etc. des Interviewers auf die Interviewsituation auswirken. Die Einflüsse sind den Forschern schon seit geraumer Zeit bekannt, doch das Ausmaß dieser Einwirkungen konnte bei persönlichen Interviews nur schwer eingeschätzt werden, weil der Interviewer in der Regel allein mit dem Befragten ist. Bei der telefonischen Befragung, sofern sie in einem zentralen Telefonlabor stattfindet, stehen die Interviewer unter der Kontrolle des Forschers oder des Supervisors. Durch deren Anwesenheit sowie die Möglichkeit des Mitschneidens kann nun die Gesprächssituation untersucht werden. „Wir ahnen erst durch die telefonischen Interviews, wie stark Interviewer bisher in face-to-face Interviews durch die Art ihrer Gesprächsführung das Instrument verändert, d. h. nicht neutral administriert haben" (Friedrichs 1990b, S. 414).

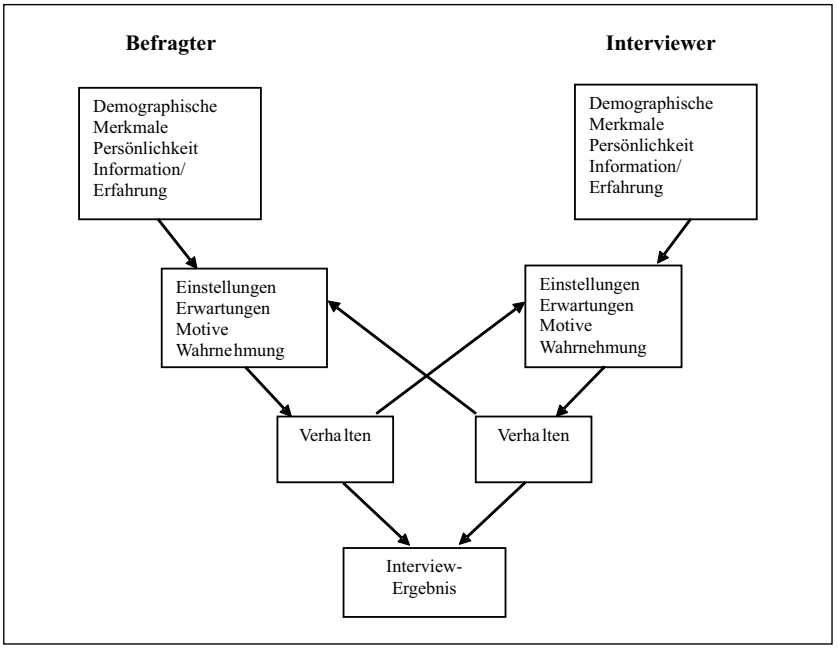


Abb. 2: Das Interview als sozialer Prozess
Quelle: Friedrichs (1990a), S. 218

Bei telefonischen Befragungen entfallen sichtbare Interviewermerkmale, wobei Sprache, Stimme, Geschlecht sowie verbal vermittelte Erwartungshaltungen und Einstellungen des Interviewers die Antworten der Befragten ebenso verzerren können. „Die genaue Formulierung der Frage, der Tonfall, in dem sie gestellt wird, die Pausen (vor allem bei offenen Fragen), bestärkende Partikel wie ‚Hmm' (…) verändern das Instrument" (ebenda, S. 414).

Folgende Merkmale des Interviewers können als bedeutsam erachtet werden, da sie einen wesentlichen Einfluss auf die Administration des Fragebogens haben:

- soziodemographische Merkmale des Interviewers (Alter, Geschlecht, Erfahrung etc.);
- Persönlichkeitseigenschaften des Interviewers (z. B. Stimmeigenschaften, rhetorische Fähigkeiten, Selbstvertrauen etc.);
- bewusstes und unbewusstes Fehlverhalten;
- Erwartungshaltungen des Interviewers an den Befragten und Einstellungen des Interviewers zum Gegenstand der Befragung.

Durch Anleitung und Vorbereitung der Interviewer sowie durch die Standardisierung der Befragung wird versucht, die Einflüsse weitgehend zu neutralisieren. Die Wirkung bestimmter individueller Merkmale kann jedoch trotz hochgradiger Standardisierung des Fragebogens nicht völlig verhindert werden.

Allen genannten Interviewermerkmalen muss im Rahmen der Qualitätsbewertung besondere Aufmerksamkeit gewidmet werden. Vor dem Hintergrund der spezifischen Situation im Telefoninterview sind jedoch die Persönlichkeitseigenschaften, also vor allem die Stimmeigenschaften und rhetorischen Fähigkeiten, der wesentliche Aspekt, auf den ich im Folgenden eingehen möchte.

2.2.1 Persönlichkeitseigenschaften

Da im telefonischen Interview alle sichtbaren Persönlichkeitsmerkmale ausgeschaltet werden, kommen vor allem die paralinguistischen Merkmale eines Interviewers wie Stimmhöhe, Modulation und Lautstärke der Stimme, Sprechgeschwindigkeit und Pausen, zum Tragen.

Stimme und Sprechweise sind ausschlaggebend beim Telefonieren, denn sie entscheiden über die Wirkung, Ausstrahlung, Überzeugungskraft und auch über die Glaubwürdigkeit des Interviewers. Die Stimme gibt dem Gesprächspartner nicht nur Auskunft über das Geschlecht des Anderen, sondern lässt ihn auch Vermutungen über sein Alter oder seine geographische Herkunft anstellen. Ebenfalls können der Gefühlszustand oder die Persönlichkeit, wie beispielsweise Selbstbewusstsein oder Schüchternheit des Sprechers eingeschätzt werden. Wie stark manche Menschen von anderen bezüglich ihrer Persönlichkeit – neben der Optik und dem Sprachinhalt – durch ihre Stimme beurteilt werden, zeigt die Abbildung 3.

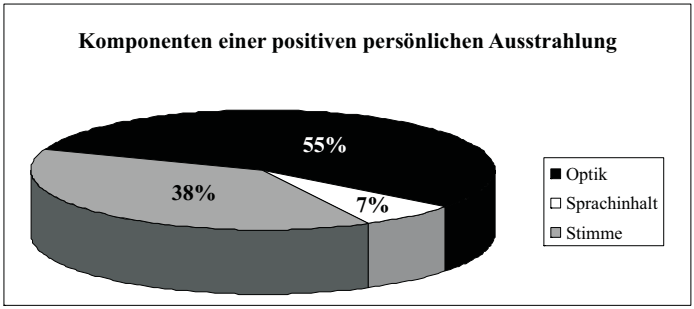


Abb. 3: Komponenten der Persönlichkeitsbeurteilung
Quelle: Amon (2000), S. 15

Hier wird sichtbar, welche Bedeutung die Stimme für die Einschätzung der Persönlichkeit eines Menschen hat. Allein das Gesicht und die Kleidung werden als wichtiger erachtet. Um die Bereitschaft für ein Interview zu erlangen, muss der Interviewer einen positiven Eindruck bei dem Befragten erzeugen, der bei CATI-Befragungen ausschließlich über die Kommunikation durch das Telefon hergestellt werden kann.

„Als Zuhörer haben wir an den Sprecher eine Reihe von Wünschen. Wir möchten ihn akustisch verstehen und seinen Worten folgen können. Dazu bedarf es einer deutlichen Artikulation, einer tragfähigen Stimme und vor allem einer Gestaltungsweise, die mitreißt, (...), jedoch keinesfalls belastet“ (Coblenzer/Muhar 1989, S. 8).

2.2.1.1 Stimmeigenschaften

Die Stimme jedes Einzelnen zeichnet sich durch unterschiedliche Stimmeigenschaften aus, die vom Hörer registriert und als angenehm oder unangenehm empfunden werden.

Stimmhöhe und Modulation

Jede Person hat einen individuell vorgegebenen Stimmumfang. Eine besonders günstige Tonhöhe, die am wenigsten anstrengt und der Konstitution des Sprechers entspricht, liegt im unteren Drittel des gesamten Stimmumfangs. Diesen Bereich nennt man Indifferenzlage. Vom Hörer wird diese Stimmhöhe als natürlich und angenehm empfunden. „Je weiter sich die Stimme gewohnheitsgemäß von der Indifferenzlage entfernt, als desto unnatürlicher wird sie vom Hörer empfunden und entsprechend negativ eingeschätzt“ (Eckert/Laver 1994, S. 39).

Spricht man immer in dieser Indifferenzlage, würde man in einen Monotonismus verfallen, der vom Hörer sehr negativ bewertet wird. Während monotones Sprechen sehr unnatürlich und unpersönlich wirkt und zudem kein Interesse erzeugt, ruft eine gute Modulation – das heißt deutliches Anheben und Senken der Stimme – Aufmerksamkeit hervor und wirkt lebendiger. Das Anheben und Absenken der Stimme sollte jedoch um die Indifferenzlage pendeln (vgl. Coblenzer/Muhar 1989, S. 12). Hörer bevorzugen meist eine etwas tiefere Stimmlage und schätzen Sprecher mit deutlichen Tonhöhenvariationen als kompetenter, selbstbewusster, kommunikationsfreudiger und wohlwollender ein (vgl. Eckert/Laver 1994, S. 39). Auch situationsbedingte Stimmungen wie Aufregung, Freude, Stress oder Frust führen zur Veränderung der Indifferenzlage. Möchte der Sprecher beispielsweise sehr höflich klingen, tut er das oft mit erhöhter Stimmlage. Gebraucht der Sprecher jedoch permanent eine erhöhte oder viel zu tief gewählte Sprechlage, wird das vom Hörer als störend und unnatürlich empfunden und weckt sogar den Eindruck von Inkompetenz.

Lautstärke

Welche Lautstärke der Sprecher zu wählen hat, ist von der gegenwärtigen Situation abhängig. In einem Zwiegespräch mit einer geringen Distanz zwischen den Personen wird man eher leise sprechen. In einer Rede oder einer Auseinandersetzung, in der man bestrebt ist, seine individuelle Meinung zu vertreten, ist dagegen eine leise Stimme weniger angebracht. Eine übertriebene Lautstärke wird negativ beurteilt, da einerseits zu lautes Sprechen sehr hektisch und aufdringlich, andererseits unangemessenes leises Sprechen sehr unsicher und inkompetent wirken kann. „Eine kräftige, laute aber nicht zu laute Stimme wird meist als Zeichen von Vitalität, Dominanz und Extravertiertheit angesehen" (ebenda, S. 42).

Während der telefonischen Befragung sollte der Interviewer demnach laut und deutlich reden, um es dem Befragten einerseits möglich zu machen die Fragestellungen akustisch gut zu verstehen und andererseits, um einen kompetenten, interessierten und kontaktfreudigen Eindruck zu machen.

Sprechtempo

Vom Interviewer wird ebenfalls verlangt, eine angemessene Sprechgeschwindigkeit zu gebrauchen. Zu schnelles Sprechen führt zu Schwierigkeiten bezüglich der akustischen und inhaltlichen Verständlichkeit der Fragen und kann demzufolge Missverständnisse zwischen Hörer und Sprecher verursachen. Zu langsames Reden hingegen wirkt oft gelangweilt und zu wenig engagiert und erzeugt häufig beim Zuhörer nur kurzfristig Aufmerksamkeit (vgl. Amon 2000, S. 113).

Damit die als positiv eingeschätzten Stimmeigenschaften völlig zum Tragen kommen, hat der Sprecher bzw. der Interviewer auf seine Körperhaltung zu achten. Der Gesprächspartner sieht den Interviewer zwar nicht, aber er hört dessen Körperhaltung, weil diese nachhaltig die Stimme beeinflusst. Sitzt ein Interviewer krumm oder gelangweilt, wirkt die Stimme gepresst und angespannt. Sitzt er hingegen gerade und entspannt, klingt die Stimme voller und etwas tiefer.

Die Wirkung der Stimme kann nicht nur durch die Körperhaltung, sondern auch durch das Telefon an sich negativ beeinflusst werden. Die Stimme am Telefon klingt „blechern" und somit unfreundlicher und unpersönlicher. Man muss davon ausgehen, dass das Telefon ca. zwei Drittel ihrer Intensität wegfiltert (ebd. S. 140). Das heißt für den Interviewer, stets zu lächeln und die Intensität seiner Stimme zu steigern, indem er beispielsweise gestikuliert.

Die Beurteilung der angeführten Stimmeigenschaften und Sprechweisen zeigt sich auch in den Analysen bezüglich telefonischer Interviews. Untersuchungen haben ergeben, dass zögerndes Sprechen und eine leise, drucklose Stimme eher zu Verweigerungen führen (vgl. Friedrichs 1990b, S. 416). Eine der wenigen Untersuchungen über den Einfluss paralinguistischer Merkmale des Interviewers auf den Verlauf des Erhebungsgesprächs wurde von Oksenberg/Cannell (1988)

durchgeführt. Sie fanden heraus, dass die Rate der Verweigerungen sinkt, wenn der Interviewer mit einer relativ hohen Sprechgeschwindigkeit und einer angenehmen Lautstärke spricht und seine Stimme hohes Vertrauen erweckt. Zudem sinkt die Verweigerungsrate, wenn mit einer hohen Sprechgeschwindigkeit begonnen und mit einer geringeren geendet wurde und stieg im umgekehrten Fall. Bei Untersuchungen dieser Art wurde festgestellt, dass die positive und negative Beurteilung der Merkmale der Stimme sehr einheitlich erfolgte und nach relativ kurzer Zeit stattfand. So konnte auch geklärt werden, warum 40 Prozent der Verweigerungen bereits während der ersten Sätze, 50 Prozent vor der ersten Frage und nur 10 Prozent während des Interviews stattfanden (vgl. Friedrichs 1990b, S. 417). Hier wird deutlich, dass die paralinguistischen Merkmale des Interviewers die Befragung in besonderer Weise steuern. „Die Stimme wirkt in der Kommunikation wie ein Schlüsselreiz, der innerhalb von Sekunden darüber entscheidet, ob und wie wir beim anderen ankommen“ (Amon 2000, S. 23).

2.2.1.2 Rhetorische Fähigkeiten

In der Präsentation der Fragestellung sind weitere vom Interviewer ausgehende Fehlerquellen zu sehen. Von ihm wird erwartet, dass er die vom Forscher vorformulierten Fragen ohne Modifikationen rezitiert. Aufgrund dessen wird von ihm verlangt, die Fragenstellungen und auch Antwortkategorien wortwörtlich, flüssig und ohne größere Pausen vorzulesen. Da jedoch das Vorlesen weniger authentisch ist als das freie Sprechen, sollte der Interviewer versuchen, es so zu gestalten, dass es dem freien Sprechen sehr nahe kommt.

Um diese Anforderung erfüllen zu können, ist darauf zu achten, die folgenden Probleme beim Vorlesen zu vermeiden (vgl. Wachtel 1998, S. 30 ff.):

- Häufige unbeabsichtigte Betonungen: Die Ursache dafür ist meist die Betonung eines Satzes in sinnwidrigen Schritten. Es scheint, als erkenne der Sprecher selbst den Sinn erst beim Vortragen.
- Unangemessenes, überhöhtes Sprechtempo: Ein hohes Sprechtempo ist nur angemessen, wenn die Betonung und die Pausen exakt auf den Inhalt und Sinn der Frage oder des vorzutragenden Textes abgestimmt sind.
- Bewusstloses Vorlesen: Das erfolgt, wenn der Text heruntergelesen wird, ohne dass der Sprecher den Inhalt erfasst.
- Atemnot: Bei falscher Atmung werden die Sätze anfangs intensiv, laut und schnell vorgetragen und verlieren am Ende ihre Intensität.
- Gleich klingende Satzmelodie: Zu kurze Atempausen lassen die Tonhöhe der Stimme stetig absinken und bewirken eine permanent gleiche Melodiebewegung.
- Gleichartige Betonung: Hier erfolgt das Absenken und Anheben der Stimme grundsätzlich und nicht auf den Sinn bezogen.

- Fehlende Pausen: Die Aussparung von Pausen vernachlässigt die Gliederung zwischen den verschiedenen Sinnschritten eines Textinhaltes.
- Zu kurze Pausen: Dem Sprecher kommen seine Pausen meist länger vor als dem Hörer und sie fallen demnach sehr knapp aus. Zu kurze Pausen machen es dem Sprecher jedoch nicht möglich, sich auf den nächsten Satz vorzubereiten.

Gutes Vorlesen beherrschen heißt in erster Linie darauf zu achten, immer dem Sinn entsprechend vorzutragen, um dem Befragten die Verständlichkeit der Frage zu erleichtern. Wichtig dabei ist die Fähigkeit, den Text in Pausen gliedern zu können. „Nur diese richtige Gliederung – nach jedem Sinnschritt eine Pause – lässt den Hörer verstehen" (Wachtel 1998, S. 51).

Solche Sinnpausen sind gleichzeitig auch Atempausen. Sind die formulierten Sätze jedoch sehr lang, kann der Sprecher auch während eines Sinnschrittes oder Satzes sehr kurze Atempausen einschieben, die aber nicht den Sinnschritt unterbrechen dürfen. Wie bereits erwähnt ist auch die Modulation der Stimme beim Vortragen der Fragen von wesentlicher Bedeutung. Um durch übermäßige und falsche Betonung nicht in einen ‚Singsang' zu verfallen, muss diese dem ursprünglich gemeinten Sinn entsprechen.

Es zeigt sich, dass die oberste Prämisse für gutes Vorlesen ein sinnerfassendes und somit vorausschauendes Lesen ist. Das verlangt vom Interviewer ein genaues Studium des Fragebogeninhaltes. Kann er die hier angegebenen Voraussetzungen erfüllen, ist es ihm möglich, die Fragen optimal zu rezitieren und trotz des wortwörtlichen Vorlesens ein angenehmes und lockeres Gesprächsklima zu schaffen? Natürlich sollte auch der Forscher bei der Frageformulierung darauf achten, diese möglichst einfach zu gestalten, um eine Modifikation der Frage durch den Interviewer nicht herauszufordern. Zudem muss eine ‚gesprochene Sprache' verwendet werden (vgl. Anders 1990, S. 430). Werden die Fragen verständlich, kurz und präzise formuliert, erhöht sich die Chance auf ihre exakte Wiedergabe und können demzufolge Antwortverzerrungen, welche durch die Modifikation der Fragen entstehen, reduziert werden.

Den Pausen kommt im Telefoninterview eine besondere Bedeutung zu. Sie sollen nicht nur den Text in Sinnschritte gliedern, sondern auch dem Befragten die Möglichkeit zum Nachdenken geben. Längere Pausen im Telefoninterview, die beispielsweise offenen Fragen folgen, werden vom Interviewer meist zu kurz gehalten. Dem Befragten bleibt zu wenig Zeit, seine Antwort zu überdenken bzw. zu ergänzen. In der persönlichen Befragung ist es leichter, längere Pausen durch Mimik und Gestik zu überbrücken. Da dies im telefonischen Interview nicht möglich ist, weiß der Interviewer meist nicht, ob er abwarten oder nachfragen soll. Das hat zur Folge, dass bei Befragungen am Telefon die Anzahl der Angaben bei offenen Fragen geringer ausfällt (vgl. auch Friedrichs 1990b, S. 418).

3 Rekrutierung und Schulung der Interviewer

Die Auswahl und auch die Schulung der Interviewer verdient unter Qualitätsgesichtpunkten besondere Aufmerksamkeit, da die Fähigkeiten der Interviewer einen wesentlichen Beitrag zur Qualitätssicherung in Telefoninterviews leisten. Wie bereits im vorangegangenen Abschnitt erläutert wurde, hat der Interviewer einen maßgeblichen Anteil am Antwortverhalten sowie der Kooperationsbereitschaft der Befragten. Um Interviewereffekte weitgehend zu neutralisieren und auf einem möglichst geringen Niveau zu halten, um somit eine hohe Datenqualität und Ausschöpfungsrate zu erlangen, werden die Interviewer von den Supervisoren und Projektleitern nach bestimmten Kriterien rekrutiert und geschult. Dabei gibt es keine festen Kriterien für die Rekrutierung und Vorgehensweisen von Schulungen der Interviewer, wodurch in der Praxis unterschiedliche Methoden angewendet werden.

3.1 Rekrutierung der Interviewer

Kommerzielle Umfrageinstitute versuchen meist durch Zeitungsannoncen geeignete Bewerber zu finden. Die Inserate sollten psychologisch so verfasst sein, dass sie kontaktfreudige Menschen aus allen sozialen Schichten und Altersgruppen ansprechen (vgl. von Kirschhofer-Bozenhardt/Kaplitza 1991, S. 127).

Das CATI-Labor des zsh, welches nahezu ausschließlich Studenten für Interviewertätigkeiten rekrutiert, versucht beispielsweise durch Aushänge in zentralen universitären Einrichtungen sowie durch eine gezielte Kontaktaufnahme mit Teilnehmern methodisch orientierter Veranstaltungen, Bewerber zu gewinnen. Die Mundpropaganda zwischen den im CATI-Labor bereits beschäftigten Interviewern und anderen Studierenden spielt bei der Rekrutierung eine nicht zu unterschätzende Rolle.

Die meisten der rekrutierten Interviewer sind Frauen. Das liegt unter anderem daran, dass ihr Interesse an flexiblen Arbeitszeiten und Teilzeitjobs sehr groß ist. Auch im universitären Bereich, wo solche Gründe keine Rolle spielen, ist eine Dominanz der Frauen zu verzeichnen. Plausible Gründe für die Überrepräsentativität der Frauen könnten zum einen in dem größeren Anteil weiblicher Studenten liegen oder zum anderen in der Annahme, dass Telefonieren im weitesten Sinn als Frauenarbeit gilt. Beobachtungen des Interviewerstamms im CATI-Labor des zsh zeigen allerdings, dass der Anteil der männlichen Interviewer in den letzten Jahren stetig zugenommen hat. Momentan sind bereits ca. 40 Prozent der Interviewer männlich.

Für die Projektleiter empfiehlt es sich, möglichst viele Interviewer zu rekrutieren, damit alle Erhebungszeiten auch tatsächlich abgedeckt werden können. Unsere Erfahrungen zeigen, dass die Interviewer aufgrund der abnehmenden Konzentration nicht länger als vier Stunden hintereinander arbeiten sollten. Das setzt

natürlich einen großen Interviewerstab voraus. Dieser ist auch von Vorteil, um beispielsweise bei sehr anspruchsvollen Projekten die Qualifiziertesten unter den Interviewern auswählen oder um das inhaltliche Interesse eines Interviewers für eine bestimmte Studie oder ein bestimmtes Thema berücksichtigen zu können.[3]

In der Literatur wird davon gesprochen, dass mindestens zwei- bis dreimal soviel Interviewer zur Verfügung stehen sollten, wie Telefonplätze im CATI-Labor eingerichtet sind (vgl. Frey/Kunz/Lüschen 1990). Um die vorhandenen Plätze des CATI-Labors des zsh effektiv nutzen zu können muss nach unseren Erfahrungen die Intervieweranzahl vier- bis sechsmal höher sein als die Anzahl der Telefonplätze. Zu diesem Ergebnis gelangten wir vor dem Hintergrund, dass zum einen die Interviewer je nach ihrer individuellen Belastungsfähigkeit nur maximal vier Stunden arbeiten sollten und es sich zum anderen bei der Interviewertätigkeit um eine Nebenbeschäftigung handelt, welche von den Terminen des Studiums abhängig ist. Natürlich ist die notwendige Größe des Interviewerstamms auch abhängig vom Umfang der Befragungszeit pro Tag. So wird für Bevölkerungsbefragungen eine Kernzeit für Interviews von ca. 4 Stunden und für Unternehmensbefragungen von ca. 8 Stunden veranschlagt.

Bei der Auswahl der Interviewer haben die Projektleiter bzw. Supervisoren auf eine Reihe von Merkmalen zu achten. Um welche Merkmale es sich dabei handelt ist bisher noch umstritten. Eine der wichtigsten Voraussetzungen für die Eignung als Interviewer ist jedoch die Stimme. Im vorangegangen Abschnitt wurde die Bedeutsamkeit der Stimme deutlich. Damit es dem Befragten möglich ist, die Fragen akustisch gut zu verstehen und ihnen folgen zu können, muss der Bewerber während des Telefongesprächs in der Lage sein, laut und deutlich zu sprechen. Eine deutliche Aussprache, das Sprechtempo sowie Klang und Modulation der Stimme sind wesentliche Merkmale für die Eignung des Bewerbers bezüglich telefonischer Befragungen.

Die beste Möglichkeit, diese Anforderungen zu prüfen ist – neben einem persönlichen Vorstellungsgespräch – ein reales Telefongespräch durchzuführen. Anders (1990) schlägt vor, die Bewerbungen grundsätzlich telefonisch entgegenzunehmen, um sich einen ersten Eindruck machen zu können. Ein anderer Vorschlag von Dillman (1978) sieht vor, diese Anforderungen in einem Telefongespräch zwischen Interviewer und einer ihm unbekannten Person zu überprüfen, indem anschließend die Stimme des Bewerbers ausgewertet wird. Während eines Probeinterviews kann untersucht werden, ob der Bewerber in der Lage ist, Frage-

3 So können zum Beispiel bei Umfragen mit politischen Inhalten bevorzugt Politikstudenten eingesetzt werden, welche bereits mit verschiedenen Termini vertraut sind und ein durch ihr Studium vermitteltes Hintergrundwissen aufweisen. Auch zeigen die Erfahrungen im CATI-Labor des zsh, dass einzelne Interviewer besser für Unternehmensbefragungen geeignet sind, während andere wiederum mehr Erfolge bei Bevölkerungsbefragungen verzeichnen können.

stellungen flüssig und ohne größere Pausen abzulesen, Antworten und wichtige Anmerkungen genau zu protokollieren sowie Anweisungen zu befolgen, ohne dass der Interviewverlauf dabei gestört wird. „Diese Probeinterviews haben Schulungs- und Testcharakter und sagen sehr viel über die Person des Bewerbers aus, ohne dass man ihn persönlich kennen muss“ (von Kirschhofer-Bozenhardt/ Kaplitza 1991, S. 131).

Eine Persönlichkeitseigenschaft, die jeder Interviewer besitzen sollte, ist die Selbstsicherheit. Da sich diese – wie schon erwähnt – in der Stimme manifestiert, ist bei der Rekrutierung darauf zu achten, dass die Bewerber ein gewisses Selbstbewusstsein besitzen. In Situationen, in denen Interviewer mit unerwarteten Fragen konfrontiert werden und dann sofort eine überzeugende Antwort bereit haben müssen, ist dies besonders wichtig. Zudem bereitet die Schulung vor der jeweiligen Erhebung darauf vor, in schwierigen Situationen angemessen reagieren zu können.

Da Telefoninterviews zum größten Teil computergestützt ablaufen, müssen die Bewerber EDV-Grundkenntnisse besitzen. Es werden zwar in der Interviewerschulung Anweisungen für den Umgang mit dem CATI-Programm gegeben, dennoch sollten sie im Umgang mit der EDV erfahren sein. Nur unter dieser Voraussetzung ist es möglich, sich allein auf das Interview konzentrieren zu können. Diese Anforderung spielt bei der Rekrutierung im universitären Bereich, wie sie im zsh durchgeführt wird, keine wesentliche Rolle, da sie von den Studenten ausnahmslos erfüllt wird. Bei der Rekrutierung von Personen aus verschiedenen sozialen Schichten und Altersgruppen, wie sie in kommerziellen Umfrageinstituten üblich ist, sollte geprüft werden, ob ein versierter Umgang mit dem Computer beim Bewerber vorausgesetzt werden kann.

Im CATI-Labor des zsh werden die Anfragen zu einem großen Teil telefonisch von den Supervisoren entgegengenommen. Während des Telefongesprächs wird bereits auf die Stimme, das Sprechtempo und einen eventuellen Dialekt geachtet. Schätzt man den Bewerber als geeignet ein, wird er gebeten, die wichtigsten Informationen, wie beispielsweise Studiensemester und Studienfachrichtung und seine Kontaktdaten in einer E-Mail festzuhalten und an das zsh zu senden. In einem dritten Schritt wird der Bewerber zu einem kurzen Vorstellungsgespräch eingeladen. In diesem Gespräch geht es darum, den Interviewer persönlich kennenzulernen und offene Fragen zur Qualifikation (z. B. Fremdsprachenkenntnisse, bereits vorhandene Erfahrungen als Interviewer) zu klären. Dabei ist zu prüfen, wie sich der Bewerber artikulieren kann. Nicht nur die Stimme, sondern auch der persönliche Eindruck ist entscheidend für die endgültige Auswahl der Interviewer. Dies ist notwendig, da der Bewerber nicht nur seine Rolle als Interviewer einwandfrei erfüllen soll, sondern auch als Mitglied des Interviewerstamms in der Lage sein muss, Rücksicht auf seine Kollegen zu nehmen und die Vorgaben und Anweisungen der Supervisoren zu befolgen.

Die tatsächliche Qualität des Interviewers lässt sich jedoch meist erst während der Erhebung einschätzen. Erst dann können sich die Supervisoren ein Bild über die Verlässlichkeit, Disziplin, Motivation und Produktivität des Interviewers machen. Es ist ihnen möglich, die jeweilige individuelle Ausschöpfungsquote des Interviewers, Freundlichkeit und Höflichkeit im Umgang mit dem Befragten, das Protokollieren von offenen Fragen usw. zu prüfen.

3.2 Schulung der Interviewer

Bewerber, welche die nötigen Anforderungen erfüllen, erhalten eine Interviewerschulung. Die Qualität der Schulung ist für den Erfolg einer Befragung entscheidend, da sie die Interviewer auf die Erhebungssituation vorbereitet.

Interviewerschulungen laufen zeitlich und inhaltlich sehr unterschiedlich ab. Werden Telefoninterviews mit Hilfe des CATI-Systems durchgeführt, müssen die Interviewer zusätzlich mit den technischen Gegebenheiten sowie mit dem programmierten Fragebogen vertraut gemacht werden.

Im CATI-Labor des zsh wird die Schulung der Interviewer in drei Phasen, der Grundschulung, der umfragespezifischen Schulung und der sprechwissenschaftlichen Schulung, durchgeführt.

Zur Grundschulung gehören:

- die Vermittlung von Grundlagen bezüglich der Interviewerrolle und der Besonderheiten der telefonischen Befragung;
- das Erläutern von Sprechtechniken;
- der Umgang mit der Interviewersoftware;
- der Umgang mit der Telefonanlage;
- das Aufklären bezüglich der Richtlinien des Datenschutzes und des Persönlichkeitsrechts der Befragten;
- das Aufklären bezüglich der Richtlinien und Regeln für das Arbeiten im CATI-Labor.

In der umfragespezifischen Schulung erfolgen:

- die Einführung in das jeweilige Thema der Befragung;
- die Vermittlung von Eckdaten und Hintergrundwissen der Befragung (Anzahl der Interviews, Stichprobenziehung);
- die Vorstellung des Fragebogens (inklusive der zur Anwendung kommenden Filterführung bzw. Besonderheiten);
- Übungsinterviews, in denen aufkommende Fragen geklärt werden.

Die Schulung aus sprechwissenschaftlicher Sicht beinhaltet:

- die Einführung in die Anatomie und Physiologie der Stimme (stimmliche Probleme (z. B. Heiserkeit) nach längerem Telefonieren, Grundsätze der Stimmhygiene);
- die verschiedenen Elemente des Sprechausdrucks (z. B. Lautstärke, Sprechtempo);
- die Kommunikation am Telefon (rhetorische Perspektive, Ebenen einer Nachricht, Analyse der Situation).

In der Grundschulung werden zum einen ganz allgemeine Probleme und Besonderheiten des telefonischen Interviews besprochen. Dazu gehört in erster Linie, den Interviewern deutlich zu machen, dass es sich bei dieser Art der Befragung um eine Situation handelt, die allein auf den akustischen Kommunikationskanal beschränkt ist. Andererseits ist eine Einweisung in die Interviewtechniken notwendig. Die Interviewer müssen lernen, einen angemessenen Konversationston, balancierten Rapport, zu finden. In dieser Phase werden die Interviewer auf ihre Rolle, welche sie in der Erhebungssituation einnehmen, vorbereitet. Des Weiteren müssen die Interviewer eine technische Einweisung erhalten sowie über die Richtlinien des Datenschutzes und des Persönlichkeitsrechts des Befragten, aber auch über die Arbeitsabläufe im CATI-Labor informiert werden. Die Interviewer müssen wissen, an wen sie sich zu wenden haben, falls Probleme auftauchen oder wem sie Bericht erstatten müssen usw.

Während die Grundschulung den Interviewern eine Einführung in die spezifische Arbeit im Rahmen von CATI-Befragungen gibt und einmalig stattfindet, wird die umfragespezifische Schulung vor jeder neuen Befragung durchgeführt. Hier führt der Forscher die Interviewer in das jeweilige Befragungsthema ein und vermittelt das nötige spezifische Hintergrundwissen. Die Interviewer erhalten Informationen zum Inhalt der Studie und Hinweise zu Besonderheiten der Befragung. Beim gemeinsamen Durchgehen des Fragebogens wird zudem auf fachspezifische Begriffe und Fragen der Interviewer eingegangen.[4] Dafür werden jede einzelne Frage und die dazugehörigen Antwortkategorien gelesen, besprochen und erläutert. Bei Telefoninterviews, die nicht computergestützt ablaufen, ist es besonders wichtig, dabei die Gabel- und Filterstrukturen des Erhebungsinstruments zu besprechen. Dies ist jedoch auch bei computergestützten Telefoninterviews nötig, damit die Interviewer einen Einblick in alle möglichen Fragenstellungen bekommen, um Unsicherheiten zu vermeiden.

Nachdem der Fragebogen ausführlich besprochen und analysiert worden ist, kann er durch Übungsinterviews geprobt werden. Während dieser Testphase

4 Der für die Übung verwendete Fragebogen sollte möglichst der letzten Fassung entsprechen, bevor er für die Untersuchung freigegeben wird.

kann das Interview für auftretende Fragen unterbrochen werden und Problemen oder Verbesserungsvorschlägen nachgegangen werden.

Zudem hat es sich als sinnvoll erwiesen, allen Interviewern ein studienspezifisches Merkblatt auszuhändigen, auf dem die wichtigsten Informationen der jeweiligen Studie nochmals festgehalten werden. Mit dieser Vorbereitung auf eine CATI-Befragung können beim Erstkontakt Rückfragen der Befragungsperson zum Inhalt und Ziel der Befragung schneller geklärt werden. Die Interviewer müssen Grundkenntnisse zum jeweiligen Thema besitzen, um kompetent auf den Befragten wirken zu können.

Vor allem in der Eröffnungsphase, auch wenn diese zum größten Teil standardisiert ist, sollten die Interviewer den Befragten entsprechende zusätzliche Informationen geben können. Häufig wird zum Beispiel nachgefragt, wie die Telefonnummer des Befragten ermittelt wurde oder welchem Zweck die Untersuchung dient. Informationen zum Sinn und Zweck der Studie helfen dem Interviewer, den Befragten von seiner Teilnahme zu überzeugen.

Diese Schulung sollte ca. eine Woche vor der tatsächlichen Feldphase liegen, damit eine gründliche Einarbeitung in die jeweilige Thematik erfolgen kann. Die Übungsphase mit dem fertigen Fragebogen sollte nicht zu kurz gehalten werden, um Einarbeitungsverzerrungen während der ersten Feldtage zu vermeiden. Der Interviewer sollte den Fragebogen mindestens einmal komplett und selbstständig durchgehen. Tauchen dennoch Fragen auf stehen den Interviewern Supervisoren zur Verfügung. Bemerken die Supervisoren Schwächen bei einigen Interviewern, kann jederzeit eine Nachschulung erfolgen.

Neben der dargestellten Grundschulung und inhaltlichen Schulung der Interviewer wird zusätzlich noch eine Schulung zum Führen von Telefoninterviews aus sprechwissenschaftlicher Sicht durchgeführt. Die Schulung findet in unregelmäßigen Abständen ca. ein- bis zweimal pro Jahr statt. Inhaltliche Schwerpunkte dieser Weiterbildung sind beispielsweise eine Einführung in die Anatomie und Physiologie der Stimme, Elemente des Sprechausdrucks und die Kommunikation am Telefon.[5] Die Interviewer erlernen hier, wie mit stimmlichen Problemen (z. B. Heiserkeit nach längerem Telefonieren) umgegangen werden muss, aber auch, wie Elemente des Sprechausdrucks im Rahmen der Kommunikation am Telefon anzuwenden sind. Die Interviewer werden in die Begriffe des Sprechdenkens und Hörverstehens eingeführt, um somit besser auf den Befragten eingehen zu können. Besonders in den ersten Minuten der Kontaktaufnahme ist es wichtig, das Vertrauen der Befragten zu gewinnen, um die Teilnahmebereitschaft bei telefonischen Befragungen zu fördern. Die Interviewer müssen in der Lage sein, in den

5 Diese Schulung wird von dem Institut für Sprechwissenschaften der Martin-Luther- Universität Halle-Wittenberg durchgeführt.

verschiedenen Gesprächssituationen – besonders aber in der Einleitungsphase – kompetent zu reagieren.

Die Rekrutierungs- und Schulungsanstrengungen sind ein sehr bedeutsames Kriterium zur Sicherstellung und Steigerung der Qualität telefonischer Interviews. In der Literatur wird jedoch immer wieder davon berichtet, dass eine spezifische Interviewerpersönlichkeit existiert, die auch nicht erlernt werden kann. Wiederum gibt es für jeden spezifischen Umfragetyp (Managerbefragungen, Bevölkerungsbefragungen, Experteninterviews) unterschiedliche Befähigungen.

Trotz einer sorgfältigen Rekrutierung und Schulung der Interviewer können, wie bei jeder anderen Befragungsform, auch im Rahmen von Telefoninterviews Interviewereinflüsse nicht ganz ausgeschaltet werden. Allerdings können diese aufgrund des zentral geführten Interviewereinsatzes vom Supervisor besser kontrolliert werden.

Wie bereits angesprochen, gibt es noch weitere Möglichkeiten, die Qualität von Interviews messen und kontrollieren zu können. Dies kann zum einen mittels Ausschöpfungsquoten erfolgen, aber auch durch sogenannte Nachbefragungen, was in den folgenden Abschnitten erläutert werden soll.

4 Die Ausschöpfungsquote als Qualitätsmaßstab

Die Ausschöpfungsquote gilt als ein wichtiger Qualitätsmaßstab für Umfragen. Allerdings müssen die Ausschöpfungsquoten von CATI-Erhebungen sehr differenziert betrachtet werden, da sie immer vom Kontext der jeweiligen Studie und den Vorgaben im Rahmen des Projektes abhängig sind. Allgemein kann gesagt werden, dass die Ausschöpfung den Grad der Realisierung einer Stichprobe beschreibt. Es haben sich gewisse Konventionen für die Berechnung von Ausschöpfungsquoten entwickelt. Danach wird die Ausschöpfung in Anlehnung an ZUMA[6] als das Verhältnis der Anzahl der ausgewerteten Interviews zur Größe der bereinigten Stichprobe definiert (vgl. Porst 1993, S. 5). Diese Definition wird auch in den folgenden Ausführungen zugrunde gelegt.

4.1 Die inhaltliche und methodisch formale Dimension von Ausschöpfungsquoten

Die Ausschöpfungsquote ist abhängig von verschiedenen Faktoren, die in zwei Kategorien eingeteilt werden können: in eine inhaltliche Dimension und in eine methodisch formale Dimension. Die folgende Gruppierung ist ein Vorschlag für die Einordnung der Faktoren in die genannten zwei Dimensionen.

6 ZUMA: Zentrum für Umfragen, Methoden und Analysen

Zur inhaltlichen Dimension gehören folgende Faktoren:

- die zu befragende Zielgruppe[7],
- das Thema der Befragung,
- die persönliche Betroffenheit,
- die gesellschaftliche Brisanz.

Die Ausschöpfungsquote variiert in Abhängigkeit von den Merkmalen der Zielgruppe. Bei Telefonumfragen beispielsweise, deren Zielgruppe Jugendliche sind, gibt es einige Unterschiede im Vergleich zu CATI-Befragungen der erwachsenen Bevölkerung. Nach Erfahrungen im zsh sind Jugendliche zwar durch ihre Mobilität schwerer erreichbar, aber leichter als Erwachsene für eine Telefonbefragung zu gewinnen, was zu einem großen Teil daran liegen könnte, dass diese Altersgruppe gern telefoniert. Außerdem sind junge Teilnehmer noch nicht in dem Maße „befragungsmüde", wie es bei der erwachsenen älteren Bevölkerung der Fall sein mag, zum großen Teil bedingt durch die vielen Anrufe kommerzieller Institute (vgl. Porst 1993, S. 3 sowie Fuchs 1995, S. 284).

Das Thema der Befragung beeinflusst die Teilnahme an einer Untersuchung insofern entscheidend, als die zu Befragenden eher an der Studie teilnehmen, wenn sie das Thema interessant finden oder persönlich betroffen sind und ihre Erfahrungen gern weitergeben. Wenn der Kontext der Befragung zur Erhebungszeit ein Thema in der Öffentlichkeit ist, so kann sich das allgemeine Interesse positiv auf die Teilnahme an der CATI-Erhebung auswirken.

Zur methodisch formalen Dimension sind folgende Faktoren zuzurechnen:

- die Art der Stichprobenziehung,
- die Größe der Stichprobe,
- die Befragungszeit bzw. der Befragungszeitraum,
- das Versenden eines Anschreibens bzw. die Mitteilung über die Presse (Kalt- und Warmkontaktierung),
- die Interviewdauer,
- Interviewerschulungen,
- die gewünschte Anzahl an Interviews und das (vorzeitige) Beenden der Befragung nach dem Erreichen der geplanten Anzahl an Interviews,
- die Tatsache, dass es sich um eine Panelbefragung mit Einverständnis zur Wiederholungsbefragung handelt.

Die Stichprobenziehung kann auf verschiedene Art erfolgen; für Bevölkerungsbefragungen zum Beispiel über das Einwohnermeldeamt, nach dem Gabler/Hä-

7 Es wird zwischen folgenden Befragungsarten unterschieden: Bevölkerungsbefragungen, Jugendbefragungen, Unternehmensbefragungen und Befragungen von Eliten.

der-Design[8] oder aus dem amtlichen Telefonbuch. Bei der Stichprobenziehung aus dem amtlichen Telefonbuch werden Ausschöpfungsquoten von 40 - 60 % (vgl. Porst 1993, S. 27) angegeben, sie können aber auch deutlich darunter oder darüber liegen[9]. Im Rahmen von Unternehmensbefragungen können die Betriebe für die Stichprobe aus der Betriebsnummerndatei der Bundesagentur für Arbeit beantragt oder aus dem Hoppenstedt-Verzeichnis gewonnen werden. Sollen bestimmte ausgewählte Personengruppen befragt werden – so beispielsweise bei Elitenbefragungen –, kann durch Recherchen im Internet oder durch Informationen der Verwaltungsämter die Stichprobe erstellt werden.

Das Versenden eines Anschreibens[10] im Vorfeld der Untersuchung (vgl. Porst 1991, S. 10 f. und Porst 1993, S. 27) mit der Ankündigung der Befragung und anderen wichtigen Informationen oder die Bekanntgabe der Studie in den Medien kann die Ausschöpfungsquote erheblich erhöhen. Allerdings muss an dieser Stelle erwähnt werden, dass die Ausschöpfungsquote bei Jugend- und anderen Bevölkerungsbefragungen nach Erfahrungen des zsh bei sogenannten „kalten Kontakten“ ohne jede Vorankündigung auch akzeptable Werte aufweisen kann, wenn der Befragungsperson das Thema der Befragung interessant erscheint.

In diesem Zusammenhang ist auch die Befragungsdauer von Bedeutung. Werden bei Bevölkerungsbefragungen bis zu 20 Minuten Interviewlänge als akzeptabel eingeschätzt, so wird im Rahmen von Eliten- und Unternehmensbefragungen eine längere Befragungsdauer (bis zu 60 Minuten, in einigen Fällen auch bis zu 90 Minuten) durchaus akzeptiert, wenn im Vorfeld eine Terminvereinbarung getroffen wurde oder die Möglichkeit besteht, das Interview zu splitten.

Um die Teilnahmebereitschaft bei telefonischen Befragungen zu fördern, ist es wichtig, das Vertrauen der Befragten in den ersten Minuten der Kontaktaufnahme zu gewinnen. Die Interviewer müssen in der Lage sein, in den verschiedenen Gesprächssituationen – besonders aber in der Einleitungsphase – kompetent zu reagieren. Neben den inhaltlichen Einführungen vor jeder beginnenden Studie nehmen die Interviewer aus diesem Grund an Schulungen zum Führen von Telefoninterviews aus sprechwissenschaftlicher Sicht teil.

---

8 Mit diesem Verfahren können diejenigen Haushalte in die Auswahl einer Telefonbefragung einbezogen werden, die nicht im Telefonbuch verzeichnet sind. Dabei werden unabhängig vom Telefonbuch Ziffernfolgen zufällig derart generiert, dass sie der Struktur von Telefonnummern entsprechen. Vgl. zu diesem Verfahren Häder 1994 und Gabler et al. 1998, S. 55 ff.

9 Porst (1991) beispielsweise berichtet im Rahmen einer Befragung in Mannheim von Ausschöpfungsquoten zwischen 26 und 39 %. Dabei wurden sogar verschiedene Möglichkeiten vorheriger Kontaktaufnahme angewandt (kein Anschreiben, kurzes Anschreiben, langes Anschreiben).

10 Bei sozialwissenschaftlichen CATI-Untersuchungen von Betrieben bietet sich die Möglichkeit des Versendens eines Anschreibens an; schwieriger gestaltet es sich bei Bevölkerungsumfragen, deren Stichprobe aus dem amtlichen Telefonbuch oder Telefon-CD-Rom gezogen wird. In diesem Fall kann die Ankündigung der Befragung über Medien wie Funk, Fernsehen oder Presse erfolgen.

Bei der Betrachtung der Ausschöpfungsquote muss auch die Tatsache beachtet werden, dass Telefonstudien nicht immer auf das Erzielen hoher Ausschöpfungsquoten, sondern häufig auf das rasche Bearbeiten festgelegter Fallzahlen ausgelegt sind (vgl. dazu auch Porst 1993, S. 26). In diesem Fall könnte die Ausschöpfungsquote höher sein, wenn nicht die Befragung durch das Erreichen der Anzahl angestrebter kompletter Interviews bzw. aus Kostengründen abgeschlossen worden wäre. Die Größe der Stichprobe spielt dabei eine wichtige Rolle.

Häufig ist nicht das Erreichen einer bestimmten Fallzahl das Ziel einer CATI-Befragung, sondern aus der vorhandenen begrenzten Stichprobe eine möglichst große Anzahl kompletter Interviews zu realisieren, möglicherweise auch mit dem Kompromiss, den Befragungszeitraum auszudehnen und die Kontakthäufigkeit zu erhöhen (über das zur Effizienz einer CATI-Studie empfohlene Maß hinaus).

Handelt es sich um eine Panel-Befragung und ist das Einverständnis der Befragungsperson bereits eingeholt, kann davon ausgegangen werden, dass die Ausschöpfungsquote in den Folgebefragungen sehr hoch ist.

Die verschiedenen Befragungsarten werden durch die genannten Faktoren unterschiedlich stark beeinflusst. Je nach Kombination der für die jeweilige Befragung relevanten Faktoren aus den beiden Dimensionen wird das Ausschöpfungsergebnis in Abhängigkeit vom Befragungstyp beeinflusst. Während bei Bevölkerungs- und Unternehmensbefragungen eine große Spannweite an Ausschöpfungsquoten vorherrscht, ist die Ausschöpfung bei den Elitenbefragungen generell erfolgreicher.

Im Folgenden sollen drei Befragungsarten gegenübergestellt und untersucht werden. Zur Auswertung werden Ergebnisse von computergestützten Bevölkerungs- und Unternehmensbefragungen und Befragungen politischer Eliten herangezogen, welche im CATI-Labor des zsh durchgeführt wurden.[11]

Diese drei Befragungsarten werden – auch in Abhängigkeit von der Anzahl der Kontakte – verglichen in Bezug auf die Ausschöpfungsquote[12] und die Ver-

11 In die Auswertung sind 11 Studien mit insgesamt knapp 14.000 abgeschlossenen Interviews einbezogen, die in den Jahren 2002 bis 2006 durchgeführt wurden.

12 An der Stelle soll darauf hingewiesen werden, dass im Rahmen der Auswertungen der zsh-Studien folgende Dispositionen zu den stichprobenneutralen Ausfällen gerechnet werden: falsche Telefonnummer bzw. gar kein Anschluss unter der angewählten Nummer, Faxanschluss oder andere technische Probleme. Außerdem beinhalten die stichprobenneutralen Ausfälle auch Verständigungsprobleme, Besetztzeichen, kein Kontakt bzw. Anrufbeantworter, Terminvereinbarungen zu einem späteren Zeitpunkt bzw. nicht genutzte Termine und „sonstiges" (verstorben, in Haft o. ä.). Zur Netto-Stichprobe zählen die Verweigerungen wie „kein Interesse", „überhaupt keine Zeit", „kein telefonisches Interview" und „ohne Antwort aufgelegt", die Abbrüche im Verlauf des Interviews und die erfolgreichen Interviews. Andere Auswertungen zu Ausschöpfungen können in den Angaben der Quote variieren, wenn die Zuordnung der Dispositionen zu den stichprobenneutralen Ausfällen bzw. zur Netto-Stichprobe in anderer Form gewählt wurde.

weigerungsquote. Dabei werden empirische Erfahrungen zur Erhöhung der Ausschöpfungsquote bei telefonischen Umfragen aufgezeigt, die zum einen der inhaltlichen Dimension und zum anderen der methodisch formalen Dimension zugehörig sind.

4.2 Ausschöpfungsquoten der CATI-Bevölkerungsbefragungen im Vergleich zu Eliten- und Unternehmensbefragungen

Im Rahmen der inhaltlichen Dimension von Ausschöpfungsquoten sind die Zielgruppe, das Befragungsthema, die persönliche Betroffenheit und die gesellschaftliche Brisanz für die Teilnahme an einer Erhebung von Bedeutung. Bei den Bevölkerungsbefragungen, die im CATI-Labor des zsh durchgeführt wurden, sind Ausschöpfungsquoten zwischen 20 und 88 Prozent zu verzeichnen. Die geringste Ausschöpfungsquote hatte eine Bevölkerungsbefragung zur Politikwahrnehmung (20 %). Die Gründe dafür sind in der Art der Stichprobenziehung (modifiziertes Gabler/Häder-Design) und im Befragungsthema zu suchen[13].

Bei der CATI-Befragung von Betroffenen der Flutkatastrophe 2002 sowie bei der Befragung von Migranten und bei der Erhebung im Rahmen von Existenzgründungen wurde das Thema der Umfrage sehr interessant empfunden[14], was die Teilnahme – trotz einer Interviewdauer von zum Teil 30 Minuten – begünstigte (vgl. Tabelle 1). Die große Spannbreite an Ausschöpfungsquoten bei CATI-Bevölkerungsbefragungen, die im CATI-Labor des zsh zu verzeichnen sind, trifft für die Elitenbefragungen nicht zu. Hier finden sich Ausschöpfungsquoten zwischen 67 und 80 Prozent, welche als sehr hoch eingestuft werden können. Aber auch bei dieser Art von telefonischer Erhebung muss bemerkt werden, dass unterschiedliche Quoten durch die Besonderheiten der einzelnen Untersuchungen bestimmt werden (vgl. Tabelle 1).

Wird die methodisch formale Dimension betrachtet, so haben u. a. das Versenden eines Anschreibens, der Erstkontakt und die Interviewdauer einen Einfluss auf die Teilnahmebereitschaft zum Gespräch. Wurde vor der Untersuchung ein Anschreiben verschickt, so gestaltete sich die Kontaktaufnahme einfacher, da die Zielpersonen bereits über die wissenschaftliche Studie informiert waren. Nach den Erfahrungen, die im Rahmen von Elitenbefragungen im CATI-Labor des zsh gesammelt werden konnten, ist es in jedem Fall notwendig, ein Anschreiben zu

13 Im Rahmen eines Projektes im zsh wurde versucht, die Verweigerungsquote durch unterschiedliche Einstiegstexte zu minimieren. In diesem Fall war es interessant, dass die Verweigerungsquote am niedrigsten war, wenn die Interviewer nur Stichpunkte zum Eingangstext vorgelegt bekamen und den Text selbst formulieren sollten. Diese Tatsache hat weiterführende Überlegungen zur Gestaltung der Kontaktaufnahme mit einem optimalen Eingangstext notwendig gemacht.

14 Die Tatsache, dass das Thema der Befragung sehr interessant empfunden wurde, gab ein Großteil der Befragten am Ende des Interviews in einer offenen Frage als Kommentar an.

versenden, um die Inhalte der Studie vorzustellen sowie die zu Befragenden von der Bedeutsamkeit der wissenschaftlichen Untersuchung und auch von der Seriosität des Forschungsinstituts zu überzeugen.

Tabelle 1: Ausschöpfungs- und Verweigerungsquoten von telefonischen Bevölkerungsbefragungen und CATI-Befragungen von politischen Eliten

Studie	Thema	Interviewdauer (Min.)	Ausschöpfung (%)	Verweigerung[15] (%)	Art der Stichprobenziehung	Vorab-Informationen
Bevölkerungsbefragungen						
SFB 580 TP A4: Bevölkerungsbefragung	Befragung zur Politikwahrnehmung	30	20	79	modifiziertes Gabler/Häder-Design	Pressemitteilung
Haushaltsbefragung in ländl. Regionen	Soziale Probleme in 3 ländl. Regionen	25	Ort 1: 56 Ort 2: 43 Ort 3: 27	40 56 71	Telefonbuch	Aushänge in den Orten
Befragung von Betroffenen der Flutkatastrophe	Situation Flutbetroffener	30	72	27	Datenbank der Hilfsorganisationen	Anschreiben
Befragung von Migranten	Ost-West-Migration junger Menschen	20	75	24	Einwohnermeldeamt	keine
Befragung von Existenzgründern	Erfolg und Misserfolg von Ich-AGs	10	88	10	Arbeitsagentur	keine
Befragungen politischer Eliten						
SFB 580 TP A4: Elitenbefragung	Lebenslauf jetziger und ehemaliger Politiker	90	1. Welle: 80 2. Welle: 76	18 21	Recherche im Internet und bei Verwaltungsämtern	Anschreiben
SFB 580 TP A6: Befragung von KWGs	KWG in Ost- und Westdeutschland	60	67	31	Recherche im Internet und bei Verwaltungsämtern	E-Mail
SFB 580 TP A6: Befragung von Fraktionsvorsitzenden	Kommunalpolitik in Deutschland	70	75	20	Recherche im Internet und bei Verwaltungsämtern	E-Mail

SFB = Sonderforschungsbereich
TP = Teilprojekt

15 Ergibt die Summe aus Ausschöpfungs- und Verweigerungsquote nicht 100 Prozent, so ist dies auf Abbrüche während des Interviews zurückzuführen.

Der Erstkontakt gestaltete sich bei den Elitenbefragungen noch relativ problemlos, allerdings mussten – oft mehrfach – Termine für ein Interview vereinbart werden. Das Zeitbudget der politischen Eliten ist in den meisten Fällen sehr knapp, wodurch Terminabsprachen schwierig sind und Terminverschiebungen keinen Einzelfall darstellen. Erwähnt werden muss an dieser Stelle auch die Interviewdauer, welche bei den dargestellten Elitenbefragungen zwischen 60 und 90 Minuten betrug (vgl. Tabelle 1). Dieser Umstand machte es – häufiger als bei Bevölkerungsbefragungen – erforderlich, ein Interview zu splitten, d.h. das Interview zu unterbrechen und zu einem späteren Zeitpunkt zu Ende zu führen.

Die Heterogenität an Ausschöpfungsquoten bei Bevölkerungsbefragungen, die im CATI-Labor des zsh zu verzeichnen sind, gilt auch für die Unternehmensbefragungen. Hier sind Ausschöpfungen zwischen 25 und 75 Prozent zu verzeichnen. Aber auch bei dieser Art von telefonischer Erhebung muss bemerkt werden, dass die unterschiedlichen Quoten durch die Besonderheiten der einzelnen Untersuchungen bestimmt werden. Die Kontaktaufnahme gestaltet sich bei größeren Unternehmen weitaus schwieriger als bei Kleinbetrieben, da man z. B. erst über eine Zentrale zur Kontaktperson verbunden wird. Das Zeitbudget der Mitarbeiter in Unternehmen ist in den meisten Fällen sehr knapp, wodurch Terminabsprachen schwierig sind und Terminverschiebungen oft vorkommen. Wurde vor der Befragung ein Anschreiben verschickt, so ist dieses häufig nicht bei der Zielperson angekommen.

Die Bedeutung der Befragungszeit für die Ausschöpfung einer Befragung darf nicht unterschätzt werden. Bei Unternehmensbefragungen empfiehlt es sich, bereits um 8.00 Uhr morgens mit dem Telefonieren zu beginnen. Im Gegensatz dazu ist die Ausschöpfung bei Bevölkerungsbefragungen und insbesondere auch bei Jugendbefragungen in den Nachmittags- und Abendstunden am höchsten. Die Anzahl der Kontaktversuche ist abhängig von der zu befragenden Zielgruppe und somit auch von der Art der Befragung.

Sind bei den Bevölkerungsbefragungen (Jugend und Erwachsene) nach dem vierten Kontaktversuch 70 bis 80 Prozent aller Interviews geführt worden und nach dem sechsten bis achten Anwählversuch zumeist 90 Prozent der Interviews komplett, so ergibt sich bei Unternehmensbefragungen doch ein ganz anderes Bild. Hier sind deutlich mehr Anwählversuche bis zum Zustandekommen eines Interviews notwendig.

Bei Befragungen von kleineren Betrieben ist es noch ähnlich wie bei den Bevölkerungsbefragungen: sechs Kontaktversuche waren bei einer Befragung von kleinen Handwerksbetrieben notwendig, um 80 Prozent und 9 Versuche, um 90 Prozent der Interviews zu realisieren. Etwas schwieriger gestaltete es sich bei einer Bildungsträgerbefragung in den ostdeutschen Bundesländern und Berlin: 9 Kontaktierungen wurden benötigt, um 80 Prozent der Interviews zu erhalten und bis zu 12 Versuche für 90 Prozent. Im Rahmen einer bundesweiten Unterneh-

mensbefragung von mittleren und Kleinbetrieben (KMU) waren erst nach 12 Kontaktversuchen 80 Prozent der Interviews geführt und nach weiteren drei Versuchen 90 Prozent (vgl. Abbildung 4). Der Grund dafür ist in der komplizierten Erreichbarkeit der Gesprächspartner zu sehen.

Die Ansprechpartner bei der Bildungsträger-Befragung und der Befragung in den kleinen und mittleren Unternehmen, die ein Interview verweigert hatten, taten dies in etwa der Hälfte der Fälle bis zum vierten Kontaktversuch. Bei der Befragung in den Handwerksbetrieben hatten bereits 70 Prozent der Verweigerer bis zum vierten Kontaktversuch ihre Nicht-Teilnahme angeführt.

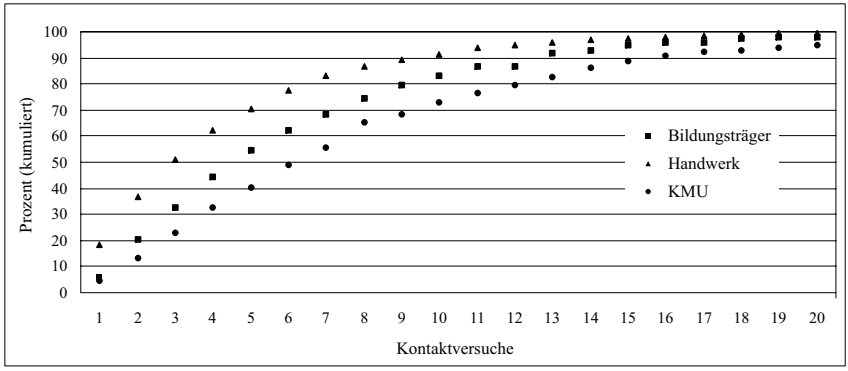


Abb. 4: Interviews nach Kontaktversuchen bei Unternehmensbefragungen

Sind bei den Bevölkerungsbefragungen nach dem vierten Kontaktversuch mindestens drei Viertel aller Interviews geführt worden und nach dem sechsten bis achten Anwählversuch zumeist 95 Prozent der Interviews komplett, so ergibt sich bei Elitenbefragungen ein anderes Bild. Es sind deutlich mehr Anwählversuche bis zum Zustandekommen eines Interviews notwendig. Nach dem vierten Kontaktversuch waren bei den Elitenbefragungen ein bis zwei Drittel der Interviews komplett und nach dem achten Versuch 70 - 80 Prozent (vgl. Tabelle 2). Hier zeigt sich deutlich, dass eine wesentlich größere Anzahl von Kontaktversuchen nötig und sinnvoll war, um die schwer erreichbaren Personen für ein Interview zu gewinnen.

Wird bei der Durchführung von Telefonumfragen die Effizienz beurteilt, so kann im Rahmen der am zsh durchgeführten CATI-Bevölkerungsstudien bestätigt werden, was in der Literatur bereits zu finden ist. Spätestens nach dem vierten Kontaktversuch ist der größte Teil der Interviews (drei Viertel bis 95 Prozent) durchgeführt und weitere Kontaktversuche könnten zu einer Ineffizienz führen (vgl. Tabelle 2).

Zur Reduzierung objektiver Ausfälle sollten allerdings mehrere Kontaktversuche unternommen werden. Kreiselmaier/Porst (1989, S. 34) berichten, dass in Studien, in denen die Zahl der Kontaktversuche erhöht worden ist, die Nicht-Erreichbarenquote gesenkt werden konnte. Aus ökonomischen Gründen sollte dies aber nicht beliebig ausgedehnt werden (vgl. Brückner et al. 1982, S. 27). Brückner et al. (1982) verweisen darauf, dass nach dem vierten Kontaktversuch eine deutliche Ineffizienz feststellbar ist, andere Angaben in der Literatur deuten darauf hin, dass die Ausschöpfung bereits nach dem dritten Kontaktversuch sinkt.

Tabelle 2: Interviewanzahl nach Kontakthäufigkeit

Studie	Thema	Komplette Interviews	
		nach 1 Kontaktversuch in %	nach 4 Kontaktversuchen in %
Bevölkerungsbefragungen			
SFB 580 TP A4: Bevölkerungsbefragung	Befragung zur Politikwahrnehmung	41	91
Befragung von Betroffenen der Flutkatastrophe	Situation Flutbetroffener	21	73
Befragung von Migranten	Ost-West-Migration junger Menschen	26	74
Befragung von Existenzgründern	Erfolg und Misserfolg von Ich-AGs	39	95
Befragungen politischer Eliten			
SFB 580 TP A4: Elitenbefragung	Lebenslauf jetziger und ehemaliger Politiker	22	66
SFB 580 TP A6: Befragung von KWGs	KWG in Ost- und Westdeutschland	5	39
SFB 580 TP A6: Befragung von Fraktionsvorsitzenden	Kommunalpolitik in Deutschland	4	32

Jedoch wird die Abhängigkeit der Ausschöpfung von der Anzahl der Kontaktversuche bei einer Studie auch von den jeweiligen Kontexten beeinflusst. Eine größere Anzahl von Kontaktversuchen wird nur dann empfohlen, wenn man schwer erreichbare Personen – wie zum Beispiel bei Elitenbefragungen – unbedingt erreichen will. Die Ergebnisse der CATI-Erhebungen des zsh bestätigen damit grundsätzlich die bereits in der Literatur beschriebenen Tendenzen, wobei die besonderen Anforderungen im Rahmen jeder Studie nicht außer Acht gelassen werden dürfen. Je nach Zielgruppe und Schwierigkeitsgrad der Erreichbarkeit muss entschieden werden, wie viele Kontaktversuche erwünscht und zugelassen sind –

im Hinblick auf die Effizienz der Ausschöpfung und natürlich auch aus ökonomischen Gründen.

5 Einschätzungen der Befragten zum Interview als Instrument der Qualitätsbewertung

Immer wieder stellt sich die Frage, warum Personen zur Teilnahme an einer Befragung motiviert sind. Eine einheitliche Theorie des Interviews gibt es auch nach jahrelanger Methodenforschung noch nicht. Zur Erklärung des Teilnahmeverhaltens an Befragungen werden zunehmend Rational-Choice-Theorien herangezogen (vgl. u. a. Schnell et al. 1995, S. 328-332)[16]. Demnach wird das Verhalten des Befragten im Interview als Ergebnis einer nach Kosten-Nutzen-Erwägungen erfolgten Entscheidung zwischen Handlungsalternativen beschrieben (vgl. z. B. Dillmann 1978, S. 12-16).

„Personen wählen die ihnen vorstellbare Handlungsalternative, die am ehesten angesichts der vorfindbaren Situationsumstände bestimmte Ziele zu realisieren verspricht" (Esser 1986, S. 321). Es handelt sich dabei immer um subjektiv vorstellbare Handlungsalternativen, um subjektiv jeweils definierte Handlungssituationen und Handlungsziele. Diese handlungstheoretische Erklärung basiert auf der Annahme, dass Personen in jeder Situation so reagieren, wie es ihnen am besten geeignet erscheint und die Situation wählen, die ihnen die stärkste Bedürfnisbefriedigung bringt (vgl. Schnell 1997). Die Entscheidung über die Teilnahme an einer Befragung kann als rationale Handlungswahl interpretiert werden. Die Teilnahmebereitschaft ist – neben personalen Eigenschaften – abhängig von der Einschätzung der Wichtigkeit und der Höhe des persönlichen Nutzens für die befragte Person, z. B. die Aussicht auf Verbesserung einer Situation, die den Befragten betrifft (vgl. Kaase 1999, S. 38).

In den folgenden Abschnitten werden diesbezüglich Auswertungen von Einschätzungen zum Interview im Rahmen von Bevölkerungsbefragungen des zsh vorgenommen[17]. Im Anschluss an die jeweilige Befragung wurden einige Fragen zum Interview selbst an den Befragten und auch an den Interviewer gerichtet. Dabei spielt die Frage nach der Belastung des Befragten (d. h. die Kosten im Sinne der Rational-Choice-Theorie) durch das Interview eine besondere Rolle. Schnell (1997, S. 166) beschreibt drei wesentliche Ursachen für Belastungen:

- die mögliche Verletzung der Privatsphäre,
- Befürchtungen über die Verwendung des Datenmaterials sowie

16 Andere Theorien des Interviews zur Erklärung des Teilnahmeverhaltens basieren beispielsweise auf sozialpsychologischen Ansätzen.

17 In diese Auswertungen sind knapp 12.000 Interviews eingegangen, die im Rahmen von 7 Studien in den Jahren 2003 und 2004 im zsh durchgeführt wurden.

- Belastungen durch den Befragungsvorgang selbst.

Zur Messung der Belastung wurden an den Interviewten Fragen gestellt, z. B. zur geschätzten Dauer der Befragung, zur empfundenen Anstrengung während des Interviews, zur Verständlichkeit der Fragen und zu Erinnerungsschwierigkeiten, beispielsweise vergangene Lebensabschnitte betreffend.

Der Interviewer hatte im Anschluss an das Gespräch ebenfalls Fragen zur geschätzten Länge des Interviews, zu Schwierigkeiten des Befragten bei der Beantwortung und zu Erinnerungsschwierigkeiten zu beantworten. Außerdem sollte der Interviewer angeben, wie anstrengend das Interview für ihn war und hatte die Möglichkeit, offene Kommentare zu geben. Um den Nutzen der Befragung zu messen, wurden den Personen im Anschluss an das Interview zudem Fragen zur Einschätzung der Wichtigkeit der Befragung, zur Interessantheit der Fragen und des Themas gestellt und die Möglichkeit offeriert, Kommentare zum Interview zu geben.

5.1 Geschätzte Interviewdauer

Über die optimale und maximale Länge telefonischer Interviews gehen die Meinungen sehr weit auseinander. In der Literatur finden sich Angaben von 20 bis 30 Minuten, aber auch bis zu 60 Minuten und mehr (vgl. Fuchs 1994, S. 56). Nach den Erfahrungen fast aller im zsh durchgeführten Studien, die eine Interviewdauer von 10 bis zu 90 Minuten aufweisen, hat die Dauer eines Interviews keinen direkten Einfluss auf das Teilnahmeverhalten. Die Interviewdauer wird nach dem Telefongespräch sowohl von den Befragten als auch von den Interviewern in fast allen Studien, die im CATI-Labor des zsh durchgeführt wurden, unterschätzt. Werden die Mittelwerte der Dauer der Befragungen differenziert nach Kohorten betrachtet, so ergibt sich folgendes Bild: In einer Jugend-Maßnahme-Befragung zur Verbesserung der beruflichen Bildungs- und Eingliederungschancen von Jugendlichen wurde die Dauer des Gesprächs durch die Befragten unterschätzt und bei einer ähnlich gelagerten Jugend-Maßnahme-Befragung zu einem Freiwilligen Sozialen Trainingsjahr überschätzt. Ein möglicher Grund könnte darin gesehen werden, dass die befragten Jugendlichen der zweiten Maßnahme nicht so viel interessante Inhalte zu berichten hatten und deshalb die Fragen der Interviewer bzw. das Gespräch als langwieriger empfanden. Die Interviewer haben bei beiden Befragungen die Gesprächsdauer kürzer eingeschätzt als sie tatsächlich war (vgl. Abbildung 5). Bei einer Jugend-Panel-Befragung unterschätzten die befragten Jugendlichen die Interviewdauer um etwa ein Drittel, bei den Interviewern war der Wert zwar geringer, jedoch wurde die Dauer auch kürzer als es tatsächlich der Fall war eingeschätzt. Gründe dafür sind z. B. in der Interessantheit und Wichtigkeit des Befragungsthemas (Mobilität ostdeutscher Jugendlicher), in der ange-

nehmen Interviewsituation (die Jugendlichen wurden von fast Gleichaltrigen befragt) und in der Beliebtheit des Telefonierens bei Jugendlichen zu sehen.

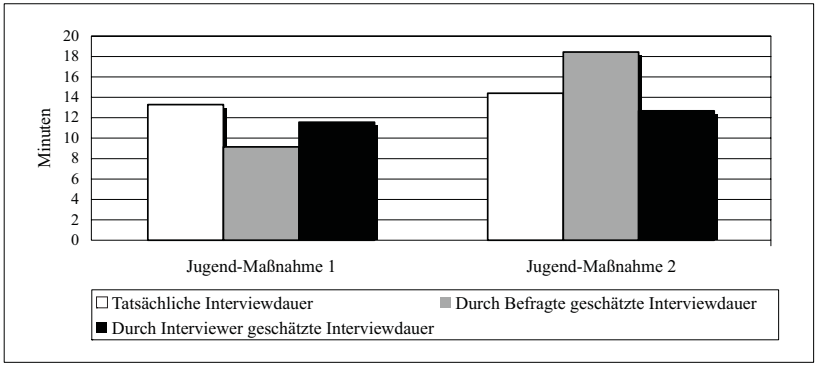


Abb. 5: Tatsächliche und geschätzte Interviewdauer bei den Jugend-Maßnahme-Befragungen

Ein ähnliches Bild zeigt sich auch für die Untersuchungen der erwachsenen Bevölkerung. Bei einer Befragung von Migranten, die aus Sachsen-Anhalt weggezogen sind, haben die Befragten die Interviewlänge ein wenig überschätzt. Jedoch war die Tendenz, bezüglich der Bereitschaft Fragen zu beantworten, sehr hoch und die Befragten gaben häufig über die gestellten Fragen hinaus Informationen zu der erfragten Situation. Möglicherweise ließ dieser Sachverhalt die geschätzte Interviewdauer bei den Teilnehmern etwas höher erscheinen. Bei einer Haushaltsbefragung in ländlichen Regionen haben die Befragten der Haushalte die Interviewdauer wesentlich unterschätzt. Die Interviewer gaben durchweg für alle Studien weniger Zeit für ein geführtes Interview an, als real zu verzeichnen war (vgl. Abbildung 6).

Bei der Befragung von Betroffenen der Flutkatastrophe im Jahr 2002 in Ostdeutschland und der Befragung zur Existenzgründung wurden den Befragten keine Fragen zum Interview gestellt.

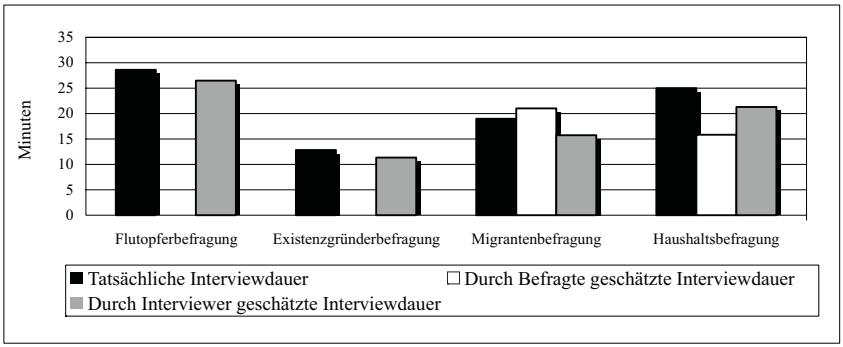


Abb. 6: Tatsächliche und geschätzte Interviewdauer bei Bevölkerungsbefragungen

5.2 Einschätzungen der Befragten zum Interview

5.2.1 Empfundene Belastung des Interviews

Mit der Frage „Wie anstrengend war das Interview?“ sollten die befragten Personen auf einer Skala von 1 bis 5 einschätzen, wie groß die empfundene Belastung für sie war. Auf der Skala von 1 bis 5 bedeutete 1 = überhaupt nicht anstrengend, 2 = kaum bzw. wenig anstrengend, 3 = teilweise anstrengend, 4 = überwiegend anstrengend und 5 = äußerst anstrengend. Bei den Jugendbefragungen zu Maßnahmen beurteilten jeweils knapp zwei Drittel, dass das Interview überhaupt nicht anstrengend war und ein weiteres Viertel gab jeweils an, dass es kaum bzw. wenig anstrengend war. Die Prozentzahl der befragten Jugendlichen, die das Interview äußerst anstrengend empfanden, lag bei beiden Erhebungen unter einem Prozent (vgl. Abbildung 7). Bei der Befragung der Migranten waren es sogar knapp drei Viertel, die das Interview überhaupt nicht anstrengend fanden und etwa weitere 20 Prozent beurteilten es als kaum bzw. wenig anstrengend (vgl. Abbildung 7). Dass die prozentuale Einschätzung eines überhaupt nicht anstrengenden Interviews bei den Jugendbefragungen leicht unter den Zahlen der Befragung von Migranten liegt, hängt damit zusammen, dass die Jugendlichen ihre Erwerbsepisoden seit dem Verlassen der Schule nacheinander zeitgemäß angeben sollten. Dabei wurde von den Jugendlichen eine relativ große Erinnerungsleistung abverlangt, was die Anstrengung erhöhte.

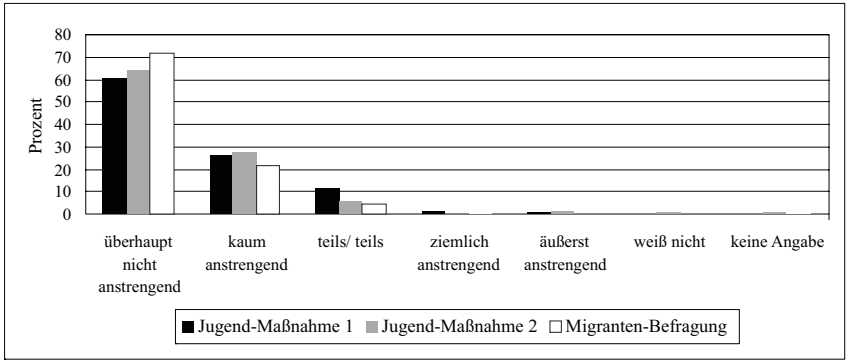


Abb. 7: „Wie anstrengend war das Interview?“

Etwas anders gestaltete sich der Grad der Anstrengung bei der Haushaltsbefragung in den drei ländlichen Regionen. Bei dieser Befragung sollten die Personen auf einer Skala von 0 bis 10 einschätzen, wie anstrengend das Interview war, wobei 0 „überhaupt nicht anstrengend“ und 10 „äußerst anstrengend“ bedeutete. In diesem Fall bewerteten 56 Prozent der Befragten den Grad der Anstrengung mit 0 (= überhaupt nicht anstrengend) und weitere 13 Prozent mit 1. Der höchste angegebene Wert auf der Skala von 0 bis 10 war der Wert 5. Die Erklärung, dass im Vergleich zu den eben angeführten Studien weniger Befragte das Interview mit „überhaupt nicht anstrengend“ bewerteten, ist auch in der inhaltlichen Gestaltung der Erhebung zu sehen. Die im jeweiligen Haushalt befragte Person sollte im Rahmen dieser Studie Auskunft über alle weiteren Haushaltsmitglieder und deren soziale Lage geben.

5.2.2 Verständlichkeit der Fragen

Die große Mehrheit der Befragten antwortete auf die Frage „Empfanden Sie es als einfach oder als schwierig, die Fragen zu verstehen?“ mit „einfach“ und „eher einfach“. Zur Auswahl wurde eine Skala mit folgenden Kategorien angeboten: „einfach“, „eher einfach“, „eher schwierig“, „sehr schwierig“. Bei den Jugendbefragungen zu Maßnahmen waren es zwei Drittel der Befragten, welche die Fragen einfach verstanden haben und etwa ein weiteres Viertel gab an, dass die Fragen „eher einfach“ zu verstehen waren. Bei der Befragung der Migranten und der Haushaltsbefragung in ländlichen Regionen waren es sogar jeweils etwa 80 Prozent der Befragten, die die Verständlichkeit der Fragen in die Kategorie „einfach“ einordneten. Im Rahmen dieser beiden Befragungen sind es immerhin fast 100 Prozent der Personen, welche die Verständlichkeit der Fragen „einfach“ oder „eher einfach“ beurteilten. Von den vier zur Auswahl stehenden Antworten bei

dieser Fragestellung gab es in keiner Befragung eine einzige Person, die mit „sehr schwierig“ antwortete und ganz wenige Befragte, die „schwierig“ angaben (vgl. Abbildung 8).

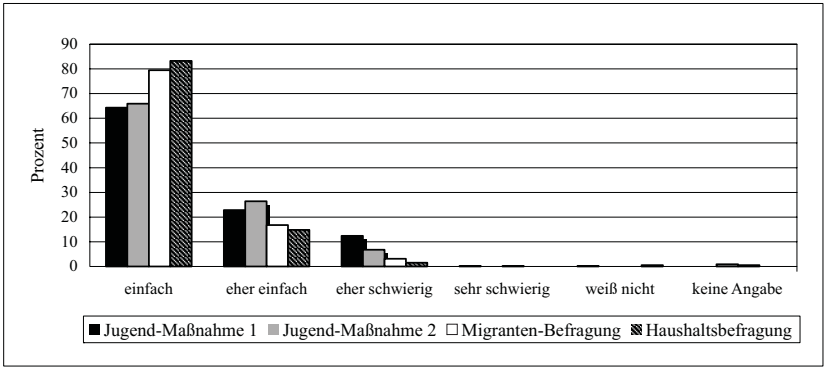


Abb. 8: Verständlichkeit der Fragen

5.2.3 Erinnerungsschwierigkeiten

Bei den Jugendbefragungen zu Maßnahmen und bei der Befragung von Migranten wurden die Teilnehmer nach bestimmten Begebenheiten aus ihrem bisherigen Leben befragt. Auf die Frage „Hatten Sie Schwierigkeiten, sich an einzelne Episoden bzw. an einzelne Details zu erinnern?“ gab es für die Interviewteilnehmer folgende Antwortoptionen: „hatte (überhaupt) keine Schwierigkeiten“, „hatte manchmal Schwierigkeiten“ und „hatte häufig Schwierigkeiten“.

Die Befragten der beiden Jugendstudien und die der Erhebung zur Migration gaben mit jeweils über 95 Prozent an, überhaupt keine oder nur manchmal Erinnerungsschwierigkeiten zu haben. Wenn auch die Anstrengung – wie bereits aufgezeigt – differenziert nach den Studien unterschiedlich ausfällt, so gaben doch die meisten Personen, die im Rahmen dieser Studien befragt wurden, keine Erinnerungsschwierigkeiten an.

5.2.4 Wichtigkeit der Befragung

Sind die Gesprächspartner zu einem Interview bereit, so kann davon ausgegangen werden, dass die Studie als wichtig eingeschätzt wird.

Auf die Frage „Für wie wichtig erachten Sie es, dass solche Befragungen zu diesem Thema durchgeführt werden?“ antwortete etwa ein Drittel der Befragten der im zsh durchgeführten Bevölkerungsbefragungen, dass sie es für sehr wichtig

erachten, Umfragen zu dem jeweiligen Thema durchzuführen. Und weiterhin knapp die Hälfte der Befragten befand die Durchführung von Befragungen eher wichtig. Insgesamt beurteilten somit über 80 Prozent der befragten Personen auf einer Skala von 1 bis 5 die Erhebung als sehr und eher wichtig. Lediglich zwischen 1 und 3 Prozent der Interviewten gaben an, dass derartige Befragungen äußerst unwichtig sind.

Ein Unterschied zwischen jugendlichen Teilnehmern und Erwachsenen lässt sich bei dieser Fragestellung nicht feststellen, so dass davon ausgegangen werden kann, dass ein Großteil derjenigen, die an CATI-Befragungen des zsh teilgenommen haben, die Wichtigkeit der Durchführung von Telefonbefragungen im Rahmen wissenschaftlicher Studien bestätigen.

Die befragten Personen gehen – wird zur Erklärung der Teilnahme an der Befragung die Rational-Choice-Theorie herangezogen – davon aus, dass sie mit ihrer Teilnahme einen wichtigen Beitrag für das Forschungsprojekt leisten und eine Veränderung in eine positive Richtung beeinflussen können. Diese Information geben die Interviewer dem Befragten beim Erstkontakt. Sie weisen die Befragten darauf hin, dass es sich um eine wissenschaftliche Studie handelt und die Befragungsergebnisse zu Forschungszwecken genutzt werden und für mögliche Veränderungen relevant sind. Um diesen Nutzen für die Befragung aufzeigen und auf evtl. Fragen der Interviewpartner reagieren zu können, erhalten die Interviewer in Schulungen vor Beginn einer Studie alle relevanten Informationen (vgl. dazu Kap. 3 dieses Beitrages). Abbildung 9 zeigt beispielhaft die Einschätzung der Wichtigkeit der Befragung der beiden Studien bezüglich der Maßnahmen für Jugendliche und der Befragung von Migranten.

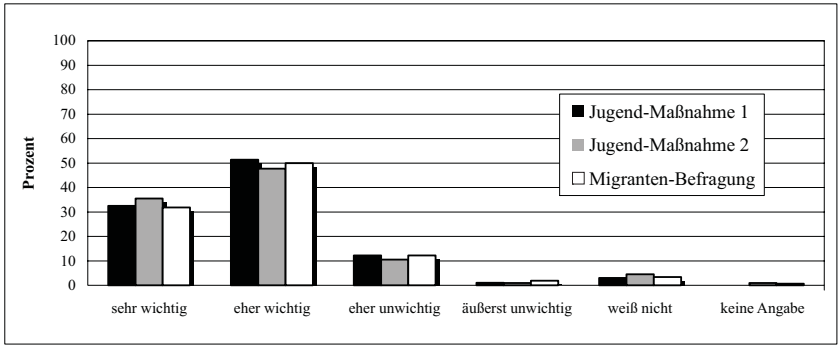


Abb. 9: Wichtigkeit der Befragung

5.2.5 Offene Kommentare zum Interview

Zum Ende des Interviews hatten die befragten Personen die Möglichkeit, ganz allgemein Kommentare oder Anmerkungen zum Interview oder zur Befragung zu geben, was in den verschiedenen Studien des zsh zwischen 10 und 20 Prozent der Gesprächspartner nutzten.

Dabei handelte es sich bei einem Großteil der Kommentare um einen Nachtrag inhaltlicher Art, um Aussagen zum Thema der Befragung bzw. zur Befragung an sich, um Aussagen über die Interviewsituation und den Fragebogen oder um Kritik und Anregungen im Rahmen der Befragung.

Des Weiteren wurden allgemeine Erfahrungen weitergegeben und in einigen wenigen Anmerkungen Verbesserungsvorschläge für den Fragebogen. Häufig wurde auf die angenehme Interviewsituation und auf ein interessantes Gespräch verwiesen.

5.2.6 Teilnahme an Befragungen zu einem früheren Zeitpunkt

Als Belastung – und somit im Sinne der Rational-Choice-Theorie den Kosten zuzurechnen – wird von Personen die häufige Aufforderung zur Teilnahme an einer Befragung, besonders auch am Telefon, erlebt. Aus diesem Grund wurden die Interviewten der Studien des zsh gefragt, ob sie bereits früher einmal an telefonischen, schriftlichen oder mündlichen Befragungen teilgenommen haben. Etwa die Hälfte der Interviewten, sowohl der Jugendbefragungen zu Maßnahmen als auch der Befragungen der erwachsenen Bevölkerung gab an, noch nie an Befragungen teilgenommen zu haben und lediglich bis zu jeweils 2 Prozent konnten sich nicht mehr erinnern oder wussten es nicht mehr genau.

Von den 50 Prozent der Personen, die bereits an Umfragen teilgenommen haben, gab der Großteil der Jugendlichen an, im Rahmen wissenschaftlicher Untersuchungen mitgewirkt zu haben, während der größere Teil der erwachsenen Bevölkerung eher für die Marktforschung an Umfragen teilnahm (vgl. Abbildung 10).

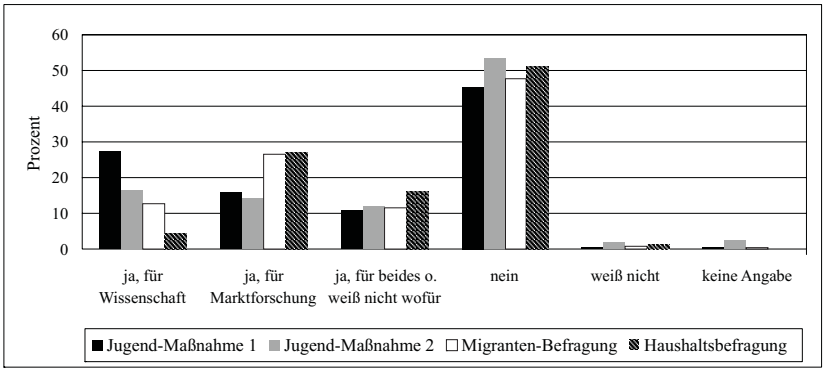


Abb. 10: Teilnahme an früheren Befragungen

5.3 Qualitätsbewertung durch Supervision

Ein wesentlicher Vorteil computergestützter Telefonbefragungen liegt in der Möglichkeit der ständigen Kontrolle durch den Supervisor, da diese heute fast ausschließlich in dafür vorgesehenen Laboren durchgeführt werden. Diese Zentralisation und die daraus resultierende Überwachung und Kontrolle der Interviewer hat wiederum zur Folge, dass die Interviewer sich stärker an die Vorgaben der Forscher halten. Während der gesamten Feldzeit ist ein Supervisor für Rückfragen anwesend, was zur korrekten Arbeit und schnellen Lösung von auftretenden Problemen beiträgt. Damit kann eine hohe Interviewerqualität und somit auch Datenqualität erzielt werden.

Die Kontrolle kann nach drei Kriterien erfolgen:

- nach der Effektivität anhand der Interviews pro Stunde,
- hinsichtlich der realisierten Ausschöpfung der ihnen zugewiesenen Interviews und
- durch eine Note von zwei Supervisoren für die Interviewqualität.

Erfüllt ein Interviewer die vorausgesetzten Anforderungen nicht bzw. liegt er unter dem Durchschnitt des Interviewerstabes, wird eine Nachschulung angeordnet. Die Abbildung 11 stellt diesen Prozess graphisch dar.

Auch im CATI-Labor des zsh wird die Arbeit der Interviewer stets kontrolliert. Dies erfolgt in erster Linie durch eine quantitative Beurteilung. Die CATI-Software macht es mittels einer Interviewer-Identifikationsnummer (ID) möglich, die Leistungen jedes einzelnen Interviewers zu kontrollieren. Folgende Informationen können genau ermittelt werden:

- die Anzahl der Interviews pro Stunde,
- die Interviewlänge,
- die Anzahl der Verweigerungen,
- die Anzahl der Wählversuche.

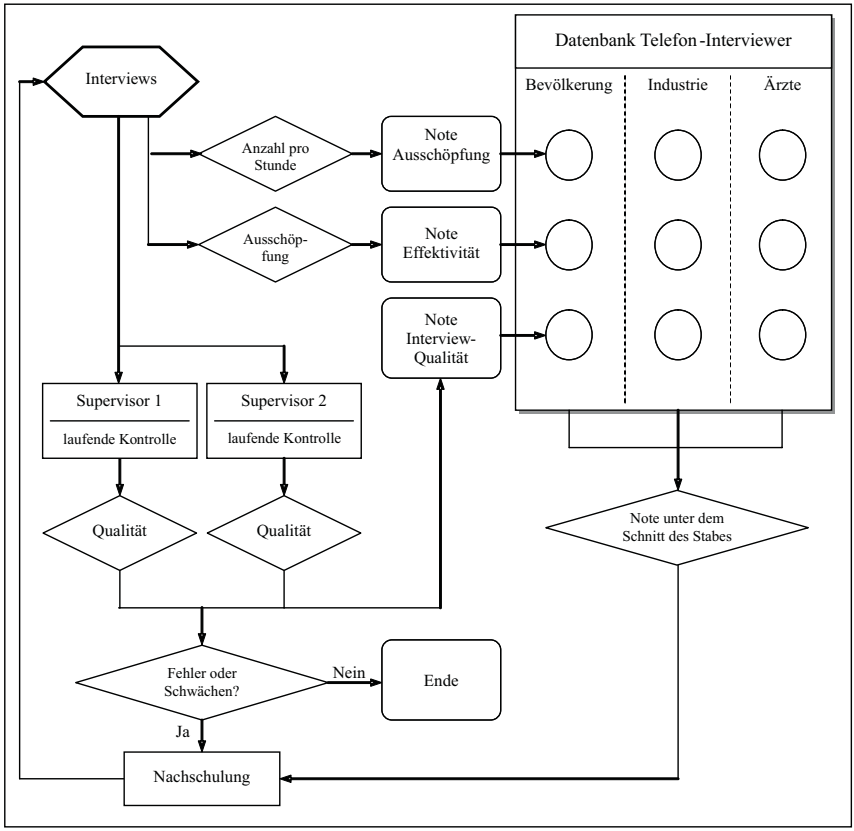


Abb. 11: Die Qualitätskontrolle der Telefoninterviewer
Quelle: Anders (1990), S. 434

Vergleicht man diese Informationen mit den Durchschnittswerten, kann zum einen geprüft werden, wie fleißig die einzelnen Interviewer arbeiten und zum anderen, wie erfolgreich und effektiv sie sind.

Des Weiteren gibt es noch die Möglichkeit einer qualitativen Bewertung durch das Mitschneiden bzw. Mithören von Interviews. Dies erfolgt jedoch aus datenschutzrechtlichen Gründen nur mit dem Einverständnis des Befragten und somit auch des Interviewers.

Diese Möglichkeit der Kontrolle muss mit dem Projektteam bzw. dem Auftraggeber abgesprochen werden.

Aufwendiger ist es, beispielsweise über Häufigkeiten von Antworten wie „weiß nicht" und „keine Angabe" Aussagen über die Qualität des Interviews und damit des Interviewers zu machen. Dazu wird derzeit im Rahmen des Sonderforschungsbereichs 580 ein Projekt bearbeitet, welches versucht, die Qualitätsfrage auf der Basis solcher Informationen zu klären.

6 Resümee

Das Telefoninterview hat sich in den letzten Jahren zu einer eigenständigen und gleichwertigen Erhebungsform neben persönlichen und schriftlichen Befragungen entwickelt und stellt als relativ neue Erhebungsart auch ganz neue Anforderungen an die Forscher und vor allem an die Interviewer. Anlass gibt unter anderem die neue Erhebungssituation, die sich durch die Beschränkung auf den akustischen Kommunikationskanal charakterisiert. Vor diesem Hintergrund muss auch die Frage nach der Qualität der Interviews und der Qualitätskontrolle neu gestellt werden.

Die Kommunikation zwischen den am Interview beteiligten Akteuren verläuft ausschließlich über die Stimme. Das Wegfallen aller äußerlichen Stimuli bringt den Vorteil mit sich, dass äußere Interviewermerkmale, die in persönlichen Befragungen die Antworten der Befragten beeinflussen können, ausgeblendet werden. Infolgedessen kommt der Stimme eine erhebliche Bedeutung zu. Sie gehört untrennbar zum Erscheinungsbild und zum Auftreten einer Person, denn die Stimme und die Sprache stehen im unmittelbaren Zusammenhang mit der Persönlichkeit. Stimmeigenschaften und rhetorische Fähigkeiten können das Befragtenverhalten im starken Maße beeinflussen. Ergebnisse zeigen, dass eine entspannte Stimmlage, eine gute Modulation, eine relativ schnelle Sprechgeschwindigkeit sowie eine angemessene Lautstärke einen positiven, kompetenten und vertrauenswürdigen Eindruck beim Befragten hinterlassen. Die Stimme muss klar und verständlich sein. Diese Anforderungen sind Voraussetzung für die Zustimmung

des Befragten zur Teilnahme an der Befragung und zur Vermeidung von Verweigerungen sowie Abbrüchen im Verlauf eines Interviews.

Während der computergestützten telefonischen Befragung wird dem Interviewer eine Reihe von Aufgaben abgenommen. Dadurch können vom Interviewer ausgehende Fehler bei der Administration des Fragebogens minimiert werden. Um jedoch die gesamte Aufmerksamkeit während des Interviews dem Befragten widmen zu können, muss der Interviewer technisch (das CATI-System betreffend) und inhaltlich (die spezielle Studie betreffend) geschult sein. Eine Schulung aus sprechwissenschaftlicher Sicht, so wie sie im CATI-Labor des zsh durchgeführt wird, ist ebenfalls empfehlenswert, da die Stimme des Interviewers, d.h. seine Sprechtechnik, aber auch seine rhetorischen Fähigkeiten, für die erfolgreiche Durchführung eines Interviews entscheidend sind.

Durch den zentralisierten Interviewereinsatz ist es möglich, die Arbeit der Interviewer zu kontrollieren und die Effekte besser zu analysieren. Untersuchungen und Analysen solcher Interviewereffekte sind jedoch in Zukunft weiter notwendig, um das Ausmaß dieser Einflüsse einschätzen zu können. Die gewonnenen Erkenntnisse könnten bei der Rekrutierung und Schulung der Interviewer angewendet bzw. berücksichtigt werden.

Um neue Erkenntnisse der Qualitätsbewertung durch den Befragten im Rahmen von telefonischen Interviews zu gewinnen, wurden bei Erhebungen im CATI-Labor des zsh im Anschluss an das Befragungsgespräch Fragen zum Interview selbst gestellt.

Anhand der Ergebnisse zu den Einschätzungen der befragten Personen zum Interview bei den Erhebungen im CATI-Labor des zsh lässt sich ein fast durchweg positives Bild aufzeigen. Das hängt zum einen damit zusammen, dass bei der ersten Kontaktaufnahme die Seriosität dieser Befragung vermittelt werden kann. Zum anderen liegt es daran, dass in den besonders positiven Beispielen jeweils Befragungsgruppen interviewt wurden, die sich im jeweiligen Thema speziell angesprochen fühlten. Die Entscheidung zur Bereitschaft am Interview führt dann auch dazu, dass der überwiegende Teil der Befragten die Forschungsprojekte für wichtig befindet. Schwierigkeiten beim Verstehen der Fragen und beim Antworten sind nur sehr selten zu verzeichnen.

Durch die Möglichkeit, offen Kommentare zum Interview oder zur Studie zu geben, können die Befragten Informationen für die Projektmitarbeiter anmerken, die sonst verloren gehen würden.

Die Angaben der Befragten zum Interview stimmen weitgehend überein mit den Einschätzungen und dem positiven Echo der Interviewer. Die Interviewer bewerten die Telefongespräche zum größten Teil als interessant. Die Anstrengung während der Befragung wird von Seiten der Interviewer als relativ gering eingeschätzt, da durch langjährige Erfahrungen die meisten Interviewer ein professionelles Arbeiten aufweisen. Auch wird von vielen Interviewern die Möglichkeit

genutzt, in einer offenen Frage Kommentare zum Interview abzugeben, um Informationen weiterzuleiten, die außerhalb des standardisierten Interviews wichtig sein könnten[18].

Bezüglich der Ausschöpfungsquote als wichtigem Qualitätsmaßstab kann gesagt werden, dass mit CATI-Befragungen gute Ausschöpfungsquoten zu verzeichnen sind. Im CATI-Labor des zsh konnten Ausschöpfungsquoten bis zu 88 Prozent erreicht werden. Die Ausschöpfung hätte in einigen Studien noch gesteigert werden können, wenn nicht durch das Erreichen der angestrebten Fallzahl die Befragung beendet werden musste. Im Hinblick auf die Effizienz der Ausschöpfung lässt sich konstatieren, dass je nach Zielperson und Schwierigkeitsgrad der Erreichbarkeit entschieden werden muss, wie viele Kontaktversuche maximal erwünscht und zugelassen sind, bis ein Interview stattfindet. Im Rahmen von Bevölkerungsbefragungen sind durchschnittlich weniger Anwählversuche dafür notwendig als bei Unternehmens- und Elitenbefragungen. Die schwierige Erreichbarkeit beispielsweise der Zielpersonen mit Positionen im Bereich der Politik und das meist enge Zeitbudget der Eliten oder auch der Geschäftsführer von Unternehmen bedingt eine höhere Anzahl von Kontaktversuchen bis zur Durchführung eines Interviews. Die Ausschöpfungsquoten sind am Ende ähnlich hoch oder sogar höher als bei anderen Befragungsarten.

Ausschöpfungsquoten müssen differenziert betrachtet werden, da sie abhängig sind von verschiedenen Faktoren, die zum einen einer inhaltlichen und zum anderen einer methodisch formalen Dimension zugeordnet werden können und zur Messung der Qualität von Umfragen dienen. CATI ist in jedem Fall ein geeignetes Instrument, um schnell und zuverlässig große Befragungen durchzuführen. Dazu muss – wie bei allen anderen Befragungsmethoden auch – ständig an der Qualitätssicherung gearbeitet werden.

7 Literaturverzeichnis

Amon, Ingrid (2000): Die Macht der Stimme: Persönlichkeit durch Klang, Volumen und Dynamik. Wien: Ueberreuter.

Anders, Manfred (1990): Praxis der Telefonbefragung. In: Forschungsgruppe Telekommunikation (Hrsg.): Telefon und Gesellschaft. Band 2: Internationaler Vergleich, Sprache und Telefon, Telefonseelsorge und Beratungsdienste, Telefoninterview. Berlin: Spiess.

18 Ausführliche Auswertungen der Einschätzungen zum Interview von Seiten der Interviewer sind nachzulesen in Buchwald, C. (2005): Telefoninterview ist nicht gleich Telefoninterview. In: Wiekert, I. (Hrsg.): Zehn aus Achtzig. Burkart Lutz zum 80. Berlin: GSFP, S. 237 ff.

Blasius, Jörg; Reuband, Karl-Heinz (1995): Telefoninterviews in der empirischen Sozialforschung: Ausschöpfungsquoten und Antwortqualität. In: ZA-Information Nr. 37, S. 64-87.

Brückner, Erika; Hormuth, Stefan E.; Sagawe, Helmut (1982): Telefoninterviews – Ein alternatives Erhebungsverfahren? Ergebnisse einer Pilotstudie. ZUMA-Nachrichten 11, Mannheim.

Buchwald, C. (2005): Telefoninterview ist nicht gleich Telefoninterview. In: Wiekert, Ingo (Hrsg.): Zehn aus Achtzig. Burkart Lutz zum 80. Berlin: GSFP, S. 209-244.

Buchwald, Christina (2006): Das Telefoninterview - Instrument der Zukunft? Forschungsberichte aus dem Zentrum für Sozialforschung Halle e. V.

Cannell, Charles F.; Kahn, Robert L. (1968): Interviewing. In: Gardner, Lindzey; Aronson, Elliot (Hrsg.): Handbook of Social Psychology. Vol. 2. Reading, Mass.: Addison-Wesley Publishing Co, S. 526-595.

Coblenzer, Horst; Muhar, Franz (1989): Atem und Stimme. Anleitung zum guten Sprechen. Wien: Österreichischer Bundesverlag.

Dillmann, Don A. (1978): Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons.

Eckert, Hartwig; Laver, John (1994): Menschen und ihre Stimmen. Aspekte der vokalen Kommunikation. Weinheim: Beltz.

Esser, Hartmut (1986): Können Befragte lügen? Zum Konzept des „wahren Wertes“ im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Jg. 38, H. 2, S. 314-336.

Frey, James H.; Kunz, Gerhard; Lueschen, Günther (1990): Telefonumfragen in der Sozialforschung: Methoden, Techniken, Befragungspraxis. Opladen: Westdeutscher Verlag.

Friedrichs, Jürgen (1990a): Methoden empirischer Sozialforschung. Opladen: Westdeutscher Verlag, 14. Aufl.

Friedrichs, Jürgen (1990b): Gesprächsführung im telefonischen Interview. In: Forschungsgruppe Telekommunikation (Hrsg.): Telefon und Gesellschaft. Band 2: Internationaler Vergleich, Sprache und Telefon, Telefonseelsorge und Beratungsdienste, Telefoninterview. Berlin: Spiess.

Fuchs, Marek (1994): Umfrageforschung mit Telefon und Computer. Einführung in die computergestützte telefonische Befragung. Weinheim: Beltz, Psychologie-Verlag-Union.

Fuchs, Marek (1995): Die computergestützte telefonische Befragung. Antworten auf Probleme der Umfrageforschung? In: Zeitschrift für Soziologie ZfS. Jg. 24, H. 4, S. 84 - 299.

Fuchs, Marek (2000): Interviewsituation in computergestützten Befragungen - zur Wirkung von Technologieeffekten. In: Statistisches Bundesamt (Hrsg.):

Neue Erhebungsinstrumente und Methodeneffekte. Wiesbaden: Metzler-Poeschel.

Gabler, Siegfried; Häder, Sabine; Hoffmeyer-Zlotnik, Jürgen H. P. (Hrsg.) (1998): Telefonstichproben in Deutschland. Opladen/Wiesbaden: Westdeutscher Verlag.

Häder, Sabine (1994): Auswahlverfahren bei Telefonumfragen. ZUMA-Arbeitsbericht 94/03, Mannheim.

Hormuth, Stephan E.; Brückner, Erika (1985): Telefoninterviews in Sozialforschung und Sozialpsychologie. Ausgewählte Probleme der Stichprobengewinnung, Kontaktierung und Versuchsplanung. Kölner Zeitschrift für Soziologie und Sozialpsychologie Jg. 37, H. 1; S. 526-546.

Hyman, J. (1954): Interviewing in Social Research. Chicago.

Kaase, Max (Hrsg.) (1999): Qualitätskriterien der Umfrageforschung: Denkschrift = Quality criteria for survey research/ Deutsche Forschungsgemeinschaft. Berlin: Akademischer Verlag.

Kreiselmaier, Jutta; Porst, Rolf (1989): Methodische Probleme bei der Durchführung telefonischer Befragungen: Stichprobenziehung und Ermittlung von Zielpersonen, Ausschöpfung und Non-Response, Qualität der Daten. ZUMA-Arbeitsbericht 89/12, Mannheim.

Oksenberg, Lois; Cannell, Charles (1988): Effects of Interviewer Vocal Characteristics on Nonresponse. In: Groves, Robert M.; Biemer, Paul P.; Lyberg, Lars E.; Massey, James T.; Nicholls, William L.; Waksberg, Joseph (Hrsg.): Telephone Survey Methodology. New York: Wiley & Sons, S. 257-269.

Porst, Rolf (1991): Ausfälle und Verweigerungen bei einer telefonischen Befragung. ZUMA-Arbeitsbericht 91/10, Mannheim.

Porst, Rolf (1993): Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Annäherung aus der ZUMA-Perspektive. ZUMA-Arbeitsbericht 93/12, Mannheim.

Reinecke, Jost (1991): Interviewer- und Befragtenverhalten. Theoretische Ansätze und methodische Konzepte. Opladen: Westdeutscher Verlag.

Schnell, Rainer; Hill, Paul Bernhard; Esser, Elke (1995): Methoden der empirischen Sozialforschung. München u. a.: Oldenbourg.

Schnell, Rainer (1997): Nonresponse in Bevölkerungsumfragen. Opladen: Leske + Budrich.

Statistisches Bundesamt (Hrg.) (2005): Datenreport 2004. Bonn: Bundeszentrale für politische Bildung.

Von Kirschhofer-Bozenhardt, Andreas; Kaplitze, Gabriele (1991): Das Interviewernetz. In: Holm, Kurt (Hrsg.): Die Befragung 1. der Fragebogen – Die Stichprobe. Tübingen: Franke Verlag, 4. Aufl.

Wachtel, Stefan (1998): Sprechen und Moderieren in Hörfunk und Fernsehen. Konstanz: UVK Medien, 3. Aufl.

Aus- und Weiterbildung als Instrument der Qualitätssicherung

Hartmut Scheffler

Hohe Qualität und damit Qualitätssicherung durch Aus- und Weiterbildung sind Voraussetzungen für sinnvolle, zielgerichtete Markt- und Sozialforschung. Sie sind auch unerlässlich angesichts des Selbstverständnisses der Markt- und Sozialforschung als „angewandter wissenschaftlicher Forschung".

Vor diesem Hintergrund ist festzuhalten, dass es für die angewandte wissenschaftliche Markt- und Sozialforschung keine einheitlichen Abschlüsse, keine einheitliche Standardausbildung und auch keine vorgeschriebene Zulassung gibt. Negativ formuliert sind die Einstiege und Zugänge in die Markt- und Sozialforschung beliebig, positiv formuliert – und dies ist ein wichtiger Aspekt für die Vielfalt des Berufsfeldes – sind sie vielfältig.

In der Markt- und Sozialforschung sind Soziologen und Psychologen beschäftigt, weiterhin Betriebswirte, Politologen, Informatiker, Linguisten, Medienwissenschaftler, Mathematiker, Ökotrophologen, Ethnologen, Neurologen, IT-Experten, Werbefachleute, Category Manager, und und und.

Es handelt sich also fraglos um ein Berufsfeld mit hoher Interdisziplinarität und diese Interdisziplinarität ist ebenso fraglos als Stärke hervorzuheben. Dies gilt umso mehr, als es sehr wohl einen eigenständigen originären Kern an Wissen für die Markt- und Sozialforschung gibt. Dies ist ein Wissenskern, der in dieser Kombination für kein anderes Berufsfeld relevant ist, der maßgeblich für die Qualität für Markt- und Sozialforschung ist, der auf wissenschaftlichen Grundlagen beruht und ein wesentliches Element der Eigenständigkeit und Legitimation des Berufsfeldes ist. Dieser originäre Kern besteht im Wesentlichen aus drei Bereichen: Der Wissenschaftstheorie und Methodologie, den Methoden und der Methodenlehre der empirischen Sozialforschung sowie dem Bereich der Statistik (deskriptiv wie schließend).

Als Zwischenresümee festzuhalten bleibt also einerseits ein originärer Kern an Wissenstatbeständen, zweitens eine hohe Interdisziplinarität mit vielfältigen Zugängen aus verschiedenen Wissenschaftsrichtungen und drittens die nicht existierende einheitliche (Mindest-)Ausbildung.

Woher kommt dann das existierende Wissen der vielen tausend Mitarbeiter in der Markt- und Sozialforschung? Die Wissensvermittlung geschieht (natürlich mit unterschiedlichen Inhalten und Schwerpunktsetzungen) an vielen Fachhoch-

schulen und Universitäten, sie wird in Zukunft verstärkt im Rahmen der klassischen dualen Ausbildung (Fachangestellte(r)) geschehen und sie geschieht durch ständige Schulungen der Verbände (hier vor allem des BVM) sowie der Arbeitgeber/Institute und Institutionen. Erneut ist aber festzuhalten: Eine einheitliche Regelung der so vermittelten Inhalte, ein verbindliches Curriculum gibt es nicht.

Unabhängig hiervon hat sich auf Seiten der Anwender/Auftraggeber von Markt- und Sozialforschung eine gewisse Polarisierung im Hinblick auf Qualität und Qualitätsbewusstsein ergeben. Positiv sind die Fälle zu nennen, bei denen die entsprechenden Abteilungen in den Unternehmen aufgestockt und qualifiziert wurden und damit hohe Qualitätserwartungen und ein hohes Qualitätsbewusstsein an den Tag legen. Leider gibt es auf der anderen Seite der Qualitäts-Bandbreite einen Zuwachs an einflussnehmenden (heißt: beauftragenden und die Ergebnisse nutzenden) Personen und Abteilungen, bei denen die Kenntnis der Qualitätskriterien und damit der Bedeutung der Qualität für die Verwendbarkeit der Ergebnisse nicht oder kaum vorhanden und als Entscheidungskriterium auch nicht relevant ist. Fehlende Kenntnisse um Qualitätskriterien, fehlende Bedeutung der Qualitätskriterien für die Tagesarbeit und damit fehlende Nachfrage nach Qualitätskriterien: ein Teufelskreis!

Leider ist es so, dass dieses partiell fehlende Qualitätsbewusstsein auch Angebote und Anbieter sich hat entwickeln lassen, die ihr Geschäftsfeld darauf aufbauen. Ein besonders negatives Beispiel ist die sogenannte „Do-it-yourself-Forschung". Hiermit sind Angebotspakete gemeint, die Anwendern von Markt- und Sozialforschung ähnlich des guten alten Prinzips des Fischer-Baukastens/ Fischertechnik suggerieren, dass mit etwas gesundem Menschenverstand jeder (scheinbar) Forschungsprojekte durchführen kann, zumindest jeder einmal üben darf. Sieht man einmal von dem vordergründigen Geschäftszweck dieser Angebote ab, dann steht dahinter eine naive Unkenntnis der methodologischen und methodischen Anforderungen, eine dramatische Unterschätzung des Qualitätskriteriums „Erfahrung" und eine ebenso naive Unkenntnis bzgl. der Auswirkungen qualitativ schlecht erhobener Daten auf Entscheidungen (ein Risiko für die gesamte Profession). Es darf allerdings zugestanden werden, dass diese Do-it-yourself-Angebotspakete nicht wirklich erfolgreich sind.

Wenn man sich fragt, warum die Markt- und Sozialforschung neben dem zunehmenden Akzeptanzproblem in der Öffentlichkeit (siehe die aktuellen Diskussionen um Telefonumfragen) auch mit Unkenntnis im Hinblick auf die Wissenschaftlichkeit im Allgemeinen und Qualitätsanforderungen und Qualitätsbewusstsein im Speziellen zu kämpfen hat, dann gilt es hier für die gesamte Branche, sich an die eigene Nase zu fassen. Es ist nicht gelungen, im Rahmen der Öffentlichkeitsarbeit eine entsprechende Kenntnis der Branche in der Öffentlichkeit wie bei Multiplikatoren (Presse, Politik) durchzusetzen. Man findet häufig auch auf Seiten der Befragten/ der Bevölkerung (natürlich) Unkenntnis im Hin-

blick auf die Qualität. Man findet bei Anwendern häufig den Vorwurf, die Markt- und Sozialforschung sei zu theoretisch, zu wenig aktions- und handlungsgetrieben. Kurz: Es herrscht insgesamt eine relativ große Unkenntnis im Hinblick auf Anforderungen, Leistungen und gesellschaftliche Relevanz.

Zurück und erneut fokussiert auf das Ausbildungsthema: Die Kenntnis um die wissenschaftlichen Hintergründe einerseits und Notwendigkeit einer umfassenden Aus- und Weiterbildung für die Qualitätssicherung andererseits ist unterbelichtet.

Eine wichtige Zwischenbemerkung: Die Markt- und Sozialforschung hat ihre Privilegien im Hinblick auf die Möglichkeiten zur Durchführung von Interviews deshalb erhalten, weil sie wohlbegründet angewandte wissenschaftliche Forschung ist (siehe Forschungsfreiheit im Grundgesetz!), sich zum strikten Anonymitätsgebot und zur Trennung von forschungsfremden Tätigkeiten (Verkauf, Direktmarketing etc.) verpflichtet. Ohne Aus- und Weiterbildung ist aber natürlich der wissenschaftliche Charakter und sind damit wichtige Argumente für die existierenden Privilegien gefährdet.

Es ist hoffentlich deutlich geworden, welche Relevanz aus ganz unterschiedlichen Gründen (Verwendbarkeit der Ergebnisse, Ansehen in der Öffentlichkeit, Privilegien) die Qualität, Qualitätssicherung und damit Aus- und Weiterbildung haben. Umso wichtiger ist es, auf gelungene Ansätze und Entwicklungen hinzuweisen.

Da ist zum einen natürlich (trotz fehlender Einheitlichkeit) auf die Ausbildung im Hinblick auf Methodologie, Methodenlehre und Statistik an den Universitäten, in den verschiedenen Fakultäten und Ausbildungsgängen hinzuweisen. Um es zu wiederholen: Wünschenswert wäre hier ein sicherer und verpflichtender Mindeststandard.

Eine sehr gute Tradition und ein herausragendes Ausbildungsbeispiel sind die verschiedenen Seminare der Verbände und hier vor allem des BVM.

Ausbildungsbeispiel 1:
Verbände - BVM

DESTATIS wissen.nutzen.

Seminarprogramm 2007

Hauptthemen	Beispiele
Grundlagenseminare	Einführung, Vertiefung
Umfrageforschung	Grundlagen und Vertiefung, Fragebogenerstellung
Qualitativ-psychologische Forschung	Grundlagen, Moderationen von Gruppendiskussionen, Frage- und Explorationstechniken, apparative und innovative Tools, Kreativitätstechnik
Statistik	Grundlagen, multivariate Analysen, Conjoint Measurement, Customer Profiling, Faktoren- und Cluster-Analyse, Discrete Choice, Regression und Varianz
Marketingforschung	Preisforschung, Marktprognosen, statistische Prognosemodelle, Neuromarketing, Image- und Markenwert, Online-Marktforschung, Mediaforschung und –planung, Konkurrenzforschung, Werbewirkungsforschung
Marktforschung	Ausgewählte Rechtsprobleme, Datenschutz in der Markt- und Sozialforschung
Der Kunde im Fokus	Messung von Kundenzufriedenheit, Mystery Shopping, Messung von Servicequalität
Kommunikation und Präsentation	Marktforschungsergebnisse präsentieren, Marktforschung professionell verkaufen

adm

Allein die Hauptthemen des Seminarprogramms 2007 zeigen die vielfältigen Anforderungen an Markt- und Sozialforschung und damit an die Ausbildung der dort beschäftigten Personen. Dies geht vom Grundlagenwissen über breites statistisches Wissen hin zu Spezialfeldern wie der Preisforschung, Marktprognosemodellen etc. und schließlich hin zu Rechtsfragen, Fragen des Datenschutzes.

Seminarprogramme wie die des BVM stellen sicher, dass regelmäßige Teilnehmer an diesen Seminaren das notwendige praxisorientierte Grundwissen erlangen.

Ein weiteres, recht neues Beispiel ist der ins Leben gerufene Ausbildungsberuf des Fachangestellten für Markt- und Sozialforschung. Hier handelt es sich um eine klassische duale Ausbildung (also Betrieb und Berufsschule), die erstmals zum 1. August 2006 angeboten wurde. Ziel ist die systematische und umfassende Vermittlung des handwerklichen Basiswissens für die Markt- und Sozialforschung. Der Ausbildungsrahmenplan darf als eindrucksvoller Beleg für die Vielfalt der Anforderungen und der zu vermittelnden Lerninhalte gelten.

Der aktuelle Status: Im Mai 2007 haben an 5 Standorten (Düsseldorf, Frankfurt, Hamburg, Köln, Nürnberg) mittlerweile 63 Auszubildende diesen Berufsweg gewählt. Obwohl die Hochschul-Zugangsberechtigung nicht zwingend ist, weisen 75 % dieser Auszubildenden die entsprechende Zugangsberechtigung auf. Von den über 60 ADM-Instituten haben sich spontan 9 Institute entschieden,

sofort selbst auszubilden. Dabei gingen zwischen 4 und 35 Bewerbungen ein: Ein Zeichen für die existierende Nachfrage. Die Industrie- und Handelskammern vor Ort, die Bundesagentur für Arbeit und die Berufsschulen haben von Anfang an engagiert mitgearbeitet und tun dies weiter. Erste Zwischenprüfungen liegen im September 2007 an, die ersten Abschlussprüfungen schon im Mai 2008. Sehr kurzfristig wird also zu erkennen sein, dass es sich hier um den richtigen Ansatz für die systematische Vermittlung von Wissen und damit Qualität in der Markt- und Sozialforschung handelt (siehe auch Anhang: Berufsbildbeschreibung und Ausbildungsrahmenplan).

Resümee

Bei der Markt- und Sozialforschung handelt es sich um angewandte Wissenschaft unter Berücksichtigung vieler interdisziplinärer Verknüpfungen bei gleichzeitig eigenem originären Kern an Fachwissen. Den (wachsenden) Anforderungen der Anwender von Forschung kann nur mit qualifizierter Aus- und Weiterbildung entsprochen werden. Nur so kann auch einem zunehmenden Risiko begegnet werden: dem wachsenden Anteil von Auftraggebern/ Anwendern, die die Notwendigkeit hoher Qualität nicht mehr einschätzen können und zum Teil Qualitäts-Anforderungen reduzieren. Ein Beispiel für diese falsche Entwicklung sind Do-it-yourself-Angebotspakete. Dieses unzureichende Qualitätsbewusstsein gefährdet die Profession.

Dies macht neben der Öffentlichkeitsarbeit, die u. a. auf den wissenschaftlichen Hintergrund und die Qualitätskriterien einzugehen hat, eine anspruchsvolle Ausbildung wichtiger als je zuvor. Leider ist die Ausbildung nicht einheitlich und systematisch – wenn auch auf insgesamt hohem Niveau. Sie wird zurzeit getragen von Hochschulen und Fachhochschulen, den Ausbildungsbetrieben sowie Verbänden und Organisationen. Eine qualitätssichernde und richtige Maßnahme ist hierbei der neue Ausbildungsberuf zum Fachangestellten für Markt- und Sozialforschung.

Die Herausforderung und gemeinsame Zukunftsaufgabe: Hochschulen, Unternehmen und Verbände müssen verstärkt die Verantwortung für eine adäquate und damit qualitätssichernde Aus- und Weiterbildung übernehmen: dies im engen gegenseitigen Kontakt.

Anhang

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

- Das Berufsbild der Fachangestellten für Markt- und Sozialforschung
  - planen, organisieren und steuern die Durchführung von Forschungsprojekten im Rahmen von Projektvorgaben
  - wenden Methoden, Erhebungstechniken und Untersuchungstypen an, recherchieren, beschaffen Daten und werten sie aus
  - wirken an der Konzeption von Fragebögen und Gesprächsleitfäden mit
  - wirken bei Auswahl, Schulung und Einsatz von Interviewern mit
  - bereiten Daten auf, kodieren und führen Plausibilitätsprüfungen durch
  - wenden Auswertungsverfahren an und führen Basisauswertungen durch
  - bereiten Untersuchungsberichte und Präsentationen vor
  - erteilen Auskünfte auch in einer Fremdsprache
  - führen qualitätssichernde Maßnahmen durch
  - wenden Informations- und Kommunikationssysteme an

adm

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

Ausbildungsrahmenplan

Der Ausbildungsbetrieb	**- Stellung, Rechtsform und Struktur des Ausbildungsbetriebes** **- Berufsbildung, arbeits-, sozial- und tarifrechtliche Grundlagen** **- Sicherheit und Gesundheitsschutz bei der Arbeit** **- Umweltschutz**
Arbeitsorganisation, Informations- und Kommunikationssysteme	**- Arbeitsorganisation** **- Informations- und Kommunikationssysteme** **- Datenschutz und Datensicherheit** **- berufsbezogene Rechtsanwendung**
Kommunikation und Kooperation	**- Kundenorientierte Kommunikation, Teamarbeit und Kooperation** **- Anwenden einer Fremdsprache bei Fachaufgaben**

adm

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

Ausbildungsrahmenplan

Aufgaben, Funktionen und Anwendungen der Markt- und Sozialforschung	**- Markt- und Sozialforschung in betriebliche Prozesse und gesamtgesellschaftliche Zusammenhänge einordnen sowie Anwendungsgebiete definieren** **- … von forschungsfremden Tätigkeiten im Rahmen von Werbung, Direktmarketing, Verkaufsförderung und Öffentlichkeitsarbeit abgrenzen …** **- Methoden, Erhebungstechniken und Untersuchungstypen der qualitativen und quantitativen Primärforschung sowie der Sekundärforschung unterscheiden und ihren Einsatz begründen**
Projektvorbereitung a) Informationsbeschaffung und Aufbereitung	**- Daten sekundärer Informationsquellen ziel- und sachgerecht auswählen, auswerten …** **- Vorhandene Untersuchungen, Untersuchungsfragen, Fragebögen und Leitfäden zum Untersuchungsgegenstand beschaffen und auf Verwertbarkeit prüfen** **- Quellen für Stichprobenziehungen festlegen**

adm

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

Ausbildungsrahmenplan

Projektvorbereitung b) Planung und Organisation	**- Methoden, Erhebungstechniken und Untersuchungstypen auf Eignung prüfen** **- Auswahl, Schulung und Einsatz von Interviewern oder Moderatoren vorbereiten** **- Fragebögen und Gesprächsleitfäden auf Mängel prüfen** **- Fragebögen und Gesprächsleitfäden gestalten** **- Kapazitäten, Zeitbedarf und Termine planen, Projektablaufplan erstellen und abstimmen** **- Informationen für die Kalkulation von Projekten einholen** **- Verfahren der Stichprobenziehung unterscheiden, insbesondere unter Berücksichtigung von Stichprobengrößen, Proportionalität sowie Ziehungs- und Auswahlverfahren** **- Probeinterviews vorbereiten, durchführen und Schlussfolgerungen für die Erhebung ziehen**

adm

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

Ausbildungsrahmenplan

Projektdurchführung **a) Prozessbegleitung**	**- Umsetzung des Projektablaufplans koordinieren** **- Prozessschritte eines Projektes … durchführen** **- Projektfortschritt kontrollieren …**
b) Datenerfassung und Codierung	**- Codeplan erstellen** **- Offene und teiloffene Fragen** **- Wörtliche Nennungen transkribieren, klassifizieren und auswerten**
c) Datenprüfung, Gewichtung	**- Plausibilitätsprüfungen durchführen** **- Implausibilitäten listen und bearbeiten** **- Informationen zur Festlegung von Gewichtungsmerkmalen und Gewichtungsmatrix beschaffen**

adm

Beispiel 2:
Fachangestellte(r) für Markt- und Sozialforschung

DESTATIS wissen.nutzen.

Ausbildungsrahmenplan

d) Datenauswertung	**- Datensätze nach vorgegebenen Spezifikationen und Formaten erstellen** **- Tabelleninhalt und –layout festlegen, Tabellen erstellen** **- Tabellen kontrollieren und Korrekturen veranlassen** **- Verfahren der beschreibenden Statistik anwenden** **- Einsatzfelder der Verfahren der schließenden Statistik unterscheiden** **- Betriebliche Analyseverfahren von Gruppendiskussionen und Einzelexplorationen anwenden**
e) Aufbereitung, Präsentation, Ergebnisbericht	**- Analyseergebnisse … in Form von Tabellen, Grafiken sowie in Textform darstellen** **- Ausgewählte Ergebnisse … grafisch darstellen**
Projektnachbereitung	**- Dokumentation** **- Projektabrechnung**

adm

Anforderungen und Nutzen der Zertifizierung in der Markt- und Sozialforschung

Hariet Köstner

1 Überblick

Im Jahr 2006 wurde die für die Markt-, Meinungs- und Sozialforschung grundlegende Norm, die *DIN ISO 20252 – Begriffe und Dienstleistungsanforderungen* – angenommen. Seit damals ist die Grundlage für eine Zertifizierung der Institute gegeben. Während des Prozesses der Normentwicklung sind die deutschen Verbände sehr aktiv gewesen, sodass eine Reihe von Kriterien, die in Deutschland als gute Marktforschungspraxis erachtet werden, aufgenommen werden konnten.

Beim Umsetzungsprozess sind andere Länder z.T. schon weiter fortgeschritten. Dass man schon weiter ist als in Deutschland, liegt u.a. an unterschiedlichen Positionen, die die nationalen Verbände einnehmen. Da eine Zertifizierung in einigen Ländern Pflicht für die Aufnahme in den Verband ist, steht hier natürlich ein deutlich höherer Druck dahinter. Mit Stand Juli 2007 sind in Deutschland zwei Institute zertifiziert. In den Niederland und UK sind ebenfalls schon Institute zertifiziert worden, Frankreich hat die Vorarbeiten abgeschlossen, sodass die ersten Zertifizierungen in naher Zukunft zu erwarten sind, und auch Spanien ist schon sehr weit fortgeschritten.

Die EFAMRO sieht sich in erster Linie als übergeordnete Plattform für den Informations- und Erfahrungsaustausch, um so eine Harmonisierung der nationalen Zertifizierungsprogramme zu erzielen (vgl. Abb. 1.1).

Der vorliegende Artikel beleuchtet die an die Zertifizierung geknüpften Anforderungen und diskutiert den erhofften Nutzen. Zwischen diesen beiden großen Blöcken stehen ein paar Worte zu falschen Hoffnungen, die an die Zertifizierung geknüpft sein können, sowie Fallstricke bei der Umsetzung.

Abb. 1.1: Fortschritt des Zertifizierungsprozesses

2 Anforderungen

Die Anforderungen, die eine Zertifizierung mit sich bringt, werden im Folgenden in Anforderungen der Zertifizierungsspezifikation, des Zertifizierungsprozesses sowie der Zertifizierungseinrichtungen unterschieden (siehe Abb. 2.1).

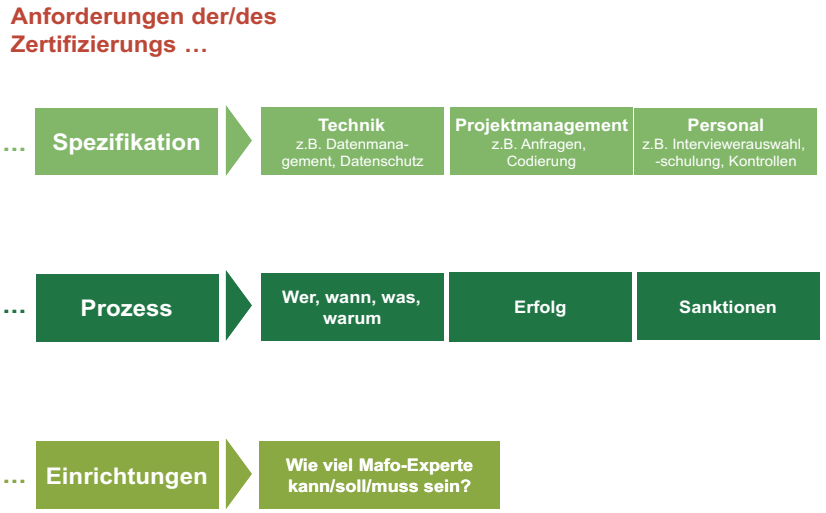


Abb. 2.1: Zertifizierungsanforderungen einzelner Bereiche

Anforderungen der Spezifikation

Jedes Institut wird von den Spezifikationen je nach Unternehmensschwerpunkt unterschiedlich stark tangiert sein. Je nachdem, ob eher techniklastige Methoden wie Online-Befragungen oder eher personalintensive wie Face-to-face- oder CATI-Interviews zum Einsatz kommen, wird im Bereich Technik oder Personal der größere Aufwand für die Zertifizierung liegen. Dazu kommt das Projektmanagement, welches alle Unternehmen gleichermaßen betreffen dürfte. In der Praxis sind die drei Bereiche sicherlich nicht disjunkt, etwaige Überschneidungen werden hier jedoch der Einfachheit halber nicht betrachtet. Es versteht sich von selbst, dass beispielsweise für die Installation technischer Geräte verantwortliche Personen von Nöten sind.

Abb. 2.2: Spezifikationsanforderungen bei Technik und Personal

Wie eben dargestellt, wird der technische Bereich Unternehmen, die ausschließlich computergestützte Umfragen durchführen besonders stark betreffen. Heutzutage stellt dies die Mehrheit der Institute dar, des weiteren müssen auch konventionell erhobene Interviews elektronisch nutzbar gemacht werden. Über dem Bereich Technik „schwebt" die zentrale Anforderung der Datensicherheit sowie der Datenschutz. Dies betrifft z.B. die Aufbewahrungsdauer sowie die Struktur

und Art der zu speichernden Unterlagen. Darüber hinaus muss laut Norm vom Institut der Schutz vor Zerstörung sowie die Vertraulichkeit sichergestellt werden. Schlussendlich muss eine sichere Speicherung bzw. Sicherheitskopien angefertigt sowie eine adäquate Verschlüsselungstechnologie verwendet werden. Neben diesen Aspekten sind Verfahren zur Fehlerlosigkeit bzw. Vollständigkeit der Daten vom Institut zu implementieren. Auf den ersten Blick Selbstverständlichkeiten, kommen hier doch immer wieder Fehler vor. Bei einer Zertifizierung nach DIN ISO 20252 wird nun verlangt, dass so viele Fehlerquellen wie möglich im Vorfeld abgestellt werden. Schlussendlich gehören zum Bereich Technik auch Bereiche, die das Marktforschungsprojekt originär betreffen, nämlich die Stichprobenziehung und etwaige eingesetzte Gewichtungsverfahren.

Der Bereich Personal betrifft primär die eingesetzten Interviewer. Im Rahmen der Zertifizierung sind Verfahren zur Rekrutierung, Schulung, Beurteilung etc. der Interviewer zu entwickeln und festzulegen. Dazu gehören auch die Aufzeichnungen über Abhör- bzw. andere Validierungsmaßnahmen.

Ein Abschnitt der Norm ist der qualitativen Marktforschung gewidmet, in dem die Anforderungen an Moderatoren von beispielsweise Gruppendiskussionen thematisiert werden.

Der dritte Bereich, das Projektmanagement, ist sicherlich der umfassendste und bietet durch seine Vielschichtigkeit und Komplexität die größte Herausforderung. Projektmanagement ist praktisch alles. Es beginnt bei der Anfrage eines Kunden, geht über die Erstellung des Fragebogens, die Feldzeit bis hin zu Auswertung und Präsentation/Beratung des Kunden. Alle Phasen eines Projektes unter Zertifizierungsgesichtspunkten zu beleuchten, würde den Rahmen dieses Artikels bei weitem sprengen. Deshalb soll anhand eines Beispiels die z.T. noch zu leistende Spezifikation der Zertifizierungsanforderungen verdeutlicht werden:

Unter Punkt 6.5.7 wird in der DIN Norm die Überprüfung von gecodeten Projekten festgelegt. Ein Auszug aus dem Abschnitt lautet wie folgt: „Wenn die Arbeit eines Coders häufige Fehler enthält, ist die Arbeit […] zu überprüfen/überarbeiten. Falls nötig, ist dem Coder eine angemessene Nachschulung zukommen zu lassen, bis eine akzeptable Fehlerquote erreicht wird. […]."

Exemplarisch seien einige Worte herausgegriffen, um die Vielschichtigkeit der Spezifikation zu verdeutlichen (vgl. Abb. 2.3):

- „*häufige*" und „*akzeptable*": Es ist nicht nur erforderlich, diesen relativen Begriffen absolute Zahlen zuzuweisen, sondern es ist zunächst einmal die Basis festzulegen. Der festzulegende Prozentwert kann bspw. pro Frage oder pro Studie berechnet werden, wobei letzteres die höheren Anforderungen stellt. Darüber hinaus ist zu überlegen, ob unterschiedlich komplexe Fragetypen unterschiedlich hohe Fehlerquoten aufweisen dürfen. Die Vercodung einer Frage zur spontanen Bekanntheit wie „Welche Firmen aus dem Bereich xy fallen Ih-

nen spontan ein?“ ist sicherlich weniger fehleranfällig als eine offene Frage der Art „Was fällt Ihnen zum Thema xy ein?“.

- „*Fehler*“: Bei der Vercodung können vielerlei Arten Fehler gemacht werden: ein Fehler kann sein, eine Antwort komplett nicht zu vercoden, oder einen Teil nicht, oder auch zuviel zu vercoden. Des Weiteren ist denkbar, einen falschen Code zu verwenden. D.h. auf die Spitze getrieben, ist pro Antwort mehr als ein Fehler, und damit pro Frage eine Fehlerquote über 100% denkbar! Dies muss bei der Festlegung der Grenzwerte bedacht werden.
- „*angemessen*“: Eine Zertifizierung nach DIN ISO 20252 verlangt eine Definition einer angemessenen Nachschulung. Dazu sollte definiert werden, dass bspw. ab x% ein Durchsprechen der Fehler geschieht, ab y% eine Neueinweisung ins Coden und ab z% an den Mitarbeiter keine Codeaufträge mehr vergeben werden.
- „*erreicht*“: Zum Abschluss muss festgelegt sein, in welchem Rahmen die Verbesserung erzielt werden muss. Das kann entweder die kontrollierte Studie selbst, eine andere oder auch lediglich eine Teststudie sein.

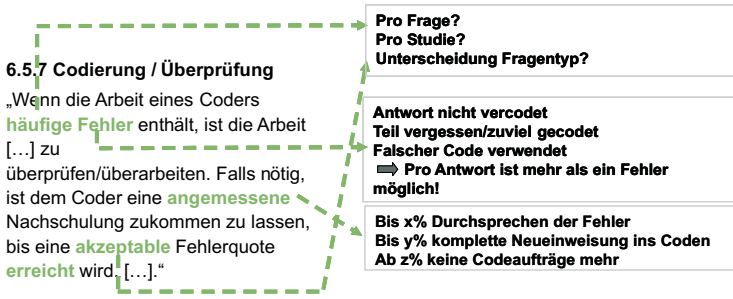


Abb. 2.3: Spezifikationsanforderungen des Projektmanagements

Anhand dieses, sicherlich extremen Beispiels sollte verdeutlicht werden, wie viel Anstrengung im Vorfeld der Zertifizierung notwendig ist, da auf diese Weise sämtliche Prozesse des Unternehmens detailliert durchleuchtet werden müssen.

Anforderungen des Prozesses

Bei den Anforderungen an den Zertifizierungsprozess gilt es auch, eine Vielzahl von Fragestellungen zu beantworten (siehe hierzu Abb. 2.4).

Es beginnt bei der Auswahl der Zertifizierungseinrichtung und der Schulung der Auditoren. Auf diesen Punkt wird im nächsten Abschnitt noch gesondert eingegangen. Weiterhin ist eine ganze Reihe an zeitlichen Fragen zu beantworten, wie z.B. der Zeitaufwand für die Spezifizierung, falls die Anforderungen nicht

unmittelbar auditierbar sind. Dass noch Spezifikationen vonnöten sind, wurde im letzten Abschnitt ausführlich dargelegt. Auch muss von der Zertifizierungseinrichtung festgelegt werden, in welchen Zeitabständen Audits und etwaige Reaudits bzw. Erhaltungsaudits stattfinden.

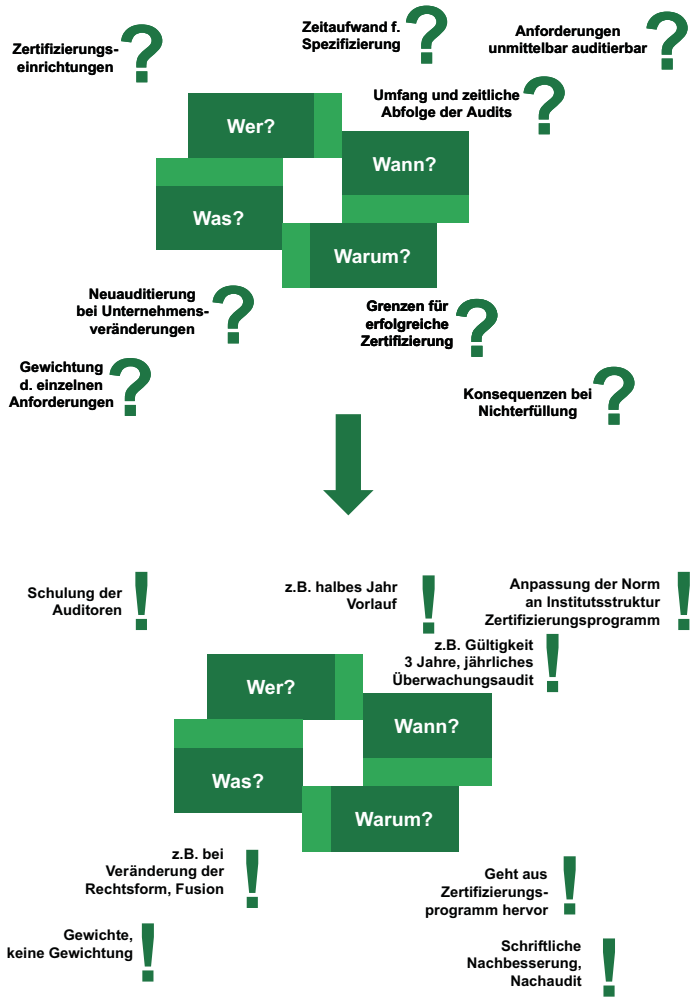


Abb. 2.4: Spezifikationsanforderungen des Zertifizierungsprozesses

Mit die wichtigste Festlegung überhaupt ist die Frage nach der Grenze für eine erfolgreiche Zertifizierung. Hiermit steht und fällt gleichzeitig die Wertigkeit der Zertifizierung und damit mittelfristig ihre Akzeptanz. Eng damit zusammen

hängt die Frage nach den Konsequenzen bei einer etwaigen Nichterfüllung der Mindestanforderungen. Hier sind Maßnahmen wie schriftliche Nachbesserungen, Nachaudits o.ä. denkbar.

Gedanken machen muss man sich auch über eine etwaige Gewichtung der einzelnen Anforderungen. Die einfachste und sicherlich nicht schlechteste Lösung wäre die Gleichgewichtung aller Anforderungen.

Schlussendlich stellt sich die Frage nach der Gültigkeit des Zertifikats bei Unternehmensveränderungen. Auch hier sollten im Vorfeld so viele Konstellationen wie möglich geklärt werden, bspw. Veränderungen der Gesellschaftsform, Unternehmensfusionen (branchenintern und branchenübergreifend) u.v.m.

Anforderungen der Einrichtungen

Die Anforderungen an die Zertifizierungseinrichtungen sind nicht minder herausfordernd, wenngleich auf einer anderen Ebene. Hier geht es weniger um qualitätssichernde Spezifikationen denn um die Zusammenarbeit zweier jeweils betriebswirtschaftlich denkender Einrichtungen. Auf der einen Seite stehen die Zertifizierungsbetriebe, die das gesamte Know-how des Zertifizierungsprozesses mitbringen. „Fusioniert" man einen Mitarbeiter dieser Unternehmen mit jenem eines Marktforschungsinstitutes, welcher das gesamte Spezialwissen des Marktforschungsprozesses mitbringt, ergibt dies den perfekten Auditor für die DIN ISO 20252 (Abb. 2.5).

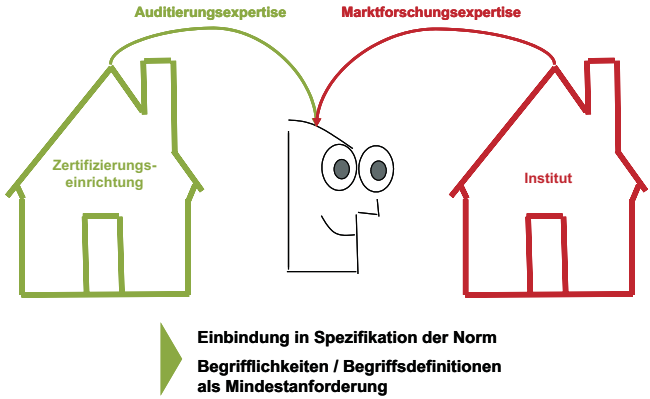


Abb. 2.5: Anforderungen an Auditoren

Einleuchtenderweise wird dies schwerlich möglich sein, ein Auditor ist sicherlich nie der gleiche Experte wie die Mitarbeiter der jeweiligen zu auditierenden Branchen. D.h. es muss ein Kompromiss gefunden werden, der unter Berücksichtigung von Zeit und Kosten ein für beide Seiten optimales Ergebnis bringt. Hierzu

ist es notwendig, dass die Auditoren „die gleiche Sprache“ sprechen und über alle grundlegenden Abläufe im Marktforschungsunternehmen Bescheid wissen. Um dies zu erreichen, ist eine enge Einbindung in den Prozess der Spezifikation der Norm zielführend; des Weiteren sollten zumindest die Begriffsdefinitionen auf den ersten Seiten der Norm verstanden sein.

3 Problemfelder der Zertifizierung

Bevor auf den Nutzen, den eine Zertifizierung mit sich bringt, eingegangen wird, sollen ein paar kritische Worte zu häufigen Irrtümern, falschen Hoffnungen und Fehlern bei der Zertifizierung gesagt werden.

Eine der häufigsten falschen Vorstellungen ist jene, dass mit einer Zertifizierung automatisch hohe Qualität verbunden sein muss. Dies ist nicht notwendigerweise der Fall, da im Rahmen einer Norm zunächst „nur“ Prozesse festgelegt und definiert werden. Die Einhaltung eines Prozesses allein garantiert aber noch keine Qualität.

Für die diskutierte DIN ISO 20252 ist jedoch zu hoffen bzw. anzunehmen, dass sich das Qualitätsniveau auf hohem Level einpendelt, da nicht nur Mindeststandards, sondern „good research practice“ festgelegt wurde.

Sind die Klippen der Anforderungen umschifft, lauern bei der Umsetzung selbst noch eine Reihe von Fallstricken. Eine umfassende Aufarbeitung ist weder Ziel des vorliegenden Artikels noch in dessen Rahmen möglich. Deshalb sollen nur die drei wichtigsten Fehler kurz umrissen werden.

Das erste betrifft die Spezifikation: Wie im ersten Abschnitt ausführlich dargelegt, steht und fällt mit einer sauberen Spezifikation die Güte und damit Akzeptanz der Zertifizierung. Die Gefahr besteht folglich darin, sich auf den kleinsten gemeinsamen Nenner zurückzuziehen, welches auf den ersten Blick sicher die einfachste Lösung darstellt, aber mittelfristig zur Erosion der Norm führt.

Die zweite Gefahr lauert in der Motivation zur Zertifizierung. Wird die Zertifizierung ausschließlich auf Druck von außen oder um ihrer selbst willen durchgeführt, wird ein Großteil des Nutzens vermutlich nicht wie erhofft eintreten. Um wirklich von der Zertifizierung zu profitieren, müssen alle betroffenen Mitarbeiter den Prozess mittragen.

Diese führt zum dritten Punkt, der Einbeziehung aller Betroffenen. Weder darf eine Zertifizierung ausschließlich „von oben“ vorordnet werden, noch kann sie allein auf Initiative der Mitarbeiter und ohne Engagement des Managements erfolgreich sein. In einer Branche wie der Marktforschung, in der noch wenig bis keine Erfahrungen mit Zertifizierungsprozessen vorliegen, ist der letztgenannte Aspekt besonders kritisch, will die Zertifizierung nicht von vornherein zum Scheitern verurteilt sein.

4 Nutzen

Der erhoffte Nutzen betrifft wiederum drei Bereiche: Von außen betrachtet gibt es Betrieb und Branche, die jeweils intern und extern unterschiedlich von der Zertifizierung profitieren. Im Folgenden wird Betriebsextern und Branchenintern als Synonym aufgefasst (dazu Abb. 4.1):

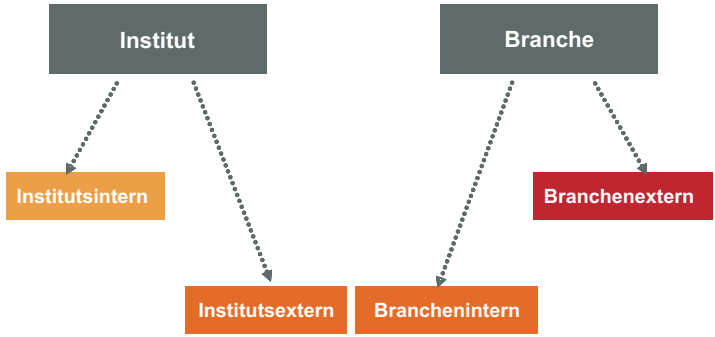


Abb. 4.1: Nutzen der Zertifizierung

Bei allen nachfolgend genannten Nutzenaspekten ist natürlich nicht davon auszugehen, dass sie eins zu eins so eintreffen werden. Es wird quasi der Idealzustand skizziert. Wie viel davon tatsächlich eintritt, ist von vielen Faktoren, nicht zuletzt dem Engagement bei der Umsetzung, abhängig.

Institutsinterner Nutzen

Zunächst sei der betriebsinterne Blickwinkel betrachtet: Was auf den ersten Blick wie zusätzliche Bürokratie aussieht, entpuppt sich auf den zweiten Blick als Optimierung und Vereinheitlichung von Abläufen, Kommunikationsflüssen oder Projektdokumentation (vgl. Abb. 4.2). Der Vorteil bspw. einer einheitlichen Dateistruktur oder einer für alle nachvollziehbaren Kommunikation vergangener Projekte erschließt sich spätestens, wenn man nicht nur in seinen eigenen Projekten etwas sucht, sondern auf die Erfahrung anderer, eventuell auch ehemaliger Kollegen, zurückgreifen möchte. Wie skizziert steigt damit in erster Linie die Effizienz, in zweiter auch die Qualität, da alle Projektbeteiligten gezwungen sind, sich an die Richtlinien zu halten. Durch die beschriebene Vereinheitlichung steigt mittelfristig auch die Mitarbeiterzufriedenheit, da keine Energie in die (erfolglose) Suche nach alten Projektunterlagen gesteckt werden muss. Letztlich ist auch eine Steigerung der Kundenzufriedenheit zu erwarten, ein Aspekt, auf den im

Folgenden bei der Betrachtung der betriebsexternen bzw. brancheninternen Sicht eingegangen wird.

Abb. 4.2: Institutsinterner Nutzen

Institutsexterner/Brancheninterner Nutzen

Für den Kunden ist durch die Schaffung der Standards und der Vereinheitlichung der Abläufe ein höheres Maß an Transparenz gegeben. Dies führt in aller Regel zu höherer Kundenzufriedenheit, da man sich auf einen standardisierten Projektablauf verlassen kann. Darüber hinaus lassen sich aus Kundensicht verschiedene Angebote besser miteinander vergleichen, da Projektabläufe und damit Leistungen bei zertifizierten Instituten in gleicher Weise im Angebot enthalten sind.

Wie schon angesprochen, ist der Nutzen von vielerlei Größen abhängig. Für den betriebsexternen Nutzen gilt, dass er sicherlich direkt abhängig ist von der Akzeptanz, die er bei Auftraggebern und –nehmern erfährt, und diese wiederum ist abhängig u.a. von der Anzahl der zertifizierten Institute (vgl. Abb. 4.3).

Abb. 4.3: Brancheninterner Nutzen

Branchenexterner Nutzen

Der mittelfristig vielleicht größte Nutzen ist im branchenexternen Umfeld zu sehen. Die Marktforschungsbranche hat sich im Laufe der Zeit einige Freiheiten „erarbeitet", die aktuell zu bröckeln drohen. In jüngster Zeit ist ein Stimmungsumschwung sowohl in der Bevölkerung als auch unter den Gerichten zu beobachten, der sich insbesondere gegen die telefonische Marktforschung richtet. Konkret äußert sich das bspw. in geringeren Ausschöpfungsraten. Demonstriert nun die Branche proaktiv ihre Unabhängigkeit und Seriosität durch eine freiwillige Selbstkontrolle, welches ja die Norm darstellt, so kann man dadurch einen positiven Einfluss auf Politik, Gesetzgebung und Rechtssprechung erhoffen. Im Idealfall lässt sich sogar die öffentliche Wahrnehmung positiv beeinflussen, ein Effekt, der sich jedoch sicherlich erst nach einiger Zeit – wenn überhaupt – einstellen wird.

Neben den dargelegten Vorteilen ist auch eine aktive Abgrenzung von anderen Anbietern des Informationssektors wie Adressbrokern oder Direktmarketinganbietern erwünscht. Dies vor allem vor dem Hintergrund, dass letztere nicht unerheblich zum schlechteren Image beitragen und auch für die zunehmend restriktive Rechtssprechung verantwortlich sind. Mit Hilfe der Norm soll folglich sowohl die öffentliche Meinung als auch die Judikative positiv beeinflusst werden (dazu Abb. 4.4):

Abb. 4.4: Branchenexterner Nutzen

5 Ausblick

Wie schon mehrfach angesprochen, krankt die Zertifizierung an der Tatsache, dass sie qualitativ hochwertige Arbeit nicht zwingend nach sich zieht. Der nun gegebene Ausblick soll als Denkanstoß dienen, welche Maßnahmen entweder der

Zertifizierung nachgelagert oder im Rahmen des ohnehin zu implementierenden Qualitätssicherungssystems aufgesetzt werden könnten.

Auf Betreiben von CCBenchmarks – einer internationalen Unternehmensvereinigung, die von Henkel ins Leben gerufen wurde – und unter der konzeptionellen Leitung von ForschungsWerk wurde ein Index entwickelt, der die tatsächlich wahrgenommene Qualität von Anrufen in Call Centern messen soll. Der so genannte „*Service Performance Index*" ist für Inbound Call Center konzipiert und fußt auf der Methode der Eventbasierten Marktforschung.

Ausgangshypothese: Auch ein zertifiziertes Unternehmen liefert nicht notwendigerweise hohe Qualität

Service Performance Index

Gütesiegel, das die Dienstleistungsqualität aus Endkundensicht erfasst

- Qualitätsniveau von Inbound Contact Centern aus Endkundensicht
- Methodische Grundlage: Eventbasierte Marktforschung
- Objektiver Qualitätsindex: Kundensicht + Systemkennzahlen
- Einheitliche, validierte Berechnung des Index
- Konzeption und Durchführung durch CCBenchmarks und ForschungsWerk

Validiertes und standardisiertes Instrument, das sich in Kürze als Maßstab in der CC Welt etablieren sollte

Als Baustein im Rahmen des Qualitätssicherungssystems nach Punkt 3.6 der Norm denkbar

Abb. 5.1: Konzept Service Performance Index

Grundidee ist eine zeitnahe Befragung von Personen, die mit dem Call Center eines Unternehmens in Kontakt getreten sind. Zeitnah bedeutet einerseits so schnell, dass die Erinnerung an das Gespräch noch frisch ist, andererseits aber weit genug weg, dass etwaige Ereignisse, die nicht ursächlich mit dem Call Center in Verbindung stehen, die Beurteilung des Gesprächs nicht überlagern. Ursache des Anrufs kann beispielsweise ein defektes Gerät sein, welches im Allgemeinen Anlass zu Ärger bietet. Im Rahmen des Service Performance Index wird dieser Ärger über das defekte Gerät jedoch so weit wie möglich ausgeblendet und lediglich das Gespräch beurteilt. Mittels des standardisierten und an einer ausreichenden Anzahl Interviews validierten Fragebogens wird besonderes Augen-

merk auf die subjektiv wahrgenommene Qualität des Gesprächs gelegt. D.h. nicht die leicht messbaren Vorgaben wie z.B. den Namen des Anrufers x-mal verwenden o.ä., sondern Dinge wie Freundlichkeit, wahrgenommene Kompetenz, Güte der Problemlösung etc. stehen im Vordergrund. Diese Aspekte bilden den Schwerpunkt des Index, hinzu kommen noch einige wenige Systemkennzahlen wie z.B. durchschnittliche Wartezeit bis zur Annahme des Anrufs o.ä. Letztere fließen aber nur mit einem geringen Gewicht ein, der überwiegende Teil ergibt sich aus der Kundenbeurteilung.

Dieses objektive Instrument ist dabei, sich in der Call Center Welt zu etablieren und dort als Qualitätskriterium der Call Center zu dienen.

Wie eingangs erwähnt, ist ein Einsatz eines ähnlichen Tools im Rahmen der Zertifizierung beispielsweise als Baustein des in Abschnitt 3.6 geforderten internen Qualitätssicherungssystems denkbar.

6 Fazit

Insgesamt befindet sich somit die Markt-, Meinungs- und Sozialforschung auf einem guten Weg, die Qualität ihrer Forschungsarbeit zu dokumentieren und zu festigen. Die Initiative zur DIN ISO 20252 ging von Deutschland aus, und die Marktforschung tut nun daran, wo es um die Umsetzung bzw. Zertifizierung und somit die Ernte der gesäten Früchte geht, diesen Prozess genauso engagiert voranzutreiben wie die Entwicklung der Norm. Damit sind die Weichen für eine weiterhin so erfreuliche Entwicklung der Branche wie in der Vergangenheit gestellt.

Referentinnen und Referenten

Christina Buchwald

Zentrum für Sozialforschung Halle e.V.
an der Martin-Luther-Universität Halle-Wittenberg
Emil-Abderhalden-Straße 6, 06108 Halle (Saale)
Tel.: 0345 / 55 266-14, Fax: 0345 / 55 266-01
E-Mail: buchwald@zsh.uni-halle.de

Christina Buchwald, geb. 1964, Diplom-Soziologin. 1983-1987 Fachschulstudium für Krankenpflege an der Medizinischen Fachschule „Dr. Salvador Allende“ in Halle. Von 1995 bis 2000 Studium der Soziologie an der Martin-Luther-Universität Halle-Wittenberg. Während des Studiums Projektmitarbeit sowie Tätigkeit als wissenschaftliche Hilfskraft und Tutorin am Institut für Soziologie. 2000/2001 wissenschaftliche Mitarbeiterin am Institut für Strukturpolitik und Wirtschaftsförderung Halle-Leipzig e.V. Seit Juli 2001 als wissenschaftliche Mitarbeiterin am Zentrum für Sozialforschung Halle tätig und für den Arbeitsbereich CATI (Computer Assisted Telephone Interviewing) zuständig.

Dr. Manfred Ehling

Statistisches Bundesamt
Gustav-Stresemann-Ring 11, 65189 Wiesbaden
Tel.: 0611/75-2903, Fax: 0611/75-3950
E-Mail: manfred.ehling@destatis.de

Manfred Ehling, geb. 1955, Studium der Sozialwissenschaften an der Ruhr-Universität Bochum 1975 – 1980, Institut für Entwicklungsforschung und Entwicklungspolitik, Ruhr-Universität Bochum, 1981-1982, Wissenschaftlicher Mitarbeiter an der Forschungsstelle für empirische Sozialökonomik e.V., Köln, 1982-1985, seit 1985 beim Statistischen Bundesamt, jetzt Leiter des Instituts für Forschung und Entwicklung in der Bundesstatistik.

Martina Hahn

Statistisches Amt der Europäischen Gemeinschaft (Eurostat)
Building BECH, A4/087, 2920 Luxembourg, Luxembourg
Tel.: +352 4301-35031
E-Mail: martina.hahn@ec.europa.eu

Martina Hahn, geboren 1967; Studium der Volkswirtschaftslehre an der Universität Konstanz und der University of Pennsylvania, USA; 1992 – 2001: Statistisches Bundesamt, Wiesbaden, zuletzt als Leiterin des Referats „Inter- und supranationale Koordinierung“ in der Gruppe IA "Grundsatzfragen der amtlichen Statistik"; 2001 bis heute: Eurostat, Luxemburg, derzeit als Leiterin der Sektion „Statistische Governance und Qualität“.

Olaf Hofmann

Deutsche Gesellschaft für Online-Forschung e.V.
c/o SKOPOS GmbH & Co. KG
Hans-Böckler-Straße 163, 50354 Hürth
Tel.: 02233/518442, Fax: 02233/518449
E-Mail: olaf.hofmann@skopos.de

Dipl.-Psych. Olaf Hofmann gründete 1995 das Full Service Marktforschungsinstitut SKOPOS in Hürth bei Köln. Die Hauptaufgabengebiete von Herrn Hofmann liegen in den Bereichen Onlinemarktforschung, Access Panels sowie Entwicklung von Qualitätsstandards.

Er studierte an der Universität Bonn Psychologie und setzte dabei Schwerpunkte in den Bereichen Methoden- und Medienforschung sowie Organisationspsychologie. Von 2001 bis 2004 war der Vater dreier Töchter Vorsitzender der deutschen Gesellschaft für Online-Forschung (DGOF). Olaf Hofmann veröffentlicht Fachpublikationen und hält Vorträge zu den Themen Kommunikations- und Online-Forschung sowie Qualitätsstandards in der Markt- und Sozialforschung. Er ist seit 2004 Trainer im Rahmen der BVM Seminarreihe für die Seminare Onlinemarktforschung sowie Apparative und Innovative Verfahren. Darüber hinaus ist Herr Hofmann Mitglied der Deutschen Delegation im Technical Committee von ISO TC 20252.

Dr. Hariet Köstner

ForschungsWerk GmbH
Nordring 100, 90409 Nürnberg
Tel.: 0911/9908-105, Fax: 0911/9908-111
E-Mail: koestner@forschungswerk.de

Hariet Köstner, geb. 1974, Studium der Diplom-Sozialwissenschaften an der Friedrich-Alexander-Universität Erlangen-Nürnberg (1994-1999), 2003 Promotion dort am Lehrstuhl für Statistik und Ökonometrie in Kooperation mit Information Resources GfK GmbH.

Information Resources GfK GmbH, Nürnberg (2000 – 2003), Junior Specialist, Analytical Consulting Group; Schwerpunkt der Tätigkeit: Konzeption und

Weiterentwicklung von Modellen für die Analyse des Verkaufserfolges von FMCG's aus Scannerdaten. Seit 09/2003 ForschungsWerk GmbH, Nürnberg, Leitung Statistik & Methoden; Schwerpunkte der Tätigkeit: Projektleitung Ad hoc und Tracking Studien; Statistische Analyse von Datensätzen, Weiterentwicklung von Analyse-Tools; Betreuung von Diplomarbeiten; Kooperation mit der Universität.

Lehraufträge an der Friedrich-Alexander-Universität Erlangen-Nürnberg, sowie der Hochschule Pforzheim zu statistischen und marktforschungsrelevanten Themen.

Dr. Silke I. Keil

Institut für Sozialwissenschaften
Universität Stuttgart
Abteilung für Politische Systeme und Politische Soziologie
Breitscheidstr. 2, 70174 Stuttgart
Tel.: 0711/ 685-83643, Fax: 0711/ 685-82333
E-Mail: silke.keil@sowi.uni-stuttgart.de

Silke I. Keil, geb. 1967, seit 2002 wissenschaftliche Mitarbeiterin am Institut für Sozialwissenschaften, Abteilung für Politische Systeme und Politische Soziologie. Seit September 2006: Geschäftsführung für die deutsche Teilstudie des European Social Survey. Bildungsgang: Studium der Politikwissenschaft und Germanistik an der Universität Stuttgart. 2002 Promotion zur Dr. rer. pol. mit einer Arbeit zum Thema „Wahlkampfkommunikation in Wahlanzeigen und Wahlprogrammen. Eine vergleichende inhaltsanalytische Untersuchung der von den Bundestagsparteien vorgelegten Wahlprogramme und Wahlanzeigen in den Bundestagswahlkämpfen 1957 bis 1998".

Andrea Kron

Statistisches Bundesamt
Gustav-Stresemann-Ring 11, 65189 Wiesbaden
Tel.: 0611/75-4683, Fax: 0611/75-4000
E-Mail: andrea.kron@destatis.de

Andrea Kron, geb. 1982, 2001-2007 Studium der Politikwissenschaft, Soziologie und VWL (M.A.) an der Johannes Gutenberg-Universität Mainz, seit 2007 wissenschaftliche Mitarbeiterin beim Statistischen Bundesamt, Institut für Forschung und Entwicklung in der Bundesstatistik.

Katja Lukanow

Zentrum für Sozialforschung Halle e.V.
an der Martin-Luther-Universität Halle-Wittenberg
Emil-Abderhalden-Str. 6, 06108 Halle/Saale
Tel.: 0345/55-2 66 22, Fax: 0345/55 2 66 01
E-Mail: katja.lukanow@zsh.uni-halle.de

Katja Lukanow, geb. 1978, Diplom-Soziologin. Von 1997-2003 Studium der Soziologie an der Martin-Luther-Universität Halle-Wittenberg. Während des Studiums Tätigkeit als wissenschaftliche Hilfskraft am Zentrum für Sozialforschung Halle. 2004 Auslandsaufenthalt. Seit Januar 2005 ist sie wissenschaftliche Mitarbeiterin am Zentrum für Sozialforschung Halle und ebenfalls für den Arbeitsbereich CATI (Computer Assisted Telephone Interviewing) zuständig.

Dr. Holger Mühlbauer

DIN Deutsches Institut für Normung e.V.
Burggrafenstraße 6, 10787 Berlin
Tel.: 030/2601 2809, Fax: 030/2601 4 2809
E-Mail: holger.muehlbauer@din.de

Holger Mühlbauer (geboren 1964) kam nach Lehre und Arbeit als Maschinenschlosser, Jura-Studium, Assistententätigkeit an der TU Berlin und Beschäftigung bei der Kassenärztlichen Vereinigung Berlin zum Deutschen Institut für Normung e.V. und ist u.a. Geschäftsführer des Normenausschusses Gebrauchstauglichkeit und Dienstleistungen (NAGD) im DIN Deutsches Institut für Normung e.V. mit dem Schwerpunkt „Normung von Endverbraucherprodukten und Dienstleistungen“. Er koordiniert beim DIN nationale, europäische und internationale Normungsverfahren auf diesem Gebiet.

Dr. Katja Neller

Universität Stuttgart
Institut für Sozialwissenschaften
Breitscheidstr. 2, 70174 Stuttgart
Tel.: 0711/685-83427, Fax: 0711/685-82333
E-Mail: katja.neller@sowi.uni-stuttgart.de

Katja Neller, geb. 1972, seit 1998 wissenschaftliche Mitarbeiterin am Institut für Sozialwissenschaften der Universität Stuttgart, Abteilung für Politische Systeme und Politische Soziologie. Bis September 2006 Geschäftsführung für die deutsche Teilstudie des European Social Survey. Aktuelle Projekte: „Der Einfluss des

politischen Vertrauens auf die Unterstützung der Reform des Sozialstaates in Deutschland“; wissenschaftliche Leitung einer Absolventenstudie an der Universität Stuttgart.

Bildungsgang: Studium der Politikwissenschaft und Geschichte an der Universität Stuttgart. 2006 Promotion zur Dr. rer. pol. mit einer Arbeit zum Thema „DDR-Nostalgie. Dimensionen der Orientierungen der Ostdeutschen gegenüber der ehemaligen DDR, ihre Ursachen und politischen Konnotationen“.

Michael Pusler

Hubert Burda Media
Marketing & Communications GmbH
Arabellastraße 23, 81925 München
Tel.: 089/ 92 50 26 12
E-Mail: michael.pusler@burda.com

Michael Pusler hat Psychologie (Schwerpunkt Wirtschaftspsychologie) und Betriebswirtschaftslehre in Konstanz, Köln und Mannheim studiert. Seine beruflichen Stationen führten ihn über die GIM (Heidelberg), Infratest Wirtschaftsforschung (München) zu Hubert Burda Media. Dort ist er seit 2004 stellvertretender Leiter der Konzernmarktforschung und seit 2007 zudem für die neu geschaffene Unit „Entwicklung neuer Forschungsmethoden & Data Intelligence“ verantwortlich.

Er ist Mitglied der technischen Kommission der Arbeitsgemeinschaft Media-Analyse (ag.ma) sowie des technischen Ausschusses der Leseranalyse Entscheidungsträger (LAE).

Seine inhaltlichen Schwerpunkte liegen im Bereich strategischer Forschungsfelder wie der inhaltlichen Konzeption und Entwicklung von Markt-Media-Studien, der Verbraucher-, Medien- und Werbewirkungsforschung sowie der Neuromarktforschung.

Hartmut Scheffler

TNS Infratest Holding GmbH & Co. KG
Stieghorster Str. 66, 33605 Bielefeld
Tel.: 0521/92 57-330, Fax: 0521/92 57-250
E-Mail: hartmut.scheffler@tns-infratest.com

Hartmut Scheffler, geb. 1955, Diplom-Soziologe und Stadt-, Raum- und Regionalplaner, Studienvertiefungen in VWL und Sozialpsychologie. Von 1979 bis 1980 Mitarbeiter des Stadtplanungsamtes Bielefeld, danach Eintritt im damaligen Emnid Institut. Seit 1990 Geschäftsführer des Emnid Institutes, später TNS Emnid. Nach vollzogenem Merger zwischen TNS Emnid und NFO Infratest zur

TNS Infratest Holding GmbH & Co. KG seit Januar 2006 Geschäftsführer der TNS Infratest Holding GmbH & Co. KG. Mitglied bei BVM, ESOMAR, Wissenschaftlicher Gesellschaft Münster, G.E.M. usw. Vorstandsmitglied der Initiative D21 sowie Vorstandsvorsitzender des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.

Ulrike Timm

Statistisches Bundesamt
Gustav-Stresemann-Ring 11, 65189 Wiesbaden
Tel.: 0611/75-4608, Fax: 0611/75-3950
E-Mail: ulrike.timm@destatis.de

Ulrike Timm, geb. 1972, 1994-2001 Studium der Soziologie an der Universität Hamburg, seit 2002 wissenschaftliche Mitarbeiterin beim Statistischen Bundesamt, Institut für Forschung und Entwicklung in der Bundesstatistik.

Erich Wiegand

ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.
Langer Weg 18, 60489 Frankfurt
Tel.: 069/97 84 31 36, Fax: 069/97 84 31 37
E-Mail: wiegand@adm-ev.de

Erich Wiegand, geb. 1950, Diplom-Soziologe, Studium an der J.W. Goethe-Universität in Frankfurt am Main; seit 1993 Geschäftsführer des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.; seit 2001 Geschäftsführer des Rates der Deutschen Markt- und Sozialforschung e.V.; seit 2005 Präsident von EFAMRO European Federation of Associations of Market Research Organisations; Leiter der deutschen Delegation im ISO Technischen Komitee 225 „Markt-, Meinungs- und Sozialforschung“; Convenor der Arbeits-gruppe 2 „Access Panels“ im ISO TK 225; Convenor der CEN Task Force 186 „Printmedienanalysen“.

Teilnehmerliste

B

Thomas Bachmann, *TNS Infratest Holding GmbH & Co. KG*, München
Jürgen Bamberger, *Statistisches Bundesamt*, Wiesbaden
Thomas Beetz, *Statistisches Bundesamt*, Wiesbaden
Stefan Bek, *Arbeitsgemeinschaft Media-Analyse e.V./ Media-Micro-Census GmbH*, Frankfurt/Main
Irmtraud Beuerlein, *Statistisches Bundesamt*, Wiesbaden
Karen Blanke, *Statistisches Bundesamt*, Wiesbaden
Petra Brien, *GESIS - ZA*, Köln
Christina Buchwald, *Zentrum für Sozialforschung Halle e.V. an der Martin-Luther-Universität Halle-Wittenberg*, Halle/Saale

C

Dr. Laura Castiglioni, *Universität Mannheim, Mannheimer Zentrum für Europäische Sozialforschung (MZES)*, Mannheim
Jürgen Chlumsky, *Statistisches Bundesamt*, Wiesbaden

D

Jörg Decker, *Statistisches Bundesamt*, Wiesbaden
Silvia Deckl, *Statistisches Bundesamt*, Bonn
Marc Deutschmann, *BIK MARPLAN Intermedia GmbH*, Offenbach/Main
Dr. Britta Dollinger, *Amt für Wahlen, Statistik und Stadtforschung*, Wiesbaden
Bonno Domke, *Hessisches Statistisches Landesamt*, Wiesbaden

E

Beate Ebel, *Amt für Statistik Berlin-Brandenburg*
Dr. Manfred Ehling, *Statistisches Bundesamt*, Wiesbaden
Maria Ehrenberg, *Medienwissenschaftliches Lehr- und Forschungszentrum der Universität zu Köln*
Dr. Uwe Engfer, *Technische Universität Darmstadt, FB Institut für Soziologie*

F

Jens Fischer-Kottenstede, *Hessisches Kultusministerium, Referat IV.6 - Statistik und Informationssysteme*, Wiesbaden
Hubert Förster, *Media Markt Analysen GmbH & Co. KG*, Frankfurt/Main

Ulrike Försterling, *Statistisches Landesamt Sachsen-Anhalt*, Halle/Saale
Sigrid Fritsch, *Statistisches Bundesamt*, Wiesbaden

G

Daniel Gebhardt, *Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit*, Nürnberg
Anke Gerhardt, *Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen*, Düsseldorf
Axel Glemser, *TNS Infratest Holding GmbH & Co. KG*, München
Stefan Gruber, *Statistisches Bundesamt*, Bonn
Winfried Gruber, *Statistisches Landesamt Baden-Württemberg*, Stuttgart
Ingrid Grüning, *Arbeitsgemeinschaft Media-Analyse e.V./ Media-Micro-Census GmbH*, Frankfurt/Main

H

Martina Hahn, *Statistisches Amt der Europäischen Gemeinschaften (Eurostat)*, Luxembourg
Prof. Dr. Barbara Hedderich, *Fachhochschule Ansbach*
Frank Hellwig, *Statistisches Bundesamt*, Wiesbaden
Ulrich Hertfelder, *ICON ADDED VALUE GmbH*, Nürnberg
Erika Herz, *ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.*, Frankfurt/Main
Olaf Hofmann, *Deutsche Gesellschaft für Online-Forschung e.V. c/o SKOPOS GmbH & Co. KG*, Hürth

J

Dr. Andrea Janßen, *GESIS - ZUMA*, Mannheim
Dr. Roland Jeske, *Statistisches Landesamt Rheinland-Pfalz*, Bad Ems
Heiner Junker, *Produkt + Markt GmbH & Co. KG*, Wallenhorst

K

Dr. Irene Kahle, *Statistisches Bundesamt*, Wiesbaden
Dr. Silke I. Keil, *Institut für Sozialwissenschaften, Universität Stuttgart*
Manfred Kiesel, *Deutsche Bundesbank*, Frankfurt/Main
Carolin Kieslich, *GfK AG - Ad Hoc Services*, Nürnberg
Cornelia Koch, *Statistisches Bundesamt*, Wiesbaden
Christian König, *Statistisches Bundesamt*, Wiesbaden
Dr. Hariet Köstner, *ForschungsWerk GmbH*, Nürnberg

Andrea Kron, *Statistisches Bundesamt*, Wiesbaden

L

Brit Liestmann, *BIK ASCHPURWIS + BEHRENS GMBH Markt-, Media und Regionalforschung*, Hamburg

Stefan Lorger, *Statistisches Bundesamt*, Wiesbaden

Katja Lukanow, *Zentrum für Sozialforschung Halle e.V. an der Martin-Luther-Universität Halle-Wittenberg*, Halle/Saale

M

Prof. Dr. Gerd Meier, *Universität Lüneburg*

Jörg Meier, *Deutsche Bundesbank*, Frankfurt/Main

Prof. Dr. Hans-Joachim Mittag, *FernUniversität Hagen, Fachbereich Wirtschaftswissenschaft*

Dr. Holger Mühlbauer, *DIN Deutsches Institut für Normung e.V.*, Berlin

Berthold Müller, *Hessisches Statistisches Landesamt*, Wiesbaden

Dr. Katja Neller, *Universität Stuttgart, Institut für Sozialwissenschaften*

N

Marianne Neuendorf, *Statistisches Bundesamt*, Wiesbaden

O

Sibylle von Oppeln-Bronikowski, *Statistisches Bundesamt*, Wiesbaden

Boris Orth, *Bundeszentrale für gesundheitliche Aufklärung (BZgA), Referat 2-25*, Köln

P

Dr. Frank Padberg, *Hessisches Statistisches Landesamt*, Wiesbaden

Hannelore Pöschl, *Statistisches Bundesamt*, Bonn

Michael Pusler, *Hubert Burda Media Marketing & Communications GmbH*, München

R

Walter Radermacher, *Präsident des Statistischen Bundesamtes*, Wiesbaden

Uwe Reim, *Statistisches Bundesamt*, Wiesbaden

Maren Risel, *Medienwissenschaftliches Lehr- und Forschungszentrum der Universität zu Köln*

Gabriele Ritter, *ag.ma/MMC*, Frankfurt/Main

S

Ralf Saathoff, *Millward Brown Germany GmbH*, Frankfurt/Main
Sabine Sattelberger, *Statistisches Bundesamt*, Wiesbaden
Angela Schaff, *Statistisches Bundesamt*, Wiesbaden
Hartmut Scheffler, *TNS Infratest Holding GmbH & Co. KG*, Bielefeld
Pia Scherhaufer, *Ipsos GmbH*, Mölln
Jürgen Schmidt, *Statistisches Bundesamt*, Wiesbaden
Marco Schwickerath, *Statistisches Bundesamt*, Wiesbaden
Till Seeck, *Statistisches Bundesamt*, Wiesbaden
Ines Seiler, *Bundesministerium für Verkehr, Bau- und Stadtentwicklung*, Bonn
Prof. Dr. Wolfgang Sodeur, *Universität Duisburg-Essen*
Volker Stander, *Bundeszentrale für gesundheitliche Aufklärung (BZgA), Referat 2-25*, Köln

T

Ulrike Timm, *Statistisches Bundesamt*, Wiesbaden

U

Dr. Alexandra Uhly, *Bundesinstitut für Berufsbildung*, Bonn

W

Wiebke Walingen, *Compagnon Marktforschung*, Stuttgart
Mario Walter, *Bayerisches Landesamt für Statistik und Datenverarbeitung*, München
Martina Wasmer, *GESIS - ZUMA*, Mannheim
Oliver Watteler, *GESIS - ZA*, Köln
Thomas Wegertseder, *Deutsches Jugendinstitut*, München
Helga Weißbecker, *ZMG Zeitungs Marketing Gesellschaft mbH & Co. KG*, Frankfurt/Main
Erich Wiegand, *ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.*, Frankfurt/Main
Thomas Wöll, *Statistisches Bundesamt*, Wiesbaden

Z

Anouk Zabal, *GESIS - ZUMA*, Mannheim
Jacek Zwania, S*tatistisches Bundesamt*, Wiesbaden

Der vorliegende Tagungsband dokumentiert die Beiträge der wissenschaftlichen Tagung „Qualitätsmanagement und Qualitätssicherung“, die am 28. und 29. Juni 2007 gemeinsam vom Statistischen Bundesamt, dem ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. und der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) in Wiesbaden durchgeführt wurde.

gesis

Lennéstraße 30 · 53113 Bonn
Telefon: 02 28 / 22 81 - 0
Telefax: 02 28 / 22 81 - 120

Ein Institut der Leibniz-Gemeinschaft

ISBN 978-3-86819-000-7
ISSN 1865-7559
EUR 10,-