Beiträge zur Gesundheitsberichterstattung des Bundes

Erster telefonischer Gesundheitssurvey des Robert Koch-Instituts – Methodische Beiträge

Martin Kohler, Angelika Rieck, Susan Borch, Thomas Ziese

Robert Koch-Institut, Berlin 2005
# Inhaltsverzeichnis

1 Der telefonische Gesundheitssurvey 2003 – Baustein zum Gesundheitsmonitoringsystem für Deutschland .............................. 5
   Thomas Ziese
   1.1 Einleitung ......................................................... 5
   1.2 Themen ............................................................. 6
   1.3 Eckdaten ............................................................ 6
   1.4 Vorteile, Grenzen und Ziele ........................................ 7
   1.5 Literatur .............................................................. 8

2 Methodische Aspekte der Fragebogenentwicklung ............................... 9
   Angelika Rieck, Susan Borch, Martin Kohler
   2.1 Einleitung ............................................................. 9
   2.2 Notwendigkeit der Berücksichtigung der Erhebungsmethode ................ 9
   2.3 Maßnahmen zur Entwicklung des Fragebogens .............................. 10
   2.3.1 Fragebogendramaturgie ........................................... 10
   2.3.2 Filterführung ....................................................... 13
   2.3.3 Optimierung der Operationalisierungen .............................. 13
   2.3.3.1 Mehrdimensionale bzw. komplexe Fragen ....................... 14
   2.3.3.2 Umfangreiche bzw. komplexe Antwortkategorien ................ 15
   2.3.3.3 Schwer memorierbare Antworten ................................ 16
   2.3.3.4 Reihenfolge von Antwortvorgaben ............................. 16
   2.3.3.5 Hinweistexte und Administrationsanweisungen .................. 17
   2.3.4 Weitere Techniken/Maßnahmen ..................................... 18
   2.4 Fazit ................................................................. 18
   2.5 Literatur .............................................................. 19

3 Feldarbeit ................................................................. 21
   Susan Borch, Angelika Rieck
   3.1 Einleitung ............................................................. 21
   3.2 Personal .............................................................. 21
   3.3 Datenerhebung und Datenhaltung ....................................... 21
   3.4 Mitarbeiterschulung ................................................. 23
   3.4.1 Interviewer ......................................................... 23
   3.4.2 Supervisoren ....................................................... 25
   3.5 Qualitätssicherung ................................................... 25
   3.5.1 Interviewführung und Schulung .................................... 27
   3.5.2 Supervision ......................................................... 28
   3.6 Fazit ................................................................. 29
   3.7 Literatur .............................................................. 29
<table>
<thead>
<tr>
<th>Kapitel</th>
<th>Titel</th>
<th>Seite</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Stichprobenkonzeption und Stichprobenrealisierung</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Einleitung</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Technische Rahmenbedingungen</td>
<td>31</td>
</tr>
<tr>
<td>4.3</td>
<td>Stichprobendesign</td>
<td>32</td>
</tr>
<tr>
<td>4.4</td>
<td>Anrufmanagement und Stichprobenbearbeitung</td>
<td>33</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Das Anruf- und Rückrufmanagement</td>
<td>35</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Prozesssteuerung im Call Back Management</td>
<td>37</td>
</tr>
<tr>
<td>4.5</td>
<td>Datenhaltung</td>
<td>44</td>
</tr>
<tr>
<td>4.6</td>
<td>Sampleausschöpfung und Response</td>
<td>44</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Leitende Kriterien für die Klassifikation der disposition codes</td>
<td>45</td>
</tr>
<tr>
<td>4.7</td>
<td>Datenkontrolle und -aufbereitung</td>
<td>48</td>
</tr>
<tr>
<td>4.8</td>
<td>Gewichtung der Daten</td>
<td>49</td>
</tr>
<tr>
<td>4.8.1</td>
<td>Designgewichtung</td>
<td>49</td>
</tr>
<tr>
<td>4.8.2</td>
<td>Anpassungsgewichtung</td>
<td>50</td>
</tr>
<tr>
<td>4.8.3</td>
<td>Güte der Gewichtung</td>
<td>50</td>
</tr>
<tr>
<td>4.9</td>
<td>Ergenisse zu Response und Nonresponse</td>
<td>51</td>
</tr>
<tr>
<td>4.9.1</td>
<td>Bekannte Gründe für Nonresponse</td>
<td>51</td>
</tr>
<tr>
<td>4.9.2</td>
<td>Analyse soziostruktureller und gesundheitsbezogener Merkmale von</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Teilnehmern und Nicht-Teilnehmern</td>
<td></td>
</tr>
<tr>
<td>4.9.3</td>
<td>Intensität der Samplenutzung</td>
<td>53</td>
</tr>
<tr>
<td>4.10</td>
<td>Zusammenfassung</td>
<td>55</td>
</tr>
<tr>
<td>4.11</td>
<td>Literatur</td>
<td>56</td>
</tr>
</tbody>
</table>
1 Der telefonische Gesundheitssurvey 2003 – Baustein zum Gesundheitsmonitoringsystem für Deutschland

Thomas Ziese

1.1 Einleitung


Diese Defizite können die in Deutschland seit Mitte der 1980er Jahre durchgeführten repräsentativen Gesundheitssurveys in weiten Teilen kompensieren. Im Gegensatz zu den meisten anderen Datenquellen lassen sich zudem subjektive Einstellungen, Wahrnehmungen und Verhaltensorientierungen berücksichtigen, die sowohl unter dem Gesichtspunkt der Prävention als auch im Hinblick auf die Heterogenität der Versorgungsbedarfe und der Selbsthilfe im Gesundheitswesen zunehmend an Bedeutung gewinnen.


Telefonische Befragungen haben nicht nur den Vorteil der deutlich geringeren Kosten, sie sind außerdem inhaltlich flexibel und können schneller realisiert werden. In den USA nutzt das Centers for Disease Control (CDC 1998) schon seit über 20 Jahren jährliche repräsentative telefonische Gesundheitssurveys, um die Daten aus den in etwa acht- bis zehnjährigen Abständen stattfindenden nationalen Health Examination Surveys zu ergänzen.

1.2 Themen

Das thematische Spektrum des telefonischen Gesundheitssurveys 2003 (GSTel03) umfasst folgende Bereiche:

**Subjektive Gesundheit**
- Allgemeine subjektive Gesundheit (European Health Module, EHM)
- Gesundheitsbezogene Lebensqualität (SF8)

**Krankheiten**
- Diabetes, -behandlung, -komplikationen
- Krankheiten der Atmungsorgane (chronische Bronchitis, Asthma bronchiale)
- Gelenkerkrankungen (Arthrose, Polyarthritis)
- Osteoporose und -mitteleinnahme
- Menopause
- Rückenschmerzen und -behandlung
- Schwindel
- Krankheiten der Augen (Nutzung von Sehhilfen, Kurz- und Weitsichtigkeit, Glaukom, Katarakt, Makula-Degeneration)
- Krankheiten der Ohren (Nutzung eines Hörgeräts, Tinnitus)
- Psoriasis
- Hepatitis
- Depressionen (Selbsteinschätzung)
- Herz-Kreislauf-Erkrankungen (Hypertonie, erhöhte Blutfette und Cholesterinwerte, Angina pectoris, Herzinfarct, Herzinsuffizienz, Herzrhythmusstörungen)
- Schlaganfall
- Krebserkrankungen

**Behinderung und körperliche Einschränkungen**
- Schwerbehinderung, Grad der Behinderung
- Minderung der Erwerbsfähigkeit, Grad der Minderung der Erwerbsfähigkeit
- Activities of daily life (ADL)

**Soziodemografie**
- Sozialschicht (Bildung, Einkommen, berufliche Stellung)
- Migrationshintergrund
- Arbeitslosigkeit
- Krankenversicherung
- Sozialhilfeempfang


1.3 Eckdaten

1.4 Vorteile, Grenzen und Ziele

Telefonische Befragungen haben nicht nur die bereits erwähnten Vorzüge, zeitnäher und mit geringerem Kostenaufwand auch aktuelle Daten zu Veränderungen in gesundheitsrelevanten Bereichen bevölkerungsrepräsentativ gewinnen zu können. Weitere Vorteile telefonischer Befragungen liegen

- in der Verfügbarkeit eines nationalen Auswahlrahmens durch das Gabler-Häder-Design, das Klumpungseffekte nahezu ausschließt. Damit sind prinzipiell die Möglichkeiten für regionalisierte Analysen gegeben.
- Telefonische Befragungen sind darüber hinaus auch effizienter in der Feldbearbeitung. Schwer erreichbare Bevölkerungsgruppen sind mit geringerem Aufwand besser zu erreichen, was allerdings entsprechend häufige Kontaktversuche voraussetzt.
- Die zentrale Unterbringung der Interviewer in einem Telefonstudio ermöglicht eine effizientere Kontrolle der Interviewdurchführung und verbessert die Datenqualität.


Im Kontext des derzeit im Aufbau begriffenen Monitoringsystems (Kurth et al. 2005) spielen die telefonischen Gesundheitssurveys eine tragende Rolle als Bindeglied zwischen den weiterhin konzeptionell vorgesehenen Examination Surveys. Damit kommt dem GSTel03 die Funktion einer Piloterhebung zu. Das ist auch der Grund, weshalb im vorliegenden Beitrag zur GBE methodische Aspekte bei Telefonbefragungen vertiefend in einem gesonderten Band behandelt werden, für die in der Darstellung der Ergebnisse des GSTel03 sonst nicht genügend Raum gegeben werden könnte.

In den hier vorgestellten Einzelbeiträgen werden zentrale methodische Aspekte computergestützter Telefonsurveys behandelt, die nicht nur für den GSTel03, sondern auch für künftige Gesundheitssurveys von Bedeutung sind.

Im ersten Beitrag geht es um die Notwendigkeit eines erhebungsmethodisch angemessenen Fragebogens und die dazu im GSTel03 umgesetzten Maßnahmen. Zu diesen zählen Fragebogendramaturgie, Filterführung und Optimierung der Operationalisierungen mit den Schwerpunkten Zerlegung von komplexen Fragen bzw. Antwortkategorien, Festlegung der Reihenfolge für die Antwortvorgaben und Vereinheitlichung retrospektiver Zeitfenster.


Der letzte Beitrag geht ausführlich auf die technischen Rahmenbedingungen, das Stichprobendesign, den Feldverlauf bzw. die Ablaufsteuerung, die Datenaufbereitung und daraus abgeleitet auf die Gewichtung des GSTel03 ein. Abschließend werden die Ergebnisse einer Non-Responderanalyse vorgestellt und die Auswirkungen der Intensität der Sampleausschöpfung diskutiert.

Die Beiträge zeigen, dass der GSTel03 hohen Qualitätsansprüchen gerecht wird. Das schließt jedoch in einzelnen Fällen nicht aus, dass methodenspezifische Effekte auch Einfluss auf die erzielten Resultate haben können. Das genauer zu bestimmen, bleibt zukünftigen Methodenstudien vorbehalten.
1.5 Literatur

2 Methodische Aspekte der Fragebogenerstellung

Angelika Rieck, Susan Borch, Martin Kohler

2.1 Einleitung


Im Folgenden wird zuerst kurz die Notwendigkeit eines erhebungsmethodisch angemessenen Fragebogens aufgezeigt, die sich insbesondere aus den Unterschieden in der kommunikativen Situation beim Telefoninterview im Vergleich zum face-to-face Interview und zur schriftlichen Befragung ergibt (Friedrichs 1990; Dillman 1978; Frey et al. 1990). Diese methodischen Besonderheiten bedürfen aufgrund ihrer Bedeutung für Datenqualität, Item-Nonresponse und Ausschöpfung bei der Entwicklung des Fragebogens besonderer Beachtung. Im weiteren Verlauf des Artikels werden die wichtigsten Maßnahmen vorgestellt, die dazu für den telefonischen Gesundheitssurvey 2003 (GSTel03) umgesetzt wurden.

2.2 Notwendigkeit der Berücksichtigung der Erhebungsmethode


Deshalb sollte der Fragebogen für Telefonsurveys noch stärker an einen natürlichen Kommunikationsfluss angepasst werden (Friedrichs 1990), sodass beispielsweise Überleitungen im Erhebungsinstrument vorgesehen sind und nicht vom Interviewer improvisiert werden müssen. Außerdem ist es für Telefoninterviews wichtig, den Fragebogen an die Alltagssituation der Probanden anzupassen (Frey et al. 1990), um einen möglichst unmittelbaren, einfachen Zugang zu den gewünschten Informationen zu erreichen. Dabei hat die Fragebogendramaturgie, d. h. die optimale Anordnung einzelner Fragen und Fragengruppen sowie die Integration von Überleitungs- und Filterfragen, eine besondere Bedeutung für die Beantwortung der Fragen und deren Qualität (Fuchs, Lamnek 1990; Frey et al. 1990; Dill-

1 CATI steht für computer-assisted telephone interviewing, also computergestütztes Telefoninterview.
Erster telefonischer Gesundheitssurvey des Robert Koch-Instituts – Methodische Beiträge

10


2.3 Maßnahmen zur Entwicklung des Fragebogens


2.3.1 Fragebogendramaturgie

Nach Dillman (1978) dient die Dramaturgie des Fragebogens bei Telefoninterviews hauptsächlich der Erhöhung der Motivation sowie der Verringerung von Verständnisproblemen der Probanden und damit zur Reduzierung der Anforderungen an die Befragten. Fühlen sich Interviewte aufgrund einer schlechten Dramaturgie überfordert, weil die Fragen nicht nach Themen bzw. Gegenständen gruppiert und innerhalb der Themen nicht nach Antwortformaten sortiert präsentiert werden, besteht die Gefahr, dass diese nicht um wiederholtes Vorlesen der Fragen bitten, sondern eine willkürliche Antwort geben (ebd.).

Deshalb wurde zur Erreichung einer guten Fragebogendramaturgie für den GSTel03 thematisch zusammengehörende Fragen nach Themen zu Blöcken zusammengestellt. So entspricht beispielsweise die gemeinsame Thematisierung von Prävalenz, Diagnostik und Therapie von Rückenschmerzen eher einem thematisch zusammenhängenden Gespräch als die separate Abhandlung in getrennten Blöcken zur Therapie, Diagnostik usw., d.h. als eine analytische Blockbil-


Der Fragebogen sollte einerseits hinsichtlich Themen und Fragearten bzw. Antwortformaten klar strukturiert und nicht zu heterogen aufgebaut sein, andererseits sollte auch keine Monotonie entstehen. Da der gesamte Fragebogen des GSTel03 fast ausschließlich aus Tatsachenfragen bestand, war es um so wichtiger, das Interview abwechslungsreich zu gestalten, indem sich Bereiche bzw. Blöcke mit anspruchsvolleren Items mit leichteren abwechselten. Beispielsweise befanden sich zwischen den beiden anspruchsvollsten Blöcken zu Diabetes und Rückenschmerzen vier leichter zu beantwortende Fragen und ein kurzer Block zu Osteoporose. Die bereits thematisierten Überleitungstexte wurden ebenfalls gezielt gegen das Aufkommen von Monotonie eingesetzt. Zum Beispiel gibt es in dem langen Fragenblock zu Rückenschmerzen mehrere Überleitungen. Vor allem innerhalb umfangreicher Blöcke, wie Diabetes mellitus, wurde darauf geachtet, lange Batterien von Fragen, die mit ja oder nein beantwortbar waren, durch andere Antwortformate aufzulockern. So kamen im Diabetesblock zusätzlich standardisierte Fragen mit bis zu fünf vorgelesenen Antwortkategorien und offene Fragen (z. B. nach der Anzahl der ärztlichen Untersuchungen der Füße...

Abbildung 1
Filter für Migrationshintergrund

Mig1: Sind Sie in Deutschland geboren?

ja nein

Mig2: Haben Sie die deutsche Staatsbürgerschaft seit der Geburt?

nein ja

Mig3: In welchem Land wurden Sie geboren?

Mig4: Welche Staatsangehörigkeit(en) haben Sie jetzt?

Deutscher Staatsbürger?

ja nein

Mig5: Seit wann haben Sie die deutsche Staatsbürgerschaft?

Nach der Wende geboren?*

ja nein

Mig6: Seit wann leben Sie in Deutschland?

Nach der Wende nach Deutschland gekommen?

ja nein

Mig7: Haben Sie 1988 auf dem Gebiet der ehemaligen DDR gelebt?

Weiter im Fragebogen

Hinweis: Fragetexte nicht vollständig abgebildet
* Diese Filterbedingung wird erst ab 2006 relevant, da Minderjährige nicht zur Grundgesamtheit gehören.
2.3.2 Filterführung


Der Beitrag der Filtertechnik zur Optimierung der Operationalisierungen ist Gegenstand des nächsten Abschnittes. Dramaturgisch wirken Filter beispielsweise bei der Einspielung unterschiedlicher Überleitungstexte in Abhängigkeit von bereits beantworteten Fragen.\(^3\)


Geschickte Filterführung verhinderte z. B. bei der Erfassung des Migrationshintergrundes die Konfrontation der Probanden mit widersinnigen Fragen, wie der schematischen Darstellung in Abbildung 1 entnehmbar ist.

Für alle anderen gegenstandsorientierten Blöcke des Fragebogens gab es ebenfalls eine gezielte Anordnung von Fragen und eine stringente Filterführung. Auch wenn aufgrund von einzelnen Filterbedingungen teilweise nur wenig eingespart wurde, führte das insgesamt zu einer erheblichen Reduktion der Anzahl gesteller Fragen für viele Probanden und damit zu einer spürbaren Entlastung.

Der konsequente Einsatz der Filtertechnik stellte jederzeit und für jede Probandengruppe die korrekte Auswahl und Aneinanderreihung der zu beantwortenden Fragen sicher (Dillman 1978), ohne dass Interviewer oder Probanden damit beansprucht wurden, wie dies beispielsweise bei einem nicht computergestützten face-to-face-Interview oder insbesondere bei einer schriftlichen Befragung der Fall wäre. Durch diese intensive Nutzung der Filtertechnik entsteht allerdings ein erhöhter Aufwand bei der Programmierung des Fragebogens und der Aufbereitung der Daten für die Auswertung.

2.3.3 Optimierung der Operationalisierungen

Warum einfach fragen, wenn es auch kompliziert geht? Hier geht es vor allem um die Zerlegung von Fragen mit zu komplexen Formulierungen oder zu komplexen Antwortkategorien, die Festlegung der Reihenfolge für die Antwortvorgaben und die Vereinheitlichung retrospektiver Zeitfenster. Hinweistexte sollten Unklarheiten minimieren, um ein einheitliches und richtiges Verständnis der Fragen zu gewährleisten.

\(^{2}\) computerassistiertes persönliches Interview bzw. computergestütztes face-to-face-Interview

\(^{3}\) Nach den einführenden Fragen zur allgemeinen subjektiven Gesundheit erschienen in Abhängigkeit von deren Beantwortung zwei unterschiedliche Überleitungen. Bei eher bessarem allgemeinen Gesundheitszustand folgte: »Auch wenn das alles möglicherweise gar nicht auf Sie zutrifft, möchte ich Sie im Folgenden kurz nach einigen Krankheiten fragen.« Bei schlechterem Zustand kam hingegen: »Als nächstes möchten wir Sie nach einigen Erkrankungen fragen, die häufiger auftreten.«
2.3.3.1 Mehrdimensionale bzw. komplexe Fragen

Komplexe Frageformulierungen sind durch sprachliche Überfrachtung bzw. Mehrdimensionalität gekennzeichnet, weshalb sie von Befragten unterschiedlich verstanden werden (Friedrichs 1980). Dies betraf insbesondere Fragen, die für schriftliche Fragebögen oder persönliche Interviews entwickelt wurden bzw. für den Einsatz per CATI noch nicht ausreichend operationalisiert waren. Da ein Ziel des GSTelo3 in der Fortschreibung wichtiger Items aus vergangenen Surveys bestand, mussten diese Fragen inhaltlich äquivalent in die CATI-taugliche Form transformiert werden. Für komplexe Frageformulierungen erfolgte eine Zerlegung in einfachere Teilfragen (Dillman 1978), die so angeordnet wurden, dass die Anzahl der zu stellenden Fragen mittels Filterführung minimiert wird.


Neben einer Vereinheitlichung der Zeitfenster geht es auch um deren optimale Festlegung in Abhängigkeit von der Erinnerungsfähigkeit der

4 zum Beispiel: »Denken Sie mal an die letzten 3 Monate. Haben Sie da Sport gemacht?«
5 beispielsweise innerhalb eines Überleitungstextes, der der Frage voranging
Befragten, was am o.g. Sportbeispiel verdeutlicht werden kann: Die wöchentliche Stundenzahl sportlicher Aktivität lässt sich für die letzten drei Monate wesentlich besser und schneller memoriern und damit valider erfassen, als für die letzten 12 Monate.6

2.3.3.2 Umfangreiche bzw. komplexe Antwortkategorien

In der Literatur werden für Telefoninterviews maximal vier bis fünf Antwortkategorien empfohlen, wobei die genaue Festlegung dieser Anzahl von deren Länge abhängt (Frey, Oishi 1995). Deshalb erfolgte für Fragen mit umfangreicheren Antwortkategorien möglichst eine Zerlegung in mehrere Teilfragen (Dillman 1978). So wurde zum Beispiel die berufliche Stellung nicht in einer Frage mit sehr vielen und langen Antwortkategorien erhoben, wie bei der schriftlichen Befragung des Bundes-Gesundheitssurveys 1998, sondern zerlegt in eine Frage nach der allgemeinen beruflichen Stellung mit den Kategorien Arbeiter, Angestellter, Beamter usw., der gefiltert nach der gewählten Kategorie eine weitere Frage zum differenzierten Beschäftigungsstatus folgt. Es gibt also eine Unterfrage für die Arbeiter, eine andere für die Angestellten, eine für die Beamten etc. Abbildung 2 demonstriert diese zweistufige Abfrage für die Arbeiter.7


Ein gutes Beispiel dafür ist auch die Erfassung des Anrufergebnisses mit Hilfe der so genannten disposition codes (Kohler 2005; Borch, Rieck 2005). 48 Einzelkategorien (z. B. Freizeichen, Anrufbeantworter, Geschäftsanschluss) erforderten eine übersichtliche und strukturierte Präsentation auf dem Bildschirm, um die Interviewer nicht zu überfordern und Qualitätsprobleme bei der Berechnung der Responserate zentraler Variablen zu vermeiden (Kohler 2005). Sie wurden so auf neun Bildschirmseiten organisiert, aneinandergereiht und gefiltert, dass die Interviewer in einer konkreten Kontaktsituation nur eine bis maximal vier hierarchisch angeordnete Seiten durchgehen mussten.

7 Für die Hauptfrage wurden „sind“ oder „waren“ situationsabhängig, d. h. filtergesteuert in Abhängigkeit vom zuvor erhobenen aktuellen Erwerbsstatus, eingeblendet.

### 2.3.3.3 Schwer memorierbare Antworten


Da schwer memorierbare Antwortvorgaben – ebenso wie anspruchsvollere Frageformulierungen – eine Verlangsamung des Interviewtempo erforderlich, sind Maßnahmen zur Regulierung des Tempos, wie eben genannt, ein wichtiger Bestandteil gelungener CATI-Fragebögen, die Interviewerschulung und Qualitätssicherung entlasten und zur Erhöhung der Güte der Daten beitragen.

### 2.3.3.4 Reihenfolge von Antwortvorgaben

Bei der Optimierung der Operationalisierungen hinsichtlich der Antwortkategorien waren neben Anzahl und Umfang auch deren Reihenfolge zu berücksichtigen. In der Literatur (z. B. Schwarz et al. 1991) gibt es vor allem für Meinungsfragen die Diskussion, dass in Abhängigkeit vom eingesetzten Erhebungsverfahren eher mit einer bevorzugten Auswahl der erst- oder letztgenannten Kategorien zu rechnen ist (response order effects). Für die im GSTel03 verwendeten Fragearten stand allerdings die Festlegung einer optimalen Reihenfolge der Antwortvorgaben unter dem Aspekt ihrer Nutzbarkeit für die Interviewer und Befragten in der interaktiven Situation am Telefon im Vordergrund, um eine hohe Güte bei der Antwortauswahl hinsichtlich der intendierten Zieldimension zu erreichen (Fowler 2001).

Beispielsweise sollte die Facharztbezeichnung des den Diabetes mellitus behandelnden Arztes mittels fünf vorgelesener Antwortkategorien ermittelt werden. Dabei sollte eine Arztgruppe ausgewählt werden und zwar jene, die die Interviewten hauptsächlich in der zu erhebenden interaktiven Situation am Telefon im Vordergrund, um eine hohe Güte bei der Antwortauswahl hinsichtlich der intendierten Zieldimension zu erreichen (Fowler 2001).

Beispielsweise sollte die Facharztbezeichnung des den Diabetes mellitus behandelnden Arztes mittels fünf vorgelesener Antwortkategorien ermittelt werden. Dabei sollte eine Arztgruppe ausgewählt werden und zwar jene, die die Interviewten hauptsächlich in der zu erhebenden interaktiven Situation am Telefon im Vordergrund, um eine hohe Güte bei der Antwortauswahl hinsichtlich der intendierten Zieldimension zu erreichen (Fowler 2001).


2.3.3.5 Hinweistexte und Administrationsanweisungen


Abbildung 3
Bildschirmansicht des programmierten Fragebogens

Administrationsanweisung

vorzulesen

nicht vorzulesen
Bedeutung für die Standardisierung der Erhebungssituation zudem Kernbestandteil der Interviewerschulung (Borch, Rieck 2005).

### 2.3.4 Weitere Techniken/Maßnahmen

Neben den bisher genannten Schritten zur Optimierung des Erhebungsinstrumentes kamen für den GSTel03 eine Reihe weiterer zur Anwendung. Sie wurden aufgrund der Schwerpunktsetzung in diesem Aufsatz (insbesondere Reduktion von Komplexität) bisher nicht thematisiert. Die wichtigsten sollen abschließend kurz erwähnt werden:


- **Komfortabel programmierte Listen**, z. B. für die Erfassung des Geburtslandes, dienten als Ersatz für eine freitextliche Erfassung.


- Eine **numerische Skala mit verbalisierten Endpunkten diente zur Erfassung der Stärke der Rückenschmerzen**.\(^8\)

- Die **visuelle Gestaltung des Fragebogens unterstüzt die Interviewer** (Dillman 1978): eine einheitliche Platzierung, Formatierung, Hervorhebung und Formulierung führte dazu, dass Administrationsanweisungen, Hinweise, Fragentexte und Antwortkategorien auf den ersten Blick gefunden, verstanden und richtig vorgelesen wurden.\(^9\)

- **Anmerkungen zu den programmierten Plausibilitätskontrollen finden sich bei Borch, Rieck (2005).**


### 2.4 Fazit


---

9 Gute graphische Gestaltungsmöglichkeiten, die Einbindung langer Listen u.ä. waren zentrale Kriterien für die Auswahl der CATI-Software.
Filtertechnik bei der Zerlegung langer Antwortlisten einzusetzen. Das bedeutet: einzelne Regeln zur Fragebogenentwicklung gelangen in Abhängigkeit vom Erhebungsmodus zu unterschiedlicher Bedeutung.


2.5 Literatur

Fowler FJ (2001) Why it is easy to write bad questions, ZUMA-Nachrichten 48: 49–66
Hippler HJ, Schwarz N (1992) The impact of administration modes on response effects in surveys. ZUMA, Mannheim
Kohler M (2005) Stichprobenkonzeption und Stichprobenrealisierung (in diesem Band)
3 Feldarbeit

Susan Borch, Angelika Rieck

3.1 Einleitung

Gesundheitssurveys haben am Robert Koch-Institut eine Tradition, die bis in die 1980er Jahre zurückreicht. Während die früheren Studien aus Examinationsmodulen, selbst auszufüllenden Probanden-Fragebögen und computergestützten ärztlichen Interviews bestanden, handelt es sich beim telefonischen Gesundheitssurvey 2003 (GSTel03) um die erste Studie, die mittels computergestützter Telefoninterviews (CATI, computer-assisted telephone interviewing) durchgeführt wurde.


Außerdem werden das eingesetzte Personal sowie die Datenerhebung und -haltung beschrieben.

3.2 Personal

Das Feldteam bestand aus folgenden Personen- gruppen: Drei wissenschaftliche Mitarbeiter gehörten zur Feldleitung, die insbesondere verantwortlich war für die Optimierung des Erhebungs instrumentes unter methodischen Aspekten (Rieck et al. 2005), die Stichprobenkonzeption (Kohler 2005), die Schulung der Interviewer und Supervisoren, Qualitätssicherung, Erstellung des Operationshandbuchs, fachliche Aufsicht sowie die Organisation und Koordination der Feldarbeit.

Mit der Supervision der Interviewer wurden fünf studentische Hilfskräfte betraut. Ihre Aufgaben bestanden vor allem in der Anleitung der Interviewer während der Datenerhebung, Qualitätssicherung der Interviewführung, Berichterstattung an die Feldleitung, Unterstützung der Interviewerschulung, Rekrutierung sowie der Organisation des Interviewereinsatzes. Dabei wurden Studentinnen und Studenten sozialwissenschaftlicher Fächer mit Interviewerfahrungen verpflichtet, die auch über ein hohes Maß an sozialer Kompetenz, Flexibilität und Einsatzbereitschaft verfügten.

Als Interviewerinnen und Interviewer wurden Honorarkräfte eingesetzt, welche sich hauptsächlich aus Studierenden sozialwissenschaftlicher Fächer rekrutierten. Die Kontaktaufnahme erfolgte telefonisch, um die Eignung z. B. anhand Stimmmodulation und Sprechtempo beurteilen zu können (Friedrichs 1990). Es kam zwar vor, dass Bewerber aufgrund eines starken Dialektes abgelehnt wurden; insgesamt spielte es jedoch eine größere Rolle, welchen Eindruck sie im Bewerbungsverfahren bezüglich ihrer Zuverlässigkeit und kommunikativen Kompetenz machten.

Alle im Projekt beschäftigten Mitarbeiterinnen und Mitarbeiter wurden hinsichtlich des Datenschutzes informiert und verpflichtet.

3.3 Datenerhebung und Datenhaltung

Die Feldzeit des GSTel03 begann am 2. September 2002 und endete am 14. März 2003. In diesem Zeitraum wurden von insgesamt 81 Interviewern...
8,362 Frauen und Männer befragt, davon waren 8,318 Interviews verwertbar. Die Geschlechtsverteilung im Interviewerstamm war ausgewogen, zum größten Teil handelte es sich um Studierende zwischen 20 und 30 Jahren. Es wurden zwei Pretest durchgeführt, um zu prüfen, ob die Fragen handhabbar waren, die Interviewerhinweise ausreichten, die disposition codes\(^3\) alle Fälle abdeckten und ob das Schulungskonzept praktikabel war. Dabei wurde u. a. festgestellt, dass die Lerninhalte durch die Aufteilung in einen »Theorietag« und einen »Praxistag« zu abstrakt bleiben, was zu einem erhöhten Aufwand für Qualitätssicherung und Nachschulung führte (vgl. Borch et al. 2004). Daher wurde die Schulung für die Hauptphase neu konzipiert; außerdem wurde das Erhebungsinstrument überarbeitet, um die Handhabbarkeit für die Interviewer zu verbessern (Rieck et al. 2005).


Für die Datenerhebung wurde eine spezielle CATI-Software eingesetzt (Kohler 2005). Ein Modul steuerte z. B. die An- und Rückrufe, wobei die Datensätze zufällig unter den Interviewern verteilt wurden und der Anwahlvorgang automatisch erfolgte (auto-dialing). Ein weiteres Modul diente zur Beobachtung der Interviewerproduktivität und anderer Kenngrößen, d. h. es zeigte interviewerspezifisch die Zahl der vollständigen Interviews, vollständigen Non-Responderinterviews und Abbrüche sowie die durchschnittliche Gesprächsdauer an.


Die durchschnittliche Länge der Interviews im GSTel03 betrug zirka 21 Minuten. Diese varierte jedoch erheblich für bestimmte Befragungsgruppen: Jüngere Probanden absolvierten die Befragung in der Regel schneller als ältere, Frauen langsamer als Männer.\(^4\)


\(^3\) Die disposition codes sind ein Klassifikationsschema zur Erfassung aller vorkommenden Anwahlergebnisse (z. B. Freizeichen, Terminvereinbarung, Interview vollständig, Ansagen wie »kein Anschluss unter dieser Nummer«). In ihnen wird der Status eines Datensatzes zur Steuerung der Rückrufe festgehalten. Die verwendeten disposition codes wurden speziell für das Projekt entwickelt und sind vergleichsweise differenziert (vgl. Kohler 2005).

\(^4\) Die durchschnittliche Interviewdauer variierte nach Altersgruppen von ca. 18 Minuten (18- bis 20-Jährige) bis ca. 30 Minuten (über 80-Jährige) und nach Geschlecht zwischen ca. 20 Minuten (Männer) und ca. 22 Minuten (Frauen).
Bei Vorhandensein bestimmter Krankheiten stieg die durchschnittliche Interviewdauer, da in diesem Fall filterbedingt mehr Fragen gestellt wurden (Rieck et al. 2005). Zur datenschutzgerechten Umsetzung der Studie gehörte einerseits die entsprechende Handhabung der Daten: Die Interviewsoftware war lokal auf den CATI-Arbeitsplätzen installiert, die Datenspeicherung erfolgte zentral und passwortgeschützt auf einem auch physisch gegen unbefugte Zugriffe gesicherten Server. Der Zugang zum CATI-System wurde beschränkt und dokumentiert. Die Verwaltung der Daten erfolgte ausschließlich anonym über die fortlaufende Datensatznummer. Außerdem wurden die Befragten vor Beginn des Interviews sowie vor der Abfrage soziodemografischer Daten darüber informiert, dass ihre Angaben freiwillig sind, nur anonymisiert ausgewertet werden und sie auch die Möglichkeit haben, einzelne Fragen nicht zu beantworten.

3.4 Mitarbeiterschulung

Die Schulung von Interviewern und Supervisoren hat eine zentrale Stellung innerhalb der Feldarbeit: Die Interviewer sollen bei der Kontaktaufnahme und bei der Befragung flexibel und aufmerksam auf den Gesprächspartner eingehen sowie souverän und seriös wirken, um die Motivation zu erhalten und Abbrüche zu vermeiden. Gleichzeitig sollen sie sich an klar definierte Regeln halten, wie z. B. die vorgegebenen Frage korrekt vorzulesen und ausschließlich erlaubte Hinweise zu geben. Die Interviewer haben also in der Erhebungs situation zwei gegensätzliche Rollen auszufüllen (Friedrichs 1990), auf die sie in der Schulung möglichst umfassend und realitätsnah vorbereitet werden müssen.

Die Supervisoren haben aufgrund ihrer Handlungsanweisungen an die Interviewer eine wichtige Funktion bei der Umsetzung der standardisierten Erhebungs situation. Zu ihren Aufgaben gehört die ständige Betreuung und Nachschulung der Interviewer im Rahmen der Qualitätssicherung. Außerdem berichten sie der Feldleitung über den Prozess. Damit die Supervisoren die methodische Relevanz von Ereignissen angemessen einschätzen können, sind kontinuierliche Schulung und Betreuung erforderlich.

3.4.1 Interviewer


6 In wenigen Fällen wurde die Erlaubnis zur Datenspeicherung von Probanden nachträglich widerrufen. Ihre Angaben wurden gelöscht.
Ziel der Schulung war die Vermittlung von Qualitätsstandards; die Lerninhalte lassen sich in folgenden Punkten zusammenfassen:

- Hintergrund der Studie, Auftraggeber und Zielsetzung,
- Aufbau, Inhalt und Besonderheiten des Erhebungsinstrumentes,
- Sicherer technischer Umgang mit der CATI-Software,
- Vorgehen in der Kontaktphase des Interviews incl. angemessenes Auftreten, Argumentationsstrategien, korrekter Durchführung der Zielpersonenermittlung und zweckmäßigen Umgang mit den disposition codes,
- Vollständige, aussagekräftige und datenschutzgerechte Dokumentation der Zielpersonenermittlung,\footnote{Sie ist notwendig, wenn das Interview zu einem anderen Zeitpunkt geführt werden soll.}
- Angemessenes Auftreten gegenüber den Befragten im weiteren Verlauf des Interviews,
- Angemessener Umgang mit Antworten wie »weiß nicht« und »keine Angabe«,\footnote{Die Antwortvorgaben »weiß nicht« und »keine Angabe« wurden von den Interviewern nicht vorgelesen, eine von den Probanden entsprechend geäußerte Antwort jedoch akzeptiert. Waren sich die Befragten unschlüssig darüber, ob sie eine Frage beantworten wollen, wiesen die Interviewer einerseits auf die Freiwilligkeit der Angaben und andererseits auf die anonymisierte Auswertung der Daten hin.}
- Vollständige, aussagekräftige und datenschutzgerechte Dokumentation von Informationen über mangelnde Qualität der Antworten beziehungsweise einzelner Fragebogenitems, kompletter Interviews oder nachträglich festgestellter Fehler, wie die Nicht-Zugehörigkeit zur Grundgesamtheit,
- Konstanthalten der Erhebungssituation.\footnote{Die zum Beispiel bei Prüfer und Stiegler (2002) dargestellten allgemeinen Regeln zur Standardisierung der Erhebungssituation wurden für den GSTel03 adaptiert und präzisiert.}

Für die Qualität der erhobenen Daten ist der letzte Punkt von essentieller Bedeutung; dieser wurde daher besonders vertiefend behandelt. Zum Konstanthalten der Erhebungssituation zählen so grundlegende Dinge, wie die Interviews ausschließlich in deutscher Sprache zu führen, die persönlichen Einstellungen gegenüber den Befragten zu verbergen sowie auf eigenmächtige Interpretationen und Erläuterungen von Fragen zu verzichten. Im programmierten Fragebogen waren daher präzise Interviewerhinweise dazu enthalten, welche Erklärungen bei Nachfragen der Probanden gegeben werden dürfen (Rieck et al. 2005).


Zur Optimierung der Ausschöpfung (Kohler 2005) wurden der korrekte Umgang mit den umfangreichen disposition codes und das Vorgehen in der Kontaktphase mit der Zielpersonenermittlung ausführlich besprochen und trainiert. Gegenstand war ebenfalls die Vermittlung eines im Pretest erprobten Vorgehens in der Kontakt-

Als Schulungsmethoden kamen Lehrgespräche, Gruppendiskussionen, praktische Arbeit mit dem Interviewprogramm sowie selbständig zu bearbeitende Übungsaufgaben zum Einsatz. Im Rollenspiel wurden unter Einbeziehung aller Teilnehmerinnen und Teilnehmer typische Fragen und Probleme simuliert und analysiert. Bei Probeinterviews befugten sich die angehenden Interviewer unter Echtbedingungen zunächst gegenseitig, um im zweiten Schritt dann Freunde und Bekannte anzurufen, um das Erlernte nochmals anzuwenden.

Teilweise hatten die Bewerber Vorervahrungen, die bei der Gestaltung der Schulung berücksichtigt werden mussten. Diese reichten von qualitativer Interviewführung bis zu CATI im Rahmen von Markt- und Meinungsforschung. Erfahrene Schulungsteilnehmer verfügten mehrheitlich über sehr gut ausgeprägte kommunikative Fähigkeiten sowie Souveränität bei der Interviewanbahnung, die unerfahrene erst erwerben mussten. Andererseits handhaben sie beispielsweise disposition codes und Verweigerungen teilweise anders als für diese Studie gefordert. Es wurde daher darauf geachtet, die Trainingsbedürfnisse aller Teilnehmergruppen abzudecken.

### 3.4.2 Supervisoren


Als Schulungsmethoden sind vor allem die persönliche Anleitung durch die Feldleitung als training on the job, die Durchführung von Interviews zu Test- und Schulungszwecken, Gruppendiskussionen im Rahmen der wöchentlichen Teamsitzungen sowie die Lektüre des Operationshandbuchs und anderer Literatur zu nennen.

### 3.5 Qualitätssicherung

### Tabelle 1
Checkliste zur Qualitätssicherung der Interviewführung

<table>
<thead>
<tr>
<th>Datum:</th>
<th>Interviewer/in:</th>
<th>Supervisor/in:</th>
</tr>
</thead>
</table>

#### Interviewtechniken

1. **Wählt die Zielperson (ZP) korrekt aus**
   - 1
   - 2
   - 3
   - 4
   - 5

2. **Führt das Auswahlverfahren der ZP reibungslos aus**
   - 1
   - 2
   - 3
   - 4
   - 5

3. **Hält sich wörtlich an die vorgegebenen Texte**
   - 1
   - 2
   - 3
   - 4
   - 5

4. **liest die Fragen vollständig vor**
   - 1
   - 2
   - 3
   - 4
   - 5

5. **liest die Antwortkategorien vollständig vor**
   - 1
   - 2
   - 3
   - 4
   - 5

6. **Führt das Interview in einem guten Tempo**
   - 1
   - 2
   - 3
   - 4
   - 5

7. **Interview ist fließend**
   - 1
   - 2
   - 3
   - 4
   - 5

8. **Lässt die ZP nicht abschweifen**
   - 1
   - 2
   - 3
   - 4
   - 5

9. **Motiviert die ZP weiterzumachen**
   - 1
   - 2
   - 3
   - 4
   - 5

10. **Erklärt situationsbedingt**
    - 1
    - 2
    - 3
    - 4
    - 5

11. **Hält sich bei Erläuterungen an die vorgegebenen Hinweise**
    - 1
    - 2
    - 3
    - 4
    - 5

12. **Formgerechte Verabschiedung**
    - 1
    - 2
    - 3
    - 4
    - 5

**Freitextliche Erläuterung zur Bewertung der Interviewtechniken:**

#### Nachfragen

1. **Frage situationsbedingt nach**
   - 1
   - 2
   - 3
   - 4
   - 5

2. **Stellt keine Suggestivfragen**
   - 1
   - 2
   - 3
   - 4
   - 5

3. **Besteht auf exakte Antwortkategorien**
   - 1
   - 2
   - 3
   - 4
   - 5

4. **Antwortet auf Nachfragen inhaltlich richtig**
   - 1
   - 2
   - 3
   - 4
   - 5

5. **Drückt sich auf Nachfragen hin verständlich aus**
   - 1
   - 2
   - 3
   - 4
   - 5

**Freitextliche Erläuterung zur Bewertung des Nachfrageverhaltens:**

#### Verhalten

1. **Ist höflich und zuvorkommend**
   - 1
   - 2
   - 3
   - 4
   - 5

2. **Klingt sicher**
   - 1
   - 2
   - 3
   - 4
   - 5

3. **Klingt nicht gelangweilt**
   - 1
   - 2
   - 3
   - 4
   - 5

4. **Spricht Fremdwörter korrekt aus**
   - 1
   - 2
   - 3
   - 4
   - 5

5. **Hat eine deutliche Aussprache**
   - 1
   - 2
   - 3
   - 4
   - 5

6. **Treibt den Probanden nicht zur Eile an**
   - 1
   - 2
   - 3
   - 4
   - 5

7. **Klingt nicht sarkastisch**
   - 1
   - 2
   - 3
   - 4
   - 5

8. **Zeigt Interesse gegenüber ZP**
   - 1
   - 2
   - 3
   - 4
   - 5

9. **Lässt sich von ZP stilistisch nicht beeinflussen**
    - 1
    - 2
    - 3
    - 4
    - 5

**Freitextliche Erläuterung zur Bewertung des Verhaltens:**

#### Traten Probleme auf?

**Wie ist der Gesamteindruck?**

Würdest Du gerne von diesem Interviewer befragt werden?  [ ] Ja  [ ] Nein

3.5.1 Interviewführung und Schulung

Qualität bezüglich der Interviewführung wurde als kontinuierliche Einhaltung der in der Schulung vermittelten Qualitätsstandards und als Befolgung der Datenschutzbestimmungen definiert.

Die Strategien zur Ermittlung von Verbesserungsbedarf bei der Interviewdurchführung fußten auf verschiedenen Säulen. So wurde Wert auf eine konstruktive und kooperative Atmosphäre gelegt, in der die Optimierung der Datenerhebung im Vordergrund stand,** denn die Betreuung durch die Supervisoren sollte als Hilfestellung bei Problemen und nicht als Überwachung empfunden werden.** In diesem Arbeitsklima machten die Interviewer häufig von der Möglichkeit Gebrauch, bei Unklarheiten im Umgang mit komplexen Situationen nachzufragen.


Neben diesen im Ablauf festgelegten Maßnahmen zur Qualitätssicherung (vgl. Abbildung 1) haben die Supervisoren natürlich auch während anderer Tätigkeiten, wie der Dokumentation, ständig mit »halbem Ohr« die Interviews verfolgt, was quantitativ die größte Rolle für die Aufdeckung von Nachschulungsbedarf spielte. Diese Nachschulungen bestanden aus einem kurzen lobenden oder kritisierenden, aber stets motivierenden Feedback zwischen zwei Interviews. Diese wurden

** Daher erfolgte die Vergütung der Interviewer im GSTel03 auch durch eine Mischkalkulation von festem Stundenlohn für das Durchführen von Anwahlversuchen und einem Leistungsanteil für vollständige Interviews und Nonresponder-Interviews. Durch den Leistungsanteil sollte eine Anreiz zum effizienten Arbeiten gesetzt werden, andererseits sorgte das Fixum dafür, dass die Interviewer qualitäts- und nicht quantitätsorientiert arbeiteten. Die Honorierung der Anwahlversuche ist notwendig, um eine gute Ausschöpfung der zufallsgenerierten Telefonnummern zu erreichen. Da für mussten die Interviewer motiviert werden, auch wenig erfolgversprechende Telefonnummern mehrfach (entsprechend den Regeln für das Anruf- und Rückrufmanagement (Kohler 2005)) anzurufen.


Damit sind die Qualitätssicherung der Interviewführung und der Schulung in einem gemeinsamen Prozess eng miteinander verknüpft (vgl. Abbildung 1).


3.5.2 Supervision

Qualität in der Supervision wurde wie folgt definiert:

- Regelmäßige und sachgemäße Qualitätssicherung bezüglich der Interviewführung,
- Ermittlung einheitlicher und sachgemäßer Anweisungen, vor allem während der Nachschnulung,
- Motivierender Kommunikationsstil gegenüber den Interviewern, insbesondere im Rahmen der Nachschnulung,
- Vollständige Durchführung der täglich vorzunehmenden technischen Vorbereitungen und Tests,
- Vollständige und aussagekräftige Dokumentation technischer Probleme, inhaltlicher Unklarheiten und des Nachschnulungsbedarfs von Interviewern,
- Sensibilität gegenüber den (insbesondere methodischen) Konsequenzen von Ereignissen und Vorgehensweisen,
- Einhaltung des Datenschutzes bezüglich der Befragten- sowie der Interviewerdaten.


CATI ist insofern keine einfach anwendbare, quasi automatisierbare Erhebungsroutine, sondern ein sensibles Instrument, dessen Vorteile nur durch erhebliche Fachkompetenz, Qualitätsbewusstsein und viel Engagement der Beteiligten ausgeschöpft werden können.

### 3.7 Literatur

- Kohler M (2005) Stichprobenkonzeption und Stichprobenrealisierung (in diesem Band)
4 Stichprobenkonzeption und Stichprobenrealisierung

Martin Kohler

4.1 Einleitung


4.2 Technische Rahmenbedingungen

Die Datenerhebung fand unter Einsatz von gesondert geschulten Interviewern im eigens am Robert Koch-Institut eingerichteten CATI-Labor direkt an den 19 CATI-Arbeitsplätzen statt. Grundlage hierzu bildete ein programmiertes Fragebogen, der vergleichbar mit computer-gestützten Interviews (CAPI) sowohl eine Datenkontrolle während der Befragung als auch die Ablaufsteuerung mittels programmiertem Filter übernimmt.


Das Interviewer-Suite Programm übernimmt auch die Anrufverwaltung und die Steuerung der »call-backs«, d. h. der Anrufe, bei denen noch kein abschließender Kontakt mit einer Zielperson im Haushalt stattfinden konnte. Die Anwahl wird dabei von einem »auto-dialing«-System übernommen, d. h. der Anwahlvorgang der eingespielten Rufnummern wird automatisch durch das Programm durchgeführt. Fehlanrufe bei Wahlwiederholung sind nahezu ausgeschlossen und eine eindeutige, nachvollziehbare Stichprobenrealisierung ist gewährleistet.

1 ausführlicher hierzu Borch, Rieck (2005) in diesem Band, CATI = computer-assisted telephone interviewing
2 CAPI = computer-assisted personal interviewing; zu den Abkürzungen vgl. auch Couper et al. (1998)
3 zur Fragebogenerstellung siehe Rieck et al. 2005, in diesem Band
4.3 Stichprobendesign


Aus dieser hypothetischen Grundgesamtheit wird dann uneingeschränkt zufällig die Anzahl der erforderlichen Telefonnummern gezogen, die sich aus der Sollzahl der vorgegebenen Interviews multipliziert mit einem Unsicherheitsfaktor ergibt. Der Unsicherheitsfaktor ist notwendig, weil sich die Existenz eines Anschlusses hinter den generierten Nummer erst im Laufe der Feldarbeit ergibt.


Die Tabelle 1 zeigt das im Rahmen des GSTel03 verwendete und von ZUMA-Mannheim zur Verfügung gestellte Sample der zufallsgenerierten Rufnummern, die im Rahmen des GSTel03 benutzt wurden (n=45.819) sowie die Eintragdichte.

4 Für das Jahr 2001 besteht der Gesamtrahmen aus 71,7 Millionen Nummern, von denen 30,7 Millionen tatsächlich eingetragen sind. (Heckel 2002)
5 Dieses so genannte „Randomized Last Digits“-Verfahren ist gegenüber einer reinen Zufallsziffernanwahl, d. h. der Generierung der gesamten Telefonnummern („Random Digit Dialing“), für Deutschland aufgrund der heterogenen Rufnummernstruktur effizienter, da zwischen den Blöcken häufig große Lücken existieren. Es wird hierbei unterstellt, dass der Anteil von nicht eingetragenen Rufnummern, die einen eigenen Block bilden, vernachlässigbar gering ist. (Heckel (2001); Häder, Gabler (2002))
6 Nach Expertenmeinung ist diese Methode für die Befragten leichter zu handhaben als die last-birthday-Methode.
7 Zentrum für Methoden und Analysen, Mannheim (ZUMA) http://www.gesis.org
8 Unter dem Begriff »Eintragdichte« versteht man den Anteil der Personen/Haushalte, der sich in ein öffentliches Telefonbuch hat eingetragen lassen.
Tabelle 1 zeigt, dass bei über einem Fünftel der nicht eingetragenen Telefonnummern ein Privathaushalt anzutreffen ist. Der zusätzliche Aufwand, dabei auch einen hohen Anteil ungültiger Nummern anzuwählen, ist damit durchaus berechtigt.

<table>
<thead>
<tr>
<th>Personen-Kontakt unter der angewählten Nummer</th>
<th>Eintrag in Telefonbuch</th>
<th>Gesamt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nein</td>
<td>Ja</td>
</tr>
<tr>
<td>Status unklar</td>
<td>0%</td>
<td>0,1%</td>
</tr>
<tr>
<td>Nein</td>
<td>77,3%</td>
<td>24,7%</td>
</tr>
<tr>
<td>Ja</td>
<td>22,7%</td>
<td>75,2%</td>
</tr>
<tr>
<td>Gesamt</td>
<td>28.444</td>
<td>17.375</td>
</tr>
<tr>
<td></td>
<td>62,1%</td>
<td>37,9%</td>
</tr>
</tbody>
</table>

Tabelle 1 zeigt, dass bei über einem Fünftel der nicht eingetragenen Telefonnummern ein Privathaushalt anzutreffen ist. Der zusätzliche Aufwand, dabei auch einen hohen Anteil ungültiger Nummern anzuwählen, ist damit durchaus berechtigt.

4.4 Anrufmanagement und Stichprobenbearbeitung

Das für den GSTel03 gewählte stichprobentheoretische Konzept bestimmt auch maßgeblich das Vorgehen während der Feldarbeit. Im Unterschied zu Ansätzen, bei denen man im Voraus bekannte Personen versucht, telefonisch zu kontaktieren – z. B. basierend auf einer Stichprobe von Einwohnermeldeamtsdaten oder durch vorhergehende Screening-Interviews –, liegt im GSTel03 als Ergebnis der Stichprobenziehung eine Menge von zu kontaktierender Rufnummern vor, bei denen zunächst nicht bekannt ist, welche tatsächlich existieren und welche davon Anschlüsse von Privathaushalten sind.

Nach dieser Klärung erfolgte dann beim ersten Kontakt mit der zufällig angetroffenen Person (Kontaktperson oder kurz: KP) die Ermittlung der Zielperson im Haushalt (kurz: ZP) nach der next-birthday-Methode. Nur mit dieser Person durfte sofort ein Interview durchgeführt werden oder es war ein geeigneter Rückruftermin zu vereinbaren, an dem die ermittelte Zielperson vor- aussichtlich zu erreichen ist.

Selbstverständlich gelingt die Klärung, ob es sich um einen Privathaushalt handelt, die Zielpersonermittlung oder gar die Durchführung

9 Zum Begriff des Privathaushalts vgl. auch Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM), Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) (1999), Statistisches Bundesamt (2004); sowie die Anweisungen im OP-Handbuch zum GSTel03, Borch et al. (2004). Ein Interview durfte nur dann unter einem Firmenanschluss durchgeführt werden, wenn unter derselben Telefonnummer auch ein Privathaushalt zu erreichen ist (z. B. bei Selbstständigen mit wenigen oder keinen Angestellten).
Abbildung 1: Anrufsteuerung im GSTelo3

1. Anwahlversuch
   - Technische Variablen
     - Verbindung?
       - ja
         - KP holt ZP
       - nein
         - Privatanschluss?
           - nein
             - raus
           - ja
             - KP wimmelt ab
     - KP # ZP
       - KP macht Angaben zu ZP
       - raus
       - CB & Lines
   - Lines

2. Anwahlversuch
   - Lines

TP-Ermittlung

KP = ZP
   - ZP kann bzw. will gerade nicht
     - CB & Lines
   - ZP stimmt zu
     - ZP stimmt zu
     - Abbruch
     - raus
   - KFB
     - KFB vollständig
     - CB & Lines
     - Lines

KP # ZP
   - KP holt ZP
   - CB & Lines

ZP verweigert

CB – call-back
KP – Kontaktperson
ZP – Zielperson
KFB – Kurzfragebogen
lines – Kommentare
des Interviews mit der Zielperson nur selten beim ersten Anwahlversuch einer Telefonnummer. Dies kann häufig erst durch mehrmali
ge Kontaktver-
suche festgestellt werden. Hier gilt es einen Mittel-
weg zu finden zwischen Effizienz der Anwahlver-
suche und den geforderten Qualitätsansprüchen. Wie an anderer Stelle dargelegt wurde, ist davon auszugehen, dass sich die schwer unter ihrer Fest-
etznummer erreichbaren Personen systematisch von den leichter zu Erreichenden unterscheiden (Häder 1996).

Die Verwaltung der Rufnummern mit dem je-
weils von den Interviewern zu protokollierenden Anwahlergebnis, in dem gegebenenfalls auch Kommentare zur Zielpersonermittlungen enthal-
ten sind, übernimmt dabei das Anruf- und Rück-
rufmanagement, das den jeweiligen Anforderun-
gen entsprechend konzipiert und programmiert werden muss.

4.4.1 Das Anruf- und Rückrufmanagement

Die Konzeption eines Anruf- und Rückrufma-
ngements bedarf einiger Vorbüberlegungen und kann aus unterschiedlichen Perspektiven betrachtet werden. Aus Sicht der Interviewer sollte es möglichst allen Eventualitäten gerecht werden, sodass eine eindeutige Zuordnung des jeweiligen Anwahlergebnisses vorgenommen werden kann. Gleichzeitig darf es aber auch nicht zu unüber-
sichtlich sein, da die Interviewer in der Kontakt-
phase besonderen Belastungen ausgesetzt sind und deren Aufmerksamkeit durch den jeweiligen Gesprächspartner in Anspruch genommen wird. Andererseits erfolgt über die Kodierung des Anwahlergebnisses (disposition codes) auch die Steuerung der Wiedervorlage der entsprechenden Rufnummer nach festzulegenden Regeln (call-
back rules) und berührt damit die Qualität der Stichprobausschöpfung. Weiterhin wird über die Vergabe der disposition codes auch die Ergeb-
nisdarstellung der Ausschöpfung und damit auch die Berechnung der Ausschöpfungsquote deter-
miniert. Von zentraler Bedeutung hierbei ist die Unterscheidung, ob sich Anrufergebnisse den Kategorien »neutraler« bzw. »nicht neutraler« Aus-
fälle zuordnen lassen. Dabei sind wiederum eige-
ne Gesichtspunkte zu beachten, auf die weiter unten noch näher eingegangen wird. Die Ent-
wicklung eines Anrufmanagementsystems stellt damit einen Optimierungsprozess teilweise diver-
gierender Zielvorgaben dar und kann nicht allen Bedürfnissen gerecht werden.

10 So wird beispielsweise differenziert, ob eine Verweige-
rung durch die – zufällig – erreichte Kontaktperson er-
folgt, bevor die Zielperson im Haushalt ermittelt werden konnte. Dieses Ergebnis bekommt einen eigenen Code und wird über die Wiederanrufregeln zeitversetzt erneut angerufen. Konnte dagegen im jeweiligen Haushalt be-
reits die Zielperson ermittelt werden und diese verwei-
gert, muss eine anderer Code vergeben werden und die Telefonnummer wird nicht mehr angewählt.

11 Wenn beispielsweise bereits mehrere Kontaktversuche bei einer Telefonnummer durchgehend das Ergebnis »Freizeichen« erzielt haben, so ist trotzdem nicht ein-
deutig zu sagen, ob der Anschluss vergeben ist und wenn ja, ob es sich nicht um einen Firmenanschluss handelt. Nur durch den mehrmaligen Wiederanruf zu verschiedenen Uhrzeiten und Wochentagen kann das Risiko minimiert werden, eine Zielperson nicht zu er-
reichen.
Anders als in den USA liegen im deutschsprachigen Raum bisher keine einheitlichen Standards für die Entwicklung von disposition codes vor. Die rules und disposition codes mussten daher speziell für den GSTelo3 entwickelt und auf Basis der Erfahrungen des Pretests verfeinert werden. Die Komplexität des Prozesses wird dabei aus dem Ablaufschema für die call-back-Verwaltung in der Abbildung 1 ersichtlich.

Prinzipiell sind für das Anruf- und Rückrufmanagement zwei Arten von disposition codes zu unterscheiden:

- Erstens Codes, die Telefonnummern zugeordnet werden, die nicht mehr kontaktiert werden sollen (z. B. nach vollständig abgeschlossenem Interview oder wenn die Zielperson aus gesundheitlichen Gründen nicht in der Lage ist teilzunehmen), in der Diktion der AAPOR als final codes bezeichnet.
- Zweitens Ergebnisse, bei denen mindestens ein weiterer Kontaktversuch stattfinden soll, also temporäre Codes, die die so genannten call-backs definieren. Call-backs können mit dem Gesprächspartner vereinbarte und vom Interviewer als Gesprächsergebnis eingegebene Rückruftermine sein, werden aber auch automatisch durch programmierte Algorithmen (call-back rules) generiert, z. B. wenn bei Besetzt- und Freizeichen kein Kontakt stattgefunden hat oder aus technischen Gründen keine Verbindung hergestellt werden konnte.

Die AAPOR unterscheidet darüber hinaus noch eine dritte Katagorie, die so genannten »action codes«, unter denen alle Ereignisse subsumiert werden, die manuelle Eingriffe erfordern und nicht über das call-back Management automatisiert verwaltet werden, z. B. das Zuweisen von Anrufen an bestimmte Interviewer oder manuelles Entfernen eines Datensatzes aus dem call-back-Pool.

Jedes Anrufergebnis wird während der Feldarbeit durch die Interviewer dokumentiert und mit einem entsprechenden Code versehen. Diesen Ergebniscodes werden dann Regeln zugeordnet, nach denen die jeweilige Rufnummer zur Wiedervorlage kommt. Diese »call-back rules« bestimmen u. a., zu welchem Zeitpunkt (nach zehn Minuten, einem Tag, mehreren Tagen etc.) eine mit call-back-Status versehene Telefonnummer wieder in den anzurufenden Pool von Datensätzen eingespielt wird.


Außerdem wird im call-backs-Modul die Priorität festgelegt mit der »frische« Datensätze – d. h. Telefonnummern, die bisher noch nicht angerufen wurden – behandelt werden. Dies ist vor allem bei einem bereits stark ausgeschöpften Sample relevant: sind die Nummern im bestehenden call-back-Pool bereits mehrmals angerufen worden, sinkt die Wahrscheinlichkeit eines Kontaktes rapide. Um die Effizienz der Datenerhebung und auch die

12 Für die USA gibt es Richtlinien der American Association of Political Opinion Research (AAPOR 2004). Die in etwa vergleichbaren Einrichtungen der ADM-Institut in Deutschland oder ZUMA in Mannheim haben keine derartigen Empfehlungen. Die umfangreiche Literatur über Telefonsurveys geht auf diesen Aspekt in der Regel nicht ein.

13 So versteht sich von selbst, dass mit Zielpersonen vereinbarte Termine mit höherer Priorität behandelt werden müssen, als beispielsweise call-backs, die wegen Besetzt- oder Freizeichen zum gleichen Zeitpunkt zur Wiederanwahl vom System eingespielt werden.


Bei Anrufergebnissen, die gegen einen vergebenen Anschluss sprechen (wie z. B. »Leitung tot« oder andere Display-Anzeigen), wird in deutlichem zeitlichem Abstand ein zweiter Kontaktversuch unternommen, um das Erstergebnis zu überprüfen.

4.4.2 Prozesssteuerung im Call-back Management

Nachfolgend wird das Anrufmanagement im GSTel03 ausführlicher behandelt. Dieser Abschnitt richtet sich vorwiegend an technisch interessierte Leser, die selbst ein Anrufmanagement zu betreuen haben oder dies beabsichtigen. Er wird an dieser Stelle behandelt, weil die Dokumentation dieses Aspekts telefonischer Umfragen üblicherweise in Publikationen zu kurz kommt, aber für die Transparenz des Vorgehens im Rahmen von telefonischen Befragungen bedeutsam ist.


Abbildung 2
Baumdarstellung Anrufsteuerung bei der Kontaktabhänung

01) OK-Verbindung

09) NOTAUSGANG = nächstes Interview
- END (1.563)

10) Freizeichen, niemand hebt ab (6x klingeln)
- END (44.960)
  7x Yes 30 Days; 9x No

30) Leitung tot
- END (428)
  2x No

31) Besetzt
- END (4.143)
  3x Yes 1 Day; 5x Yes 18 Days; 9x No

11) Besetzt bzw. kein Freizeichen – Display

12) Ansage (kein AB od. Warteschleife)

13) AB (privat od. geschäftl.)

14) Fax/Modem
- END (14.215)
  3x Yes 8 Days; 5x Yes 18 Days; 7x Yes 30 Days; 9x No

Tel. Nr. importiert

INT 10 1. Kontakt

DIAL

CATI
Kontaktaufnahme

Display-Anzeige

INT 30

INT 40 Ansagetexte

INT 20

INT 30

INT 10

41) Nr. nicht vergeben
- END (428)
  2x No

42) Teiln. hat neue Nr.; wird genannt
Tel. Nr. eingeben
- END (17)

43) Teiln. hat neue Nr.; wird im Laufe des Tages bekanntgegeben
- END (35)
  2x No

44) Anschluss vorübergehend nicht erreichbar
- END (211)
  1x No

45) Kein Anschluss unter dieser Nr.
- END (23)
  2x No

46) Sonstige
Supervisor informieren!
- END (62)
  2x No

21) AB privat
- END (11.115)
  3x Yes 4 Days; 4x Yes 9 Days; 5x Yes 18 Days; 9x No

22) AB nicht eindeutig zu klären
- END (6.179)
  3x Yes 18 Days; 5x Yes 30 Days; 9x No

23) Nr. nicht vergeben
- END (428)
  2x No

24) Teiln. hat neue Nr.; wird genannt
Tel. Nr. eingeben
- END (17)

25) Teiln. hat neue Nr.; wird im Laufe des Tages bekanntgegeben
- END (35)
  2x No

26) Anschluss vorübergehend nicht erreichbar
- END (211)
  1x No

27) Kein Anschluss unter dieser Nr.
- END (23)
  2x No

28) Sonstige
Supervisor informieren!
- END (62)
  2x No

29) Nr. nicht vergeben
- END (428)
  2x No

30) Teiln. hat neue Nr.; wird genannt
Tel. Nr. eingeben
- END (17)

31) Teiln. hat neue Nr.; wird im Laufe des Tages bekanntgegeben
- END (35)
  2x No

32) Anschluss vorübergehend nicht erreichbar
- END (211)
  1x No

33) Kein Anschluss unter dieser Nr.
- END (23)
  2x No

34) Sonstige
Supervisor informieren!
- END (62)
  2x No

35) Nr. nicht vergeben
- END (428)
  2x No

36) Teiln. hat neue Nr.; wird genannt
Tel. Nr. eingeben
- END (17)

37) Teiln. hat neue Nr.; wird im Laufe des Tages bekanntgegeben
- END (35)
  2x No

38) Anschluss vorübergehend nicht erreichbar
- END (211)
  1x No

39) Kein Anschluss unter dieser Nr.
- END (23)
  2x No

40) Sonstige
Supervisor informieren!
- END (62)
  2x No

41) Nr. nicht vergeben
- END (428)
  2x No

42) Teiln. hat neue Nr.; wird genannt
Tel. Nr. eingeben
- END (17)

43) Teiln. hat neue Nr.; wird im Laufe des Tages bekanntgegeben
- END (35)
  2x No

44) Anschluss vorübergehend nicht erreichbar
- END (211)
  1x No

45) Kein Anschluss unter dieser Nr.
- END (23)
  2x No

46) Sonstige
Supervisor informieren!
- END (62)
  2x No

- Weiter mit Abbildung 3

DS – Datenschutzerklärung
INTxx – neue Bildschirmseite
KFB – Kurzfragebogen
KP – Kontaktperson
YN01 – Kommentar ja/nein
ZP – Zielperson
- – Rückrufregeln, näheres siehe Text
Abbildung 3
Baumdarstellung Anrufsteuerung bei der Kontaktanbahnung

02) KP = ZP + bereit /DS Interview
53) KP wimmelt ab, lässt nicht ausreden (ZP unklar)

YN01(5 252) 2 x No

04) ZP ans Tel. geholt & bereit /DS Interview
61) ZP kurzfr. krank /Termin-Kalender (265)
62) ZP nicht da/keine Zeit /Termin-Kalender (13.098)

63) ZP versteht zu wenig deutsch YN01 (248)
64) ZP kann nicht antworten (Krankheit, Alter) YN01 (351)
65) ZP langfr. abwesend YN01 (329)
66) KP verweigert angebl.
stellvertretend f. ZP YN01 (347) 2 x No

58) keine ZP im Haushalt (z.B. < Altersgrenze) END (1.291)
59) Kein Privat-Haushalt END
83) Geschäftsanschluss END (1.054)

DS – Datenschutzerklärung
INTx = neue Bildschirmseite
KFB – Kurzfragebogen
KP = Kontakt-Person
YN01 = Kommentar ja/nein
ZP = Zielperson
* = Rückrufregeln, näheres siehe Text
Die Abfolge der Bildschirmseiten wird nachfolgend exemplarisch für das Anwahlereignis 62 »Zielperson hat gerade keine Zeit« und anschließender Terminvereinbarung dargestellt.


Die Abbildung 4 zeigt den Bildschirm nach dem Anwahlvorgang. Es kommt eine Verbindung zustande.

Abbildung 4
Bildschirm Anwahlvorgang

Abbildung 5
Bildschirm Zielpersonenermittlung
Darauf hin folgt die Ausdifferenzierung der Gründe der aktuellen Nichtteilnahme(-bereitschaft) auf einer weiteren Bildschirmseite. Hier werden die Gründe der Nichtteilnahme der Zielperson ausdifferenziert (Abbildung 6).

Da die Zielperson gerade keine Zeit hat, aber zu einem späteren Zeitpunkt bereit wäre, ein Interview zu führen, wird ein Termin zu einem späteren Zeitpunkt mit ihr vereinbart (Abbildung 7).
<table>
<thead>
<tr>
<th>Anwahlergebnis</th>
<th>Häufigkeit</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>10) Freizeichen, niemand hebt ab (6 x Klingeln)</td>
<td>44.960</td>
<td>27,6</td>
</tr>
<tr>
<td>32) Falsche Nummer</td>
<td>29.020</td>
<td>17,8</td>
</tr>
<tr>
<td>14) FAX/Modem</td>
<td>14.215</td>
<td>8,7</td>
</tr>
<tr>
<td>62) ZP nicht da/ hat gerade keine Zeit</td>
<td>13.098</td>
<td>8,0</td>
</tr>
<tr>
<td>21) AB privat</td>
<td>11.115</td>
<td>6,8</td>
</tr>
<tr>
<td>97) Hauptfragebogen komplett</td>
<td>8.351</td>
<td>5,1</td>
</tr>
<tr>
<td>22) AB nicht eindeutig zu klären</td>
<td>6.179</td>
<td>3,8</td>
</tr>
<tr>
<td>99) endgültiger Abbruch</td>
<td>5.528</td>
<td>3,4</td>
</tr>
<tr>
<td>53) KP wimmelt ab, lässt nicht ausreden (ZP noch nicht bekannt)</td>
<td>5.252</td>
<td>3,2</td>
</tr>
<tr>
<td>37) keine Antwort</td>
<td>4.870</td>
<td>3,0</td>
</tr>
<tr>
<td>31) Leitung besetzt (später nochmals versuchen)</td>
<td>4.143</td>
<td>2,5</td>
</tr>
<tr>
<td>09) NOTAUSGANG = nächstes INTERVIEW/kein Kommentar</td>
<td>1.563</td>
<td>1,0</td>
</tr>
<tr>
<td>38) keine Antwort</td>
<td>1.243</td>
<td>0,8</td>
</tr>
<tr>
<td>58) keine ZP im Haushalt (z. B. unter Altersgrenze)</td>
<td>1.291</td>
<td>0,8</td>
</tr>
<tr>
<td>35) Bitte auflegen!</td>
<td>1.108</td>
<td>0,7</td>
</tr>
<tr>
<td>98) Kurzfragebogen beendet</td>
<td>792</td>
<td>0,5</td>
</tr>
<tr>
<td>83) Geschäftsanschluss</td>
<td>513</td>
<td>0,3</td>
</tr>
<tr>
<td>81) Geschäftsanschluss (inklusive Warteschleife)</td>
<td>441</td>
<td>0,3</td>
</tr>
<tr>
<td>08 INTERVIEW beenden ===&gt; Wiederaufnahme möglich</td>
<td>428</td>
<td>0,3</td>
</tr>
<tr>
<td>64) ZP nicht in der Lage zu antworten (wg. Krankheit o.ä.)</td>
<td>265</td>
<td>0,2</td>
</tr>
<tr>
<td>66) KP verweigert stellvertretend für ZP</td>
<td>248</td>
<td>0,2</td>
</tr>
<tr>
<td>65) ZP während Feldzeit nicht erreichbar</td>
<td>248</td>
<td>0,2</td>
</tr>
<tr>
<td>72) Zielperson hat gerade keine Zeit -&gt; Termin</td>
<td>287</td>
<td>0,2</td>
</tr>
<tr>
<td>54) KP versteht nicht ausreichend deutsch</td>
<td>287</td>
<td>0,2</td>
</tr>
<tr>
<td>61) ZP kurzfristig krank/ fühlt sich unwohl</td>
<td>254</td>
<td>0,2</td>
</tr>
<tr>
<td>57) KP ist nicht auskunftsfähig (z. B. Kind, verwirrte)</td>
<td>265</td>
<td>0,2</td>
</tr>
<tr>
<td>63) ZP versteht nicht ausreichend deutsch</td>
<td>248</td>
<td>0,2</td>
</tr>
<tr>
<td>44) Anschluss ist vorübergehend nicht erreichbar</td>
<td>211</td>
<td>0,1</td>
</tr>
<tr>
<td>41) Nummer nicht vergeben</td>
<td>140</td>
<td>0,1</td>
</tr>
<tr>
<td>59) kein Privathaushalt (Anstalt o.ä.)</td>
<td>120</td>
<td>0,1</td>
</tr>
<tr>
<td>56) KP ist nur zufällig anwesend (z. B. Handwerker o.ä.)</td>
<td>94</td>
<td>0,1</td>
</tr>
<tr>
<td>55) Technische Verständigungsprobleme mit KP</td>
<td>80</td>
<td>0</td>
</tr>
</tbody>
</table>

*Fortsetzung S. 43*

Abschließend werden die Ergebnisse aller Anwahlvorgänge tabellarisch nach Auftretenshäufigkeit sortiert aufgelistet, die während der gesamten Feldzeit im GSTel03 durchgeführt wurden (Tabelle 2). Insgesamt wurden während der 112 Tage Feldzeit über 160.000 Anwahlvorgänge durchgeführt.


Die Dokumentation der Anwahlvorgänge in der call history erfüllt aber noch andere Zwecke. So ergeben sich aus der Häufigkeit der Anwahlergebnisse im Feldverlauf bereits erste Aufschlüsse über die Qualität der Stichprobenbasis. Darüber hinaus lassen sich auch Fehlsteuerungen im call-back Management identifizieren. Eine bedeutende Rolle spielt die Verlaufsdatei darüber hinaus im Kontext der feldbegleitenden Qualitätssicherung. So lassen sich aus der ungewöhnlichen Häufung bestimmter Anwahlergebnisse Defizite in der Schulung der Interviewer erkennen oder aber auch eventuelle Täuschungsmanöver identifizieren. Die Qualitätsstandards im BRFSS sehen dafür auch spezielle Quotienten der Anwahlergebnisse pro Interviewer vor, die während der Feldzeit laufend kontrolliert werden sollten.
4.5 Datenhaltung


4.6 Sampleausschöpfung und Response


Für die Darstellung der Ausschöpfungsquoten von telefonischen Befragungen sind grundsätzlich folgende Punkte von Bedeutung:

- Das gewählte Stichprobenkonzept (Random vs. Register)
- Die Festlegung der maximalen Anwahlversuche
- Die Differenziertheit der disposition codes
- Damit in Verbindung stehend die Feldsteuerung, d. h. die Regeln, nach denen eine Telefonnummer zur Wiedervorlage zugespielt wird (call-back rules)
- Das Verfahren, nach dem bei Annäherung an die Sollzahl das Verhältnis »neue« zu »alte« Nummern gesteuert wird (wind-down-procedure)
- Damit einhergehend die Aufsplittung des Gesamtsamples in Teilsample oder Quoten
- Die fernmeldetechnische Ausstattung der Telekommunikations-Anlage (insbesondere: Analog- vs. ISDN-Anschluss)
- Damit verbunden, dass verwendetet Wählverfahren (manuell, auto-dialing, predictive-dialing)
- Funktionalität und Stabilität der Soft- und Hardware sowie die Netzstabilität und
- nicht zuletzt: die Qualität der Schulung von Interviewerinnen und Supervisoren (Feldteam-Performance)\(^\text{15}\)

Wegen der Vielzahl zu berücksichtigender Faktoren hat sich bisher noch kein allgemeiner Standard für die Darstellung der Ausschöpfung bei Telefonstichproben durchgesetzt, zumal sich die angeführten Faktoren auch wechselseitig beeinflussen. Die zentralen Aspekte für die Ergebnisaufbereitung und -darstellung liegen jedoch beim Stichprobenkonzept sowie in der Differenziertheit der disposition codes, sofern mit einem call-back-Managementsystem gearbeitet wird.

\(^\text{15}\) Vgl. auch Borch, Rieck (2005) in diesem Band

Dem Vorgehen im GSTelo3 am nächsten kommt das ADM-Verfahren, welches auf einem modifizierten Gabler-Häder Stichprobendesign beruht (Heckel 2002; Heyde 2002). Die wesentlichen Unterschiede bestehen in der Differenziertheit der disposition codes sowie in der Anzahl der maximalen Kontaktversuche pro zufällsgenerierter Telefonnummer.

Die folgende Darstellung der Ausschöpfungsquote orientiert sich primär an der ADM-Schematik, weil deren Erhebungen sowohl hinsichtlich des Stichprobendesigns als auch in den Durchführungsbedingungen dem Vorgehen beim GSTelo3 am nächsten kommen. Aus der ADM-Schematik wird ergänzend die Unterscheidung zwischen Haus- halts- und Zielpersonenkontakten übernommen. Der Hauptgeißpunkt dabei ist, dass erst durch Personenkontakt sicher gestellt werden kann, ob die zufällsgenerierten Nummern Privathaushalte, also Personen unserer Grundgesamtheit enthalten. Erst dadurch wird eine sichere Bestimmung der stichprobennutralen Ausfälle möglich.

4.6.1 Leitende Kriterien für die Klassifikation der disposition codes


Unter verlaufsbezogener Betrachtung stellt sich zudem die Frage, welche Einordnung Ergebnissen zukommt, deren Status nicht endgültig geklärt werden konnte, bzw. wegen der Regeln und maximaler Anzahl der Anrufe während der begrenzten Feldzeit offen bleiben.

Daraus folgt, dass sowohl die call-back rules als auch die maximale Zahl der Anwahlversuche sowie die Verweildauer eines Datensatzes während der Feldarbeit im Nummern-Pool, Auswirkungen auf die Berechnung der Responsequote haben. Um diesen unterschiedlichen Perspektiven in der Auswertung der call history gerecht werden zu können, wurden die in den disposition codes festgehaltenen Anwahlergebnisse nach folgenden Gesichtspunkten verdichtet. Dabei wurde auf die in den USA durch die AAPOR und CASRO entwickelten Standards soweit möglich Bezug genommen (AAPOR 2004; CASRO 1982).

16 Institut für Biometrie und Epidemiologie München u. a.: (Weitkunat 1998; Fischer et al. 2001; Wiedenmayer 2002)
17 ADM: Arbeitsgemeinschaft deutscher Markt- und Meinungsforschungsinstitute
18 beim ADM-Ansatz: Anzahl der disposition codes in der komprimierten Darstellung = 24, maximale Anzahl Kontakte = 5, gegenüber 46 disposition codes und bis zu 15 Kontaktversuchen beim GSTelo3.
19 Council of American Survey Research Organizations
Ein wesentlicher Unterschied zu den deutschsprachigen Veröffentlichungen beruht dabei auf der Differenzierung zwischen »temporary«-, »final«- und »action«-Codes. Dies ist deshalb wichtig, weil der Endzustand der call history nur unzureichende Informationen über den Status der jeweiligen zufallsgenerierten Rufnummer erlaubt. Die ausschließliche Betrachtung des letzten Anwahlergebnisses und dessen Zuordnung zu einer der vier Hauptgruppen

- Interview – teilweise oder vollständig (eligible, with response)
- Element der Grundgesamtheit, aber kein Kontakt (eligible, no response)
- Status der Nummer konnte nicht geklärt werden (unkown eligibility)
- Nummer nicht Element der Grundgesamtheit (not eligible)


---

20 Human-contact-over-other-outcome rule (AAPOR 2004, S. 9)
Tabelle 3
Ausschöpfung des Nummernsamples im GSTelo3 – Endzustand der Call-Results

<table>
<thead>
<tr>
<th>Ausschöpfung</th>
<th>Absolut</th>
<th>Brutto-stichprobe</th>
<th>Netto-stichprobe</th>
<th>Personen-kontakt</th>
<th>Zielpersonen-kontakt</th>
<th>Vollständige Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruttoansatz</td>
<td>45.819</td>
<td>100,0 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telekom-Ansage/Falsche Nummer</td>
<td>18.604</td>
<td>40,6 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max-Kontakt erreicht</td>
<td>4.863</td>
<td>10,6 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAX/Modem</td>
<td>715</td>
<td>1,6 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kein Privathaushalt</td>
<td>2.601</td>
<td>5,7 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doppeladresse</td>
<td>–</td>
<td>0,0 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falsche Regionalzelle/Quote voll</td>
<td>9</td>
<td>0,0 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sprachprobleme</td>
<td>252</td>
<td>0,5 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonstige neutrale Ausfälle</td>
<td>1.268</td>
<td>2,8 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutrale Ausfälle insg. (Summe 1–8)</td>
<td>28.312</td>
<td>61,8 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bereinigtes Brutto</td>
<td>17.507</td>
<td>38,2 %</td>
<td>100,0 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teilnehmer nimmt nicht ab, Status noch offen</td>
<td>242</td>
<td>0,5 %</td>
<td>1,4 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anrufbeantworter</td>
<td>293</td>
<td>0,6 %</td>
<td>1,7 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personenkontakte insg. (Brutto Privat-Haushalte)</td>
<td>16.957</td>
<td>37,0 %</td>
<td>96,9 %</td>
<td>100,0 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terminvereinbarung nicht wahrnehmbar</td>
<td>73</td>
<td>0,2 %</td>
<td>0,4 %</td>
<td>0,4 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kein Termin möglich während Feldzeit</td>
<td>313</td>
<td>0,7 %</td>
<td>1,8 %</td>
<td>1,8 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP ist beschäftigt, nicht erreichbar</td>
<td>–</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP gibt keine Auskunft: Thema/Projekt</td>
<td>–</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP gibt keine Auskunft: anderer/kein Grund</td>
<td>69</td>
<td>0,2 %</td>
<td>0,4 %</td>
<td>0,4 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP verweigert absolut jede Teilnahme</td>
<td>584</td>
<td>1,3 %</td>
<td>3,3 %</td>
<td>3,4 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zielpersonenkontakte insg. (Brutto Zielperson)</td>
<td>15.918</td>
<td>34,7 %</td>
<td>90,9 %</td>
<td>93,9 %</td>
<td>100,0 %</td>
<td></td>
</tr>
<tr>
<td>ZP ist beschäftigt, nicht erreichbar</td>
<td>107</td>
<td>0,2 %</td>
<td>0,6 %</td>
<td>0,6 %</td>
<td>0,7 %</td>
<td></td>
</tr>
<tr>
<td>ZP gibt keine Auskunft: Thema/Projekt</td>
<td>–</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td></td>
</tr>
<tr>
<td>ZP gibt keine Auskunft: anderer/Grund</td>
<td>–</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td>0,0 %</td>
<td></td>
</tr>
<tr>
<td>ZP verweigert Teilnahme*</td>
<td>449</td>
<td>1,0 %</td>
<td>2,6 %</td>
<td>2,6 %</td>
<td>2,8 %</td>
<td></td>
</tr>
<tr>
<td>ZP ist krank, nicht in der Lage</td>
<td>361</td>
<td>0,8 %</td>
<td>2,1 %</td>
<td>2,1 %</td>
<td>2,3 %</td>
<td></td>
</tr>
<tr>
<td>Interview-Verweigerung zu Beginn/Abbruch</td>
<td>5.531</td>
<td>12,1 %</td>
<td>31,6 %</td>
<td>32,6 %</td>
<td>34,7 %</td>
<td></td>
</tr>
<tr>
<td>Interview nicht auswertbar</td>
<td>44</td>
<td>0,1 %</td>
<td>0,3 %</td>
<td>0,3 %</td>
<td>0,3 %</td>
<td></td>
</tr>
<tr>
<td>System. Ausfälle insg. (Summe 9–11, 12–17, 18–24)</td>
<td>8.081</td>
<td>17,6 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durchgeführte Interviews (Netto Zielperson)</td>
<td>9.426</td>
<td>20,6 %</td>
<td>53,8 %</td>
<td>55,6 %</td>
<td>59,2 %</td>
<td>100,0 %</td>
</tr>
<tr>
<td>davon Vollinterviews (Netto Stichprobe)**</td>
<td>8.318</td>
<td>18,2 %</td>
<td>47,5 %</td>
<td>49,1 %</td>
<td>52,3 %</td>
<td>88,2 %</td>
</tr>
<tr>
<td>davon Kurzfragebögen</td>
<td>1.108</td>
<td>2,4 %</td>
<td>6,3 %</td>
<td>6,5 %</td>
<td>7,0 %</td>
<td>11,8 %</td>
</tr>
</tbody>
</table>

* nach Statusklärung, ob es sich um einen Privathaushalt handelt
** endgültig, nach Bereinigung
4.7 Datenkontrolle und -aufbereitung


<table>
<thead>
<tr>
<th>Ausschlussgrund</th>
<th>Anzahl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geschlecht falsch kodiert</td>
<td>10</td>
</tr>
<tr>
<td>Verständnisprobleme (mental oder körperlich)</td>
<td>10</td>
</tr>
<tr>
<td>Nichtzugehörigkeit zum Haushalt, Geschäftsanschluss, Anstaltsbewohner oder Fehler bei Zielpersonenermittlung</td>
<td>8</td>
</tr>
<tr>
<td>Sprachliche Probleme</td>
<td>6</td>
</tr>
<tr>
<td>Versteckte oder latente Verweigerer (Interviewabbruch, Zweifel an der Ernsthaftigkeit)</td>
<td>6</td>
</tr>
<tr>
<td>Falsches Geburtsjahr</td>
<td>2</td>
</tr>
<tr>
<td>Betrunken</td>
<td>1</td>
</tr>
<tr>
<td>Ausschluss wegen fehlender Angaben in zentralen Merkmalen</td>
<td>1</td>
</tr>
<tr>
<td><strong>Summe</strong></td>
<td><strong>44</strong></td>
</tr>
</tbody>
</table>

22 vgl. Borch, Rieck 2005 in diesem Band sowie das OP-Handbuch zum GSTel03 (Borch et al. 2004)
4.8 Gewichtung der Daten


Vor der Anpassungsgewichtung muss jedoch zunächst eine Kompensation der rein designbedingten Verzerrungen des jeweils verwendeten Stichprobenverfahrens vorgenommen werden (Gabler, Häder 1999; Glemser 2002). Die so genannte Designgewichtung wird nachfolgend beschrieben.

4.8.1 Designgewichtung


In Verbindung mit der Zunahme der ISDN-Anschlüsse, die es auch Privathaushalten ermöglicht, kostengünstig über mehrere Telefonnummern zu verfügen, haben Haushalte mit ISDN-Anschluss bei mehreren Privatnummern im selben Haushalt eine höhere Wahrscheinlichkeit befragt zu werden. Diese erhöhte Auswahlwahrscheinlichkeit war daher durch den Kehrwert der »Anzahl der privat genutzten Festnetzanschlüsse im Haushalt« zu kompensieren

Vor dem Hintergrund des Ziels, bevölkerungsrepräsentative Aussagen auf Personenebene für die erwachsene Bevölkerung in der Bundesrepublik zu gewährleisten, ist in einem weiteren Schritt eine Anpassung an die Haushaltsgröße erforderlich, da Telefonstichproben prinzipiell als Haushaltsstichproben angesehen werden können. Personen, die in Haushalten mit mehreren Erwachsenen leben, haben eine um die Anzahl der

Erwachsenen im Haushalt geringere Chance, befragt zu werden. Somit ergäbe sich ein designbedingter Gewichtungsfaktor $GWd_i$ als

$$GWd_i = \frac{AEPHH_i}{ATHHi},$$

mit

$AEPHH_i$ Anzahl der erwachsenen Personen im Privat-Haushalt $i$,

$ATHHi$ Anzahl der privat genutzten Festanschlüsse im Haushalt $i$.


23 nach neuesten Ergebnissen der Einkommens- und Verbraucherstichprobe verfügen mittlerweile 23,4 % der Privathaushalte über ISDN-Anschlüsse (Statistisches Bundesamt 2003)

24 Im GSTel03 wurde die Anzahl der Festnetzanschlüsse, die ausschließlich privat genutzt werden als 2-stellige offene Angabe mit einem Grenzwert von 20 festgehalten. Von einer Kategorisierung z. B. 1, 2, 3 und mehr Anschlüsse wurde wegen der weiten Verbreitung von ISDN-Anschlüssen abgesehen. Es war allerdings überraschend, dass auch Angaben von mehr als 20 Telefonanschlüssen genannt wurden. Darüber hinaus entstanden Ungenauigkeiten auch bei der Abgrenzung von Telefonanschlüssen, die nicht ausschließlich als Fax- oder Modemanschluss betrieben und auch zum Telefonieren genutzt wurden. Weiterhin wurde beobachtet, dass in einigen Fällen die Rufnummern, die angewählt wurden, dem oder der Befragten selbst gar nicht bekannt war oder ist. Dieses Phänomen tritt in den laufenden CATI-Projekten mit zunehmender Tendenz auf.

25 u. a. Kappungsgrenzen bei 3, 4 oder 5 und mehr Anschlüssen, log-Transformationen
Extremen ergebenden Faktoren, bei Beibehaltung der Größenrelation. Darüber hinaus ist dieses Verfahren sowohl transparent als auch reversibel, was auf das in vergleichbaren Situationen zumeist verwendete Verfahren der Begrenzung auf einen Maximalwert (z. B. 1, 2, 3 und mehr Telefonanschlüsse etc.) nicht zutrifft.

Zusätzlich ist bei der Designgewichtung noch die Normierung auf den Stichprobenumfang $n = 8.318$ zu berücksichtigen. Somit ergibt sich das im GSTel03 verwendete adjustierte und normierte Designgewicht $GWd$, nach der Formel:

$$GWd^{*}_i = \sqrt{\frac{AEPH_i}{ATHH_i}} \times \frac{n}{\sum \sqrt{\frac{AEPH_i}{ATHH_i}}}$$

### 4.8.2 Anpassungsgewichtung


$$GWr^{*}_i = \frac{P^{RAS\text{ soll}}_i}{P^{RAS\text{ soll}}_i}, \text{mit}$$

$P^{RAS\text{ soll}}_i$ rel. Zellbesetzung der Kombination aus den Merkmalen Region, Alter, Geschlecht in der Bevölkerung

$P^{RAS\text{ soll}}_i$ rel. Zellbesetzung der Kombination aus den Merkmalen Region, Alter, Geschlecht in der designgewichteten Stichprobe

Das für die Auswertungen zu verwendende Endgewicht $w03$ ergibt sich schließlich nach der Formel:

$$W03_i = GWd^{*}_i \times GWr^{*}_i$$

### 4.8.3 Güte der Gewichtung

Als Gütekriterien für Gewichtungen werden in der Literatur (Rösch 1994) der minimale und maximale Gewichtungsfaktor, die Spannweite und Streuung der Faktoren sowie das Effizienzmaß $E$ genannt. Letzteres ist definiert als die durchschnittliche Summe der Abweichungsquadrat vom »Idealwert« pro Zelle in Prozent:

$$E = \frac{\sum w03^2}{\sum w03^2} \times 100$$

Je näher die Effektivität am Wert 100 % liegt, desto geringer sind die gewichtungsbedingten Verzerrungen der Ausgangsdaten. Werte des Effizienzmaßes über 80 % gelten als gute Anpassung.

Das Ergebnis zur Güte der Gewichtung des GSTel03 zeigt die Tabelle 5 – Gewichtungsvariable $w03$. Zum Vergleich werden die Gewichtungsergebnisse des BGS98 mit aufgeführt – Gewichtungsvariable $w98$ (Stolzenberg 2002).

<table>
<thead>
<tr>
<th>Kriterien</th>
<th>$w03$</th>
<th>$w98$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effektivität</td>
<td>87,10 %</td>
<td>84,10 %</td>
</tr>
<tr>
<td>Min. Faktor</td>
<td>0,482</td>
<td>0,275</td>
</tr>
<tr>
<td>Max. Faktor</td>
<td>1,669</td>
<td>4,15</td>
</tr>
<tr>
<td>Fälle</td>
<td>8.318</td>
<td>7.124</td>
</tr>
</tbody>
</table>

26 Wir danken dem Statistischen Bundesamtes für die Bereitstellung der Sonderauswertungen
27 bis 29 Jahre, 30 bis 39 Jahre, 40 bis 49 Jahre, 50 bis 59 Jahre, 60 bis 69 Jahre sowie 70 und mehr Jahre

4.9 Ergebnisse zu Response und Nonresponse

Die Ausschöpfungsquote sagt, unabhängig von deren Höhe, noch nichts über das Ausmaß möglicher Selektivität aus (DFG 1999). Daher stellt sich die Frage: Was wissen wir über die Nonresponder und welche Quellen sind verfügbar? Im Vordergrund stehen dabei vor allem gesundheitliche Gründe, die eine Teilnahme verhinderten sowie soziodemografische Merkmale der Nichtteilnehmer. Neben Geschlecht und Alter interessieren in diesem Zusammenhang auch sprachliche Probleme, die eine Beteiligung der mittlerweile auf nahezu 10% gestiegenen Anteil der ausländischen Bevölkerung in der Bundesrepublik am GSTelo3 beeinträchtigen können.

Erste Hinweise auf systematische Selektionseffekte ergeben sich aus der Dokumentation der Anwahlergebnisse, wie sie sich in den disposition codes (Tabelle 2) wieder finden. So konnte mit 252 kontaktierten Zielpersonen wegen sprachlicher Gründe kein Interview durchgeführt werden. Weitere 361 ermittelte Zielpersonen sahen sich aus gesundheitlichen Gründen nicht in der Lage, an der Befragung teilzunehmen. Über diese Personen liegen keine weiteren Angaben vor.


4.9.1 Bekannte Gründe für Nonresponse


Tabelle 6

Gründe für Nicht-Teilnahme (Mehrfachnennungen)

<table>
<thead>
<tr>
<th>Begründung für Nicht-Teilnahme</th>
<th>Anzahl</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>aus zeitlichen Gründen</td>
<td>675</td>
<td>57,9</td>
</tr>
<tr>
<td>nehme grundsätzlich nicht teil</td>
<td>259</td>
<td>22,2</td>
</tr>
<tr>
<td>aus gesundheitlichen Gründen</td>
<td>83</td>
<td>7,1</td>
</tr>
<tr>
<td>sehe keinen Nutzen darin</td>
<td>116</td>
<td>9,9</td>
</tr>
<tr>
<td>sonstige Gründe</td>
<td>115</td>
<td>9,9</td>
</tr>
<tr>
<td>Antwortende insgesamt</td>
<td>1.166</td>
<td>100</td>
</tr>
</tbody>
</table>

29 Die Abweichung zu der in Tabelle 2 ausgewiesenen Zahl von 1.108 ergibt sich deshalb, weil nicht alle Teilnehmer den Kurzfragebogen vollständig beantwortet haben.
30 Im BGS98 lag der Anteil der gesundheitlichen Gründe für die Nicht-Teilnahme bei über 10%, Stolzenberg (2002)
4.9.2 Analyse soziostukturrelle und gesundheitsbezogener Merkmale von Teilnehmern und Nicht-Teilnehmern


Demgegenüber waren weniger Gesunde und durch mindestens eine chronische Erkrankung betroffene Personen eher bereit, an der Befragung teilzunehmen. Befragungsteilnehmer weisen darüber hinaus einen höheren Anteil von Rauchern aus.

Vor allem bei hochbetagten Frauen und Personen mit niedriger Bildung dürften mögliche Ver-

Tabelle 7
Soziodemografische Merkmale und Befragungsbeteiligung

<table>
<thead>
<tr>
<th>Teilnahme</th>
<th>Geschlecht</th>
<th>Alter</th>
<th>Anzahl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frauen</td>
<td>Männer</td>
<td>–&lt;40</td>
</tr>
<tr>
<td>Nonresponder</td>
<td>56,2%</td>
<td>43,8%</td>
<td>1.541</td>
</tr>
<tr>
<td>Responder</td>
<td>53,4%</td>
<td>46,6%</td>
<td>8.362</td>
</tr>
<tr>
<td>Gesamt</td>
<td>53,9%</td>
<td>46,1%</td>
<td>9.903</td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,046</td>
<td>P = 0,000</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bildung</th>
<th>Anzahl</th>
<th>Ausländer</th>
<th>Anzahl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niedrig</td>
<td>69,1%</td>
<td>30,9%</td>
<td>1.117</td>
</tr>
<tr>
<td>Hoch</td>
<td>58,6%</td>
<td>41,4%</td>
<td>8.196</td>
</tr>
<tr>
<td>Gesamt</td>
<td>59,8%</td>
<td>40,2%</td>
<td>9.313</td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,000</td>
<td>P &gt; 0,05</td>
<td></td>
</tr>
</tbody>
</table>

Tabelle 8
Gesundheitsbezogene Merkmale und Befragungsbeteiligung

<table>
<thead>
<tr>
<th>Teilnahme</th>
<th>Subjektive Gesundheit</th>
<th>Anzahl</th>
<th>Chron. Erkrankung</th>
<th>Anzahl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Schlecht</td>
<td>Mittel</td>
<td>Gut</td>
<td>Nein</td>
</tr>
<tr>
<td>Nonresponder</td>
<td>8,1%</td>
<td>18,5%</td>
<td>73,4%</td>
<td>1.123</td>
</tr>
<tr>
<td>Responder</td>
<td>21,7%</td>
<td>50,3%</td>
<td>27,9%</td>
<td>8.348</td>
</tr>
<tr>
<td>Gesamt</td>
<td>20,1%</td>
<td>46,6%</td>
<td>33,3%</td>
<td>9.471</td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,000</td>
<td>P = 0,046</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Raucher</th>
<th>Anzahl</th>
<th>Gewicht (BMI)</th>
<th>Anzahl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nein</td>
<td>Ja</td>
<td>Unter</td>
<td>Norm</td>
</tr>
<tr>
<td>Nonresponder</td>
<td>71,8%</td>
<td>28,2%</td>
<td>1.180</td>
</tr>
<tr>
<td>Responder</td>
<td>66,0%</td>
<td>34,0%</td>
<td>8.360</td>
</tr>
<tr>
<td>Gesamt</td>
<td>66,8%</td>
<td>33,2%</td>
<td>9.540</td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,000</td>
<td>P &gt; 0,05</td>
<td></td>
</tr>
</tbody>
</table>

* Wechselnde Fallzahlen ergeben sich daraus, dass nicht alle Befragten den Kurzfragebogen vollständig beantwortet haben
zerrungen, aufgrund deren geringerer Beteiligung an der Befragung am stärksten sein.

Einschränkend ist an dieser Stelle noch anzu-
merken, dass die hier vorliegenden Ergebnisse nur erste Hinweise auf gruppenpezische Nonre-
sponse geben können. Zum einen stellen die Non-
responder eine Selektion insofern dar, als sie zu-
mindest noch bereit waren, ein Minimum an
Fragen zu beantworten. Zum anderen liegen über
den größeren Teil der nicht Teilnahmewilligen kei-
ne zusätzlichen Informationen vor. Hier können
erst weitere Analysen im Vergleich zu externen
Daten, z. B. dem Mikrozensus, tieferen Aufschnal
über gruppenspezifischen Nonresponse liefern.

4.9.3 Intensität der Samplenutzung

Selektivität bei – telefonischen – Befragungen
kann auch aus der systematisch besseren Erreich-
barkeit bestimmter Bevölkerungsgruppen resulti-
tieren. In diesem Zusammenhang spielt die
Intensität der Sampleausschöpfung und die
Kontakthäufigkeit einer Telefonnummer eine be-
deutende Rolle. Die maximale Kontaktzahl im
GSTel03 wurde aus Qualitätsgründen auf 15 fest-
gesetzt und geht damit – wie oben bereits erwähnt
– deutlich über das übliche Maß hinaus.

Nachfolgend wird anhand ausgewählter
Merkmale untersucht, welche Konsequenzen sich
aus der intensiveren Sampleausnutzung auf die
Ergebnisse des GSTel03 ergeben. Hierzu wurde
die Anzahl der Kontakte die zur Durchführung des
Interview benötigt wurden in drei Gruppen – bis
zu 5, zwischen 5 bis 9 und mehr als 9 Kontaktver-
suche – zusammengefasst und in den Zeilen der
Tabellen 8 bis 13 dargestellt. Die letzte Zeile in den
abgebildeten Tabellen gibt immer die Prozent-
satzdifferenz der ersten Zeile, bei bis zu 5 Kon-
taktversuchen und der aktuellen Verteilung – mit
maximaler Anzahl der Kontakte – des jeweils dar-
gestellten Merkmals an.

Die intensivere Sampleausnutzung im
GSTel03 verändert die soziodemografische Zu-
sammensetzung in Richtung jüngerer, männ-
licher Teilnehmer und führt zu einer leichten Zu-
nahme des Anteils der Migrant. Die mittleren
und höheren Bildungsgrade steigen ebenfalls mit
zunehmender Anzahl der Kontaktversuche an
und führen in der Folge zu einer Zunahme des
Anteils der höheren Sozialschicht. Dies schlägt
sich auch bei den Krankenversicherungsarten in
einer Zunahme des Anteils der privat Versicher-

<table>
<thead>
<tr>
<th>Anrufe</th>
<th>Geschlecht</th>
<th>Alter in Dekaden</th>
<th>Migrant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Männer</td>
<td>Frauen</td>
<td>−29</td>
</tr>
<tr>
<td>Bis 5</td>
<td>45,0%</td>
<td>55,0%</td>
<td>13,3%</td>
</tr>
<tr>
<td>5–9</td>
<td>53,3%</td>
<td>46,7%</td>
<td>20,5%</td>
</tr>
<tr>
<td>Mehr als 9</td>
<td>54,8%</td>
<td>45,2%</td>
<td>17,4%</td>
</tr>
<tr>
<td>Gesamt</td>
<td>46,6%</td>
<td>53,4%</td>
<td>14,4%</td>
</tr>
<tr>
<td></td>
<td>1,6%</td>
<td>-1,6%</td>
<td>1,1%</td>
</tr>
<tr>
<td>Signifikanz (χ²)</td>
<td>P = 0,000</td>
<td></td>
<td>P = 0,000</td>
</tr>
</tbody>
</table>
ten nieder. Hinsichtlich der Erwerbstätigkeit ist ein Ansteigen der Vollzeiterwerbstätigen mit zunehmender Kontakthäufigkeit festzustellen.


Andererseits ergibt sich hinsichtlich der demografischen Zusammensetzung ein überlagerner Effekt aus der geringeren Teilnahmebereitschaft von Personen mit niedriger Schulbildung und deren Kompensation durch höher Gebildete der Mittel- und Oberschicht im Zuge der intensiveren Sampleausschöpfung. Dies trägt auch zur Erklärung der bereits anderenorts festgestellten höheren Beteiligungsbereitschaft der mittleren und oberen Sozialschichten an telefonischen Befragungen bei (Schulte 1997; Porst et al. 1998; Wüst 1998). Für die Auswertungen und Analyse der Daten des GSTelo3 folgt daraus, dass bei schichtsensitiven Fragestellungen die soziale Schicht immer in die Modellierung einbezogen werden sollte.

Tabelle 10
Sozialschichtbezogene Merkmale und Kontakthäufigkeit

<table>
<thead>
<tr>
<th>Anrufe</th>
<th>CASMIN-Bildungsgruppen</th>
<th>Soz. Schicht n. Winkler-Index (adj.)</th>
<th>Krankenversicherung verdichtet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Einfache Bildung, Mittlere Bildung, Höhere Bildung</td>
<td>Fehlend, US (3–8), MS (9–14), OS (15–21)</td>
<td>AOK, EKK, BKK, GKV, PKV, Beihilfe, Rest</td>
</tr>
<tr>
<td>Bis 5</td>
<td>31,9 % 47,5 % 20,6 % 1,8 % 15,4 % 53,4 % 29,4 % 23,5 % 37,3 % 16,8 % 6,5 % 13,8 % 2,0 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–9</td>
<td>25,0 % 51,5 % 23,5 % 1,3 % 14,3 % 50,7 % 33,7 % 20,9 % 35,5 % 19,1 % 6,6 % 16,3 % 1,7 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mehr als 9</td>
<td>23,4 % 48,9 % 27,7 % 1,9 % 11,0 % 45,9 % 41,2 % 19,1 % 35,9 % 16,6 % 6,2 % 21,2 % 1,1 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gesamt</td>
<td>30,6 % 48,1 % 21,3 % 1,7 % 15,0 % 52,7 % 30,6 % 22,9 % 37,0 % 17,1 % 6,5 % 14,6 % 1,9 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1,3 % 0,6 % 0,8 % -0,1 % -0,4 % -0,8 % 1,2 % -0,6 % -0,3 % 0,3 % 0,0 % 0,7 % -0,1 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,000</td>
<td>P = 0,000</td>
<td>P = 0,000</td>
</tr>
</tbody>
</table>

Tabelle 11
Erwerbsbeteiligung und Kontakthäufigkeit

<table>
<thead>
<tr>
<th>Anrufe</th>
<th>Rest</th>
<th>Ausbildung</th>
<th>Vorruhestand</th>
<th>Rente</th>
<th>Arbeitslos</th>
<th>Hausfrau</th>
<th>Teilzeit &lt;15 h</th>
<th>Teilzeit &gt;15 h</th>
<th>Vollzeit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bis 5</td>
<td>0,1 %</td>
<td>2,8 %</td>
<td>3,5 %</td>
<td>19,1 %</td>
<td>4,0 %</td>
<td>8,8 %</td>
<td>7,2 %</td>
<td>11,5 %</td>
<td>42,9 %</td>
</tr>
<tr>
<td>5–9</td>
<td>0,1 %</td>
<td>3,7 %</td>
<td>1,6 %</td>
<td>10,8 %</td>
<td>3,4 %</td>
<td>4,9 %</td>
<td>7,5 %</td>
<td>11,0 %</td>
<td>57,0 %</td>
</tr>
<tr>
<td>Mehr als 9</td>
<td>0,2 %</td>
<td>2,8 %</td>
<td>1,5 %</td>
<td>8,5 %</td>
<td>2,5 %</td>
<td>7,0 %</td>
<td>5,9 %</td>
<td>9,1 %</td>
<td>62,4 %</td>
</tr>
<tr>
<td>Gesamt</td>
<td>0,1 %</td>
<td>2,9 %</td>
<td>3,1 %</td>
<td>17,5 %</td>
<td>3,9 %</td>
<td>8,2 %</td>
<td>7,2 %</td>
<td>11,3 %</td>
<td>45,8 %</td>
</tr>
<tr>
<td></td>
<td>0,0 %</td>
<td>0,1 %</td>
<td>-0,3 %</td>
<td>-1,6 %</td>
<td>-0,2 %</td>
<td>-0,6 %</td>
<td>0,0 %</td>
<td>-0,2 %</td>
<td>2,8 %</td>
</tr>
<tr>
<td>Signifikanz (chi²)</td>
<td>P = 0,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.10 Zusammenfassung

Die Ergebnisse zur Response und Teilnahmebereitschaft am GSTel03 lassen eine hohe Übereinstimmung mit den Befunden, die aus der Literatur zu telefonischen Befragungen entnommen werden können, erkennen. Es erweist sich, dass diese Methode nicht allein wegen der geringeren Erhebungskosten und der größeren Zeitnähe gut für den Aufbau eines kontinuierlichen Gesundheitsmonitoring geeignet ist, sondern auch wegen der leichten Verfügbarkeit eines nationalen Auswahlrahmens, der regionalisierbare Analysen ermöglicht. Das setzt allerdings eine entsprechende Sorgfalt bei der Durchführung der Erhebung voraus. Dies betrifft nicht nur die Intensität der Sampleausschöpfung und die Feldsteuerung sondern beinhaltet den gesamten Erhebungsprozess, beginnend mit der Erstellung des Erhebungsinstruments unter Berücksichtigung methodenspezifischer Aspekte, der Schonung der eingesetzten Erhebungskräfte sowie der feldbegleitenden qualitätssichernden Maßnahmen, die in eigenen Beiträgen dieses Bandes ausführlicher behandelt werden (vgl. Borch, Rieck 2005).
4.11 Literatur


Gabler S, Hoffmeyer-Zlotnik HJP (1997) Stichproben in der Umfragepraxis. Westdeutscher Verlag, Opladen


Erster telefonischer Gesundheitssurvey
des Robert Koch-Instituts – Methodische Beiträge

November 2005

Berlin: Robert Koch-Institut
ISBN 3-89606-163-1
The Telephone Health Survey 2003 (GSTel03) is the first nationwide survey conducted by the Robert Koch-Institute using computer-assisted telephone interviewing (CATI). The procedures used in telephone surveys differ in many aspects from face-to-face or mail surveys. The articles in this volume discuss methodological aspects of the GSTel03 related to questionnaire design, data collection procedures, sampling design and response rate. One article examines the need for a questionnaire that is suitable for the respective survey method and discusses the measures taken to achieve this in the preparation for the GSTel03. These include questionnaire dramaturgy, filtering and the optimization of the operationalizations; they focus on breaking down complex questions and/or response categories into smaller units, deciding on the sequence of the fixed answer options, and standardizing retrospective time windows. Another contribution describes the following aspects of data collection procedures: assigned staff, data collection and storage, interviewer and supervisor training, quality assurance on interviewing, training procedures and supervision. A further article discusses in detail the technical framework, sampling design, call-back management, and, derived from this, the weighting procedures used in the first nationwide telephone health survey. Finally, the results of a nonresponse analysis are presented, and there is a discussion on the effects of the number of attempts made to contact households. It is shown that telephone surveys are an appropriate and important method to establish a continuous health monitoring system.