

Gruppenbezogene Menschenfeindlichkeit

Methodenbericht zur repräsentativen Befragung vom Mai/Juni 2003

Durchgeführt im Auftrag
der Universität Bielefeld
Prof. Dr. Wilhelm Heitmeyer

Infratest Sozialforschung

Dr. Thomas Gensicke

München, 08. Juli 2003
Mb10-31817

Inhaltsverzeichnis	Seite
Vorbemerkung	3
1. Methodische Anlage telefonischer Bevölkerungsumfragen von NFO Infratest	4
1.1 Überblick	4
1.2 Grundgesamtheit und Hochrechnungsrahmen	4
1.3 Stichprobenanlage und Stichprobensteuerung	5
1.4 ADM-Standard des Random-Digit-Dialing	6
1.5 Haushalts- und Personenstichprobe (Umwandlung)	6
2. Überarbeitung des CATI-Fragebogens nach dem Pretest	7
2.1 Fragebogenlänge: Reduktion durch Kürzung des Fragebogens	7
2.2 „Fragebogendramatik“: Entspannung des Interviews	8
2.3 Filterführung	9
3. Interviewerschulung	9
4. Durchführung der Befragung	9
4.1 Interviewdauer	9
4.2 Abbrüche von Interviews	10
4.3 Stichprobenausschöpfung	10
5. Erfassung von Ausländern	14
6. Die Bruttodaten	14
7. Gewichtung der Nettostichprobe	17
Anhang/Fragebogen	19

Vorbemerkung

Dieser Methodenbericht dokumentiert die **zweite Repräsentativbefragung** im Rahmen des Forschungsprojekts „Langzeituntersuchung des Einstellungssyndroms ‚Gruppenbezogene Menschenfeindlichkeit‘ in der Bevölkerung (2002 – 2011)“, durchgeführt unter Leitung von Prof. Dr. W. Heitmeyer, Universität Bielefeld, Institut für interdisziplinäre Konflikt- und Gewaltforschung.

Das Projekt beinhaltet eine Trendbeobachtung von Einstellungen über einen Zeitraum von 10 Jahren. Grundlage sind telefonisch durchgeführte Repräsentativbefragungen der Bevölkerung ab 16 Jahren mit einem Stichprobenumfang von jeweils 3.000 Interviews. Der Zeitabstand zwischen den Befragungen beträgt jeweils ein Jahr.

Infratest Sozialforschung wurde als Kooperationsinstitut mit der Durchführung der Befragungen beauftragt. Der Auftrag umfasst zunächst die Erhebungen in den drei Jahren 2002 – 2004. In diesem Zeitraum wird ergänzend zu den jährlichen Querschnittsstichproben zusätzlich eine Längsschnittbefragung mit drei Wellen durchgeführt.

Die erste Befragung mit 3.000 Interviews erfolgte im Mai/Juni 2002, die aktuelle im Mai/Juni 2003. Die Daten wurden dem Auftraggeber Anfang Juli übergeben. Im vorliegenden Bericht werden Anlage und Durchführung der Befragung unter methodischen Gesichtspunkten dokumentiert und erläutert.

1. Die methodische Anlage telefonischer Bevölkerungsumfragen von NFO Infratest

1.1 Überblick

NFO Infratest arbeitet in großem Umfang mit der Methode der telefonischen Befragung. Im Jahr 2002 wurden von sechs Telefonstudios aus (München, Frankfurt, Bielefeld, Berlin, Parchim, Güstrow) insgesamt rd. 1.000.000 Telefoninterviews durchgeführt. Die Bandbreite der Anwendungen reicht von kontinuierlichen Mehrthemenbefragungen (InfraScope) bis zu komplexen sozialwissenschaftlichen Erhebungen, die auch in der Fachwelt große Beachtung gefunden haben.¹

Die hohe Qualität der telefonischen Befragungen wird bei NFO Infratest durch ein Zusammenwirken verschiedener Steuerungsinstrumente erreicht:

- das CATI (Computer-Assisted Telephone Interview)
- das ITMS (Infratest-Telefonhaushalts-Master-Sample)
- das SMS (Sample-Management-System)
- den Autodialer (automatisches Anwählen der zufällig ausgewählten Telefonnummern)
- das System der kontinuierlichen Kontrolle, Unterstützung und gegebenenfalls Schulung der Interviewer online (Mithören während des Interviews) und offline (qualitätsorientierte Leistungskennziffern).

Gegenstand des vorliegenden Projekts ist die Durchführung einer repräsentativen Bevölkerungsumfrage mit 3.000 befragten Personen. Die Stichprobenanlage ist disproportional im Sinne eines überproportionalen Anteils der ostdeutschen Bevölkerung. Zu befragen waren 2.000 Personen in den alten Bundesländern (einschließlich Berlin) und 1.000 Personen in den neuen Bundesländern.

1.2 Grundgesamtheit und Hochrechnungsrahmen

Grundgesamtheit von Bevölkerungsbefragungen sind in der Regel alle in Privathaushalten lebenden deutschsprachigen Personen ab 14 Jahren. Im vorliegenden Projekt wurde eine altersmäßige Eingrenzung auf Personen ab 16 Jahren vorgenommen.

Die telefonische Durchführung der Befragung reduziert die Grundgesamtheit auf die Auswahlgrundlage „deutschsprechende Personen ab 16 Jahren in Telefonhaushalten“.

Sowohl bei der Stichprobenanlage (Schichtung a priori) als auch bei der Gewichtung (Schichtung a posteriori) wird die Stichprobe jedoch nach der Verteilung der Privathaushalte bzw. den

¹ Vgl. etwa C. Babka von Gostomski, J. Hartmann und M. Thum: Die Mannheimer Scheidungsstudie: Aspekte der Durchführung der telefonischen Befragung zu Determinanten der Ehescheidung. In: ZUMA-Nachrichten Nr. 41, November 1997.

soziodemographischen Strukturen der deutschen Wohnbevölkerung und nicht etwa nach den Telefonhaushalten bzw. der deutschen „Telefonbevölkerung“ ausgerichtet. Der Anteil der Telefonhaushalte an den Privathaushalten liegt gegenwärtig bei über 98% im Westen und ca. 96% im Osten. Merkmalsunterschiede zwischen Telefonhaushalten und Privathaushalten sind deshalb ohnehin klein und können nach dem vorgestellten Stichproben- und Gewichtungsverfahren für die allermeisten Merkmale praktisch vernachlässigt werden.

1.3 Stichprobenanlage und Stichprobensteuerung

Die Befragung basiert auf dem Infratest-Telefon-Master-Sample (ITMS), das für derartige Untersuchungen aufgebaut wurde und zu verzerrungsfreien Stichproben (insbesondere weitgehende Vermeidung des Not-at-home-Bias) ohne Klumpeneffekte führt. Es entspricht hinsichtlich des Random-Digit-Dialing dem ADM-Standard, der nach dem Gabler-Häder-Verfahren definiert wurde (siehe dazu Abschnitt 1.4).

Das ITMS ist so differenziert geschichtet, dass praktisch jede Gemeinde über 5.000 Einwohnern in der Stichprobe vertreten ist und eine eigene Schicht bildet, in der durch eine systematische Zufallsauswahl Telefonhaushalte proportional zur Zahl aller Privathaushalte in der Gemeinde ausgewählt werden. Somit erfolgt für 85% der Auswahleinheiten eine einstufige Auswahl der Zielhaushalte. Lediglich bei den Gemeinden unter 5.000 Einwohnern, in denen nur etwa 15% der Wohnbevölkerung leben, wird aus Gründen der Forschungsökonomie eine mehrfach geschichtete („multistratifizierte“) Stichprobe von Gemeinden (mit Auswahl-Wahrscheinlichkeiten proportional zur Zahl der Privathaushalte pro Schicht) gezogen, und daraus werden in der zweiten Auswahlstufe die Zielhaushalte ausgewählt. In jedem Fall erfolgt die Auswahl repräsentativ für die betreffende Gemeinde, auch wenn in einem Ortsnetz mehrere Gemeinden enthalten sind bzw. wenn die Gemeinde sich auf mehrere Ortsnetze aufteilt. Die Zufallsauswahl innerhalb der Gemeinden erfolgt ungeklumpt; somit handelt es sich um eine Stichprobe ohne Klumpeneffekte.

Das ITMS besteht also aus einer mikrostratifizierten, weitgehend ungeklumpten Stichprobe, die sich proportional zur Zahl der Privathaushalte auf die Mikrozellen aufteilt. Die Multistratifikation und Aufteilung der Stichprobe auf die Zellen erfolgt vollautomatisch über ein Allokationsprogramm.

Innerhalb jeder Zelle sind die Datensätze der Telefonhaushalte nach Zufallszahlen sortiert. Somit bildet jede Zelle eine Urne im klassischen Sinne. Nicht erreichte Haushalte werden „zurückgelegt“ und kommen in größerem zeitlichem Abstand zu anderen Tageszeiten zur „Wiedervorlage“. Die an einem bestimmten Tag nicht erreichten Haushalte werden durch solche substituiert, die an anderen Tagen nicht erreicht werden. Damit entfällt der sogenannte „Not-at-home-Bias“ weitgehend (nur Haushalte, die auch nach dem 12. Kontakt nicht angetroffen werden, werden angesteuert).

Das ITMS ist als EDV-Datei für EDV-gestützte zentrale Telefonumfragen konzipiert. Die Aussteuerung der Stichprobe nach den Zellen der Multistratifikationsmatrix erfolgt dabei unmittelbar und unabhängig durch das Steuerungsprogramm. Es ist somit gewährleistet, dass in jeder Zelle

die erforderliche Zahl von Interviews durchgeführt wird. Von diesem Programm wird auch, falls nötig, die notwendige Gleichverteilung der Interviews auf die Befragungstage gesteuert.

Um mögliche Einflüsse der Tageszeit auf Untersuchungsergebnisse von vornherein auszuschalten, wird die Stichprobe nach einem Verfahren der „dynamischen Repräsentativität“ bezüglich der Besetzung der Zellen des Multistratifikationstableaus optimiert, so dass sich für jedes Stundenintervall vorgebenproportionale Teilstichproben ergeben.

1.4 ADM-Standard des Random-Digit-Dialing

Die Auswahlgrundlage der Arbeitsgemeinschaft ADM-Telefonstichproben enthält eingetragene und generierte Nummern. Generierte Nummern werden erzeugt, indem die letzten beiden Stellen aller eingetragenen Rufnummern jeweils durch die Ziffernfolge 00 bis 99 ersetzt werden. Entstehende Doubletten werden bereinigt. Dieses Verfahren nennt man „Random Last Two Digits – RL2D“. Es garantiert, dass auch die Telefonanschlüsse in der Auswahlgrundlage enthalten sind, die nicht in Verzeichnisse eingetragen sind.

Die Auswahlgrundlage enthält lediglich (eingetragene oder generierte) Telefonnummern, aber keine Adressen, da diese für eine anonyme Befragung unerheblich sind. Aus den Einträgen werden jedoch Kennzeichen für die Art des Eintrags (geschäftlich/privat) sowie die Gemeindekennziffer übernommen.

Sofern die Nummer nicht eingetragen ist, wird ihr eine Gemeinde zugeordnet, und zwar abhängig davon, in welchen Gemeinden sich die anderen eingetragenen Rufnummern desselben Nummernstammes (= um die letzten zwei Stellen reduzierte Telefonnummer) befinden; Kennzeichen für die Art des Anschlusses können hierbei natürlich nicht vergeben werden.

Die Nummern aller Teilnehmer, die auf keinen Fall in Zusammenhang mit einer Befragung angerufen werden wollen, werden in der Auswahlgrundlage gesperrt.

1.5 Haushalts- und Personenstichprobe (Umwandlung)

Das beschriebene Auswahlverfahren führt zu einer *haushaltsrepräsentativen* Stichprobe, wobei jeder Haushalt die gleiche Chance hat, in die Auswahl zu kommen.

In jedem der ausgewählten Haushalte wird durch ein Zufallsverfahren (Schwedenschlüssel) *eine* Person als Zielperson ausgewählt. Dies gilt unabhängig davon, wie viele zur Grundgesamtheit gehörende Personen in dem betreffenden Haushalt leben. Die Chancen für die in Privathaushalten lebenden Personen der Grundgesamtheit, als Befragungsperson der Stichprobe ausgewählt zu werden, sind demnach umgekehrt proportional zur Zahl der zur Grundgesamtheit gehörenden Personen in ihren Haushalten.

Um eine repräsentative Personenstichprobe zu erhalten, wird die erstellte Stichprobe mathematisch im Nachhinein so umgeformt, dass jede Person der Grundgesamtheit stichprobentheoretisch die gleiche Auswahlchance erhält. Dabei wird die geringere Auswahlchance von Personen in Mehrpersonen-Haushalten durch entsprechende Gewichtung korrigiert. Durch diesen Umwandlungsschritt im Rahmen des Gewichtungsverfahrens entsteht eine personenrepräsentative Stichprobe.

2 Überarbeitung des CATI-Fragebogens nach dem Pretest

Die Befragung wurde mit der Methode des computer-unterstützten telefonischen Interviews (CATI) durchgeführt. Der CATI-Fragebogen wurde im Rahmen des Pretests erprobt, worüber ein gesonderter Bericht erstellt wurde (siehe Pretestbericht vom Mai 2003). Nach Durchführung des Pretests wurden zahlreiche Änderungen und Kürzungen im Fragebogen vorgenommen.

2.1 Fragebogenlänge: Reduktion durch Kürzung des Fragebogens

Im Pretestbericht musste eindringlich auf die deutlich erhöhte Interviewlänge von 35 Minuten hingewiesen werden. Diese hatte eine dramatische Steigerung der Abbruchquote auf 23% (Pretest 1992: 10%) zur Folge. In der Hauptstudie gelang es, die Netto-Interviewlänge auf durchschnittlich 31 Minuten zu senken. Das ist zwar immer noch 1 Minute über der vertraglich vereinbarten Länge von 31 Minuten, aber dennoch ein bedeutender Erfolg. Die Abbruchquote der Interviews bezogen auf die realisierten Interviews sank in der Hauptstudie gegenüber dem Pretest von 23% auf 10%. Damit ist diese vergleichbar mit derjenigen der Hauptstudie 2002 (9%).

Es ist zu begrüßen, dass das Forscherteam sich zur den Kürzungen entschließen konnte. Damit wurde sichergestellt, dass die befragten Populationen 2002 und 2003 nicht gegeneinander verzerrt wurden. Die überproportionale Abbruchneigung von Personen mit politisch eher „rechts“ einzustufenden Einstellungen, die sich im Pretest 2003 besonders drastisch gezeigt hatte, konnte in der Hauptstudie 2003 auf ein vergleichbares Maß wie in der Hauptstudie 2002 eingegrenzt werden. Durch die Besonderheiten der Bildungsgewichtung beider Datensätze auf Basis des Mikrozensus, die Personen mit einfachen Bildungsabschlüssen deutlich höher gewichtet und Personen mit höheren Abschlüssen deutlich niedriger, wird dieser Ausgleich weiter verstärkt, da Personen mit höheren Bildungsabschlüssen im Gegensatz zu solchen mit einfachen Abschlüssen vermehrt zu eher links-liberalen Positionen neigen.

2.2 „Fragebogendramatik“: Entspannung des Interviews

Im Pretestbericht musste neben der Interviewlänge auch die **Fragebogendramatik** problematisiert werden. Die Interviewlogik der Pretestinterviews lief auf eine ständige Spannungserhöhung hinaus und zwar im **negativen** Sinne. Die Interviewten wurden stetig stärker psychologisch belastet. Und zwar nahm im Laufe des Interviews sowohl der Grad des verbindlichen Zugriffs der Fragen auf die moralische Selbstschreibung der Befragten zu als auch die Ballung negativer Statements. Ziel der Hauptstudie war daher (in Korrespondenz mit der Kürzungsabsicht),

- die Streichung besonders belastender Fragen,
- eine Auflockerung des Bogens durch Einschub „harmloserer“ Fragen zwischen die schwierigeren,
- sowie die Umformulierung einer Reihe negativ formulierter Items in positiv formulierte.

Streichung und positive Umformulierung: Sinnvoll, weil psychisch entlastend, war die Streichung eines besonders verhaltensnahen Blocks von Aussagen, mit Aussagen wie „Wenn mich ein Obdachloser anspricht, reagiere ich abweisend.“ oder „Ich behandle Muslime abweisender als andere.“. Ansonsten war eine Reihe von Streichungen von Items in der Liste „Werte“ vor allem längenmäßig entlastend. Das betrifft auch besonders die Liste „Desintegration“, die bis auf 2 anders verbrachte Items ganz gestrichen wurde. Außerdem war die Änderung einer Reihe der auf den Islam bzw. auf Moslems bezogenen Items von negativen in positive Formulierungen für die Interviewgestaltung hilfreich. U.E. kann damit auch dem möglichen Vorwurf begegnet werden, das GMFS-Forscherteam habe die Einstellungen der Befragten zum Islam/zu Moslems nur auf negativ getönte Weise erfasst, wo doch diese Einstellungen höchstwahrscheinlich von eher ambivalenten Gefühlen bestimmt werden.

Umstellungen: Die Umstellungen erzeugten insgesamt einen akzeptablen Rhythmus von „schwierigen“ und „leichteren“ Fragekomplexen. Die unkomplizierte, so genannte Bildungsitem-Liste wurde weiter nach vorne gezogen, desgleichen die Werteliste von ganz hinten weit nach vorn zwischen den Islam/Moslem-Block und den Ausländerblock, die besonders schwierig sind. Zwischen den Ausländerblock und den ebenfalls schwierigen Bevölkerungsgruppenblock (inkl. Rassismus) wurde ein leichter Block mit Politikfragen (Demokratieentleerung 1) gesetzt. Zwischen den Bevölkerungsgruppenblock (inkl. Rassismus) und den schwierigen Gewaltblock wurde ein weiterer leichter Politikblock (Demokratieentleerung 2) platziert. Zwischen den Gewaltblock und die wieder schwierige Liste über Juden wurde schließlich die leichtere Liste Machtlosigkeit und Gerechtigkeit eingefügt.

Insgesamt wurde somit durch Streichung, Positivierung und Auflockerung erfolgreich die „negative Dynamik“ aus dem Interview genommen. Diese Maßnahmen können auch in Zukunft sicherstellen,

- dass Personen mit eher menschenfeindlichen Einstellungen (soweit wie mit Umfragen des vorliegenden Typs irgend möglich) auch **wirklich erfasst** werden;
- dass diese Erfassung über alle Wellen des GMF-Surveys auch **gleichermaßen** erfolgt und somit **echte Trends bzw. Konstanzen** abbildet und nicht Artefakte unterschiedlicher Mobilisierung von Meinungsträgern.

2.3 Filterführung

Die Filterung erfolgte weitgehend so wie in der ersten Welle 2002. Ausländerbezogene Einstellungen wurden Ausländern nicht vorgelegt, jüdenbezogene nicht Juden. Aussiedler beurteilten bei den ausländerbezogenen Fragen in Q43B Türken, die anderen Ausländer beurteilten in Q043C Aussiedler. Beide letztere Blöcke erhielten extra Einleitungstexte. Diesmal wurden außerdem konsequent den Moslems keine moslembezogenen Fragen gestellt.

Das Geschlecht musste aus filtertechnischen Gründen vor der Werteliste abgefragt werden, weil diese nach Männern und Frauen getrennte Formulierungen enthielt.

3. Interviewerschulung

Wie im Pretest wurde auch in der Hauptstudie eine intensive Interviewerschulung betrieben. Das bedeutet, dass die Interviewer über die Studie informiert und für das schwierige Thema sensibilisiert und motiviert wurden. Im Anschluss erhielten sie die Möglichkeit, den Fragebogen am Computer gründlich durchzuarbeiten. Dabei konnten Fragen gestellt werden und Unklarheiten beseitigt werden. Viele Interviewer waren allerdings schon an den vorhergehenden Befragungen des Projekts beteiligt (Querschnitt 2002/Längsschnitt 2003) und von daher bereits erfahren und sensibilisiert. Die Zielpersonenbefragung wurde von Telefonstudios in München und Parchim durchgeführt.

4. Durchführung der Befragung

4.1 Interviewdauer

Insgesamt wurden 3.000 Interviews in der Zeit vom 21.5. bis zum 29.6. realisiert. Die vereinbarte Netto-Interviewdauer (ohne Kontaktphase) betrug 30 Minuten. Die Befragungsdauer von der ersten bis zur letzten Frage des Bielefelder Bogens betrug 31,3 Minuten. Interessant ist, dass die Spannbreite zwischen kürzestem und längstem Interview sehr groß ist. Das kürzeste Interview dauerte lediglich 18 Minuten, wogegen das längste 114 Minuten dauerte. Das zeigt, wie unterschiedlich intensiv sich die Zielpersonen mit dem Fragebogen auseinandersetzen. Auch in der Hauptstudie konnten wir beobachten, wie schwer es vielen Befragten fiel, im Rahmen der Skalenausprägungen auf die schwierigen Statements zu antworten. Die Konsequenz war, dass die Befragten auswichen, versuchten, sich für ihre Antwort zu rechtfertigen, und dadurch die Befragungsdauer verlängerten.

4.2 Abbrüche von Interviews

Wir registrierten 302 **abgebrochene Interviews**, was einer Abbruchquote von 10% der realisierten Interviews entspricht. Diese Abbruchquote liegt in der Nähe der letzten Hauptstudie. Allerdings liegt sie nach wie vor klar über der üblichen Abbruchquote, was deutlich macht, dass auch 2003 anfangs befragungsbereite Zielpersonen überdurchschnittlich oft das Interview abbrechen. Je länger die Befragung dauert, desto mehr Zielpersonen beenden das Interview vorzeitig. Dies ist ein eindeutiges Indiz dafür, dass die Länge und Schwierigkeit des Interviews zu Ermüdungseffekten führen.

4.3 Stichprobenausschöpfung

Die 3.000 durchgeführten Interviews bilden die sog. „Nettostichprobe“. Diese ist der erfolgreich bearbeitete Teil der erheblich größeren Ausgangsstichprobe, auch „Bruttostichprobe“ genannt. Die oben in Kapitel 1 beschriebenen Verfahren der Stichprobenbildung beziehen sich auf die Bruttostichprobe. Im Zuge der Feldarbeit kommt es dann zu Ausfällen verschiedener Art. Der Erfolg der Feldarbeit dokumentiert sich in der Relation von erfolgreich durchgeführten Interviews zu (bereinigtem) Brutto, der sog. „Stichprobenausschöpfung“.

Die folgende Übersicht **„Rücklaufmeldung“** zeigt den Bruttoansatz der Stichprobe, die verschiedenen Arten von Ausfällen und die Berechnung der Stichprobenausschöpfung.² Wiedergegeben ist das Ergebnis des letzten Kontaktversuchs mit dem jeweiligen Haushalt bzw. der jeweiligen Zielperson.

Die Ausgangsstichprobe („Bruttoansatz“) umfasste 13.450 Telefonnummern. Nach Abschluss der Feldarbeit sind hiervon über 10.450 Ausfälle. Die Ausfallgründe sind in der Rücklaufmeldung nach drei Oberkategorien gruppiert, die in statistischer Hinsicht unterschiedlich zu werten sind:

(a) Nicht verwendbare Telefonnummern

Über den RLD-Ansatz der Stichprobenbildung werden teilweise auch Telefonnummern generiert, die real gar nicht existieren oder bei denen es sich um Fax- oder Modem-Anschlüsse handelt. Dies trifft für mehr als 5.663 der eingesetzten Telefonnummern des Bruttoansatzes zu. Es verbleiben 7.787 Telefonnummern, die wir als das „verwendbare Brutto“ bezeichnen.

(b) Neutrale Ausfälle

Unter diesem Oberbegriff fassen wir Fälle zusammen, bei denen es zwar eine verwendbare Telefonnummer gibt, der Haushalt aber nicht zur befragbaren Stichprobe entsprechend der definierten Grundgesamtheit gehört. Wir ordnen hier zwei Arten von Ausfallgründen zu:

² Die Ausfallgründe sind teilweise zusammengefasst. Die in Klammern genannten Code-Nummern zeigen die Ausfallkategorien an, wie sie im Bruttoband dokumentiert sind.

- Telefonnummern, bei denen es sich nicht um einen Privathaushalt handelt und der Computer dies vorab nicht erkannte. Dies ist bspw. möglich, wenn eine Person als Selbstständiger zuhause ein Versicherungsbüro betreibt (Codes 30, 37).
- Fälle, in denen der Haushalt bzw. die Zielperson aus sprachlichen Gründen (Code 70) oder aus mental-gesundheitlichen Gründen (Code 73) zu einem Interview nicht in der Lage ist.
- Zusammengenommen wurden 914 Telefonnummern als „neutrale Ausfälle“ gewertet. Das verwendbare Brutto, vermindert um die neutralen Ausfälle, ergibt das „bereinigte Brutto“.

(c) Ausfälle im bereinigten Brutto

Es verbleibt als „bereinigtes Brutto“ eine Zahl von 7.126 Telefonnummern. Im Laufe der Feldarbeit wurden hiervon 4.126 Telefonnummern zu einem Ausfall. Wir können vier Kategorien von Ausfällen im bereinigten Brutto unterscheiden.

- (1) Im Rahmen der Feldarbeit telefonisch kein Kontakt herzustellen
Gründe dafür können sein, dass der Teilnehmer nicht abnimmt, der Telefonanschluss besetzt ist oder ein Anrufbeantworter eingeschaltet ist. Weitere Kontaktversuche über einen längeren Zeitraum hätten möglicherweise diese Quote noch leicht reduziert, doch bestand hierfür im Rahmen der begrenzten Feldzeit kein Spielraum.
- (2) Kontaktperson verweigert Auskunft oder ist momentan zu beschäftigt
Diese Ausfallgründe sind quantitativ die stärkste Barriere zur Realisierung von Interviews. In all diesen Fällen wird die Kontaktsituation von Seiten des Haushalts sofort beendet, ehe überhaupt geklärt werden kann, wie viele Personen hier leben und welche davon nach dem gewählten Zufallsverfahren („Schwedenschlüssel“) als Zielperson für das Interview zu bestimmen wäre.
- (3) Zielperson grundsätzlich oder momentan nicht zum Interview bereit
Dies kann verschiedene situative Gründe haben (Codes 21, 25, 72) oder eine grundsätzliche Ablehnung gegenüber solchen Befragungen zum Ausdruck bringen (Codes 62, 63, 64). Das Umfrageinstitut muss die Verweigerung akzeptieren, ja sogar auf die Freiwilligkeit der Teilnahme selbst hinweisen.
- (4) Abbrüche des Interviews
In 302 Fällen wurde das Interview begonnen, dann aber vor Abschluss abgebrochen. Die Abbruchquote ist mit 4,2% (bezogen auf das bereinigte Brutto von 7.126, 10% bezogen auf das Netto von 3.000) deutlich höher als in den meisten – meist kürzeren – Telefonbefragungen. Hierauf wurde in Kapitel 4.3 bereits eingegangen.

Um 3.000 Interviews zu realisieren, wurde ein Bruttoansatz von 13.450 bzw. eine bereinigte Bruttostichprobe von 7.787 Telefonhaushalten benötigt. Die Stichprobenausschöpfung beträgt demnach 42,1%. Sie liegt damit etwas über der vorjährigen von 40,6%.

Dieser Wert ist niedriger, als er aus anderen Telefonumfragen meist berichtet wird. Grund ist vor allem die für ein Telefoninterview relativ lange Interviewdauer. In der Kontaktsituation wird der

Interviewer häufig gefragt, wie lange das Interview denn dauern wird. Im vorliegenden Projekt kann der Interviewer diese Hürde nicht mit dem Hinweis überwinden, dass es „ganz schnell gehen“ würde. Man muss vielmehr wahrheitsgemäß die vermutliche Dauer von 30 Minuten nennen. Unter diesen Umständen sind viele Befragungspersonen zu einem Interview nicht bereit.

Die Stichprobenausschöpfung gilt in der Umfrageforschung häufig als Qualitätsindikator. Tatsächlich sagt sie für sich genommen allerdings wenig über die Qualität der Stichprobe aus. Ob die Struktur der realisierten Nettostichprobe verzerrt ist, hängt nicht von der Anzahl, sondern der Struktur der Ausfälle ab. Nur wenn die Ausfälle sich in untersuchungsrelevanten Merkmalen systematisch von den Teilnehmern an der Befragung unterscheiden, wird die Nettostichprobe relevante Verzerrungen aufweisen. Auf diese Fragen wird in Abschnitt 5.3 „Gewichtung der Nettostichprobe“ weiter eingegangen.

	<u>Rücklaufmeldung</u>		
Projekt:	GMF-Survey 2003/Querschnitt (08.10.31817)		
Projektleiter:	Dr. Gensicke		
Feldzeit:	27.05.03 – 24.06.03		
Zielgruppe:	Bevölkerung ab 16 Jahren		
Befragungsgebiet:	Bundesrepublik Deutschland		
Auswahlverfahren:	ADM-Sample		
Telefon-Studio:	München/Parchim		
Ergebnis des letzten Kontaktversuchs			
Bruttoansatz		13.450	100,0%
	Falsche Telefonnummer (31, 88, 89)		35,6%
	Fax/Modem (35)		4,1%
	Doppeladresse (39)		0,1%
	Falsche Regionalzelle (42)		-
	Regionalzelle nicht zu ermitteln (43, 65)		-
	Maximale Kontakzahl (>12) erreicht (15, 16, 17, 18)		2,3%
	Gesamt: Nicht verwendbare Tel.-Nr.		42,1%
Verwendbares Brutto		7.787	100,0%
	Zelle voll (49)		-
	Kein Privathaushalt (30, 37)		4,5%
	Keine Pers. i. d. Zielgruppe (41)		-
	im HH keine Verständigung (70)		2,6%
	mit ZP keine Verständigung (71)		0,6%
	ZP nicht in der Lage (73)		0,9%
Neutrale Ausfälle gesamt		661	8,5%
Bereinigtes Brutto		7.126	100,0%
	Teilnehmer nimmt nicht ab (10, 82, 84)		7,4%
	Besetzt (11, 85, 87)		0,2%
	priv. Anrufbeantworter (19)		6,9%
	Abbruch Interview (20, 26, 90)		4,2%
	ZP vorübergehend krank (21)		0,0%
	KP beschäftigt (24)		2,4%
	ZP beschäftigt (25)		2,4%
	KP verweigert Auskunft (51, 53, 54)		25,8%
	ZP verweigert Auskunft (62, 63, 64)		5,8%
	Kein Termin möglich (72)		2,8%
	Sonstige Ausfälle		-
Ausfälle gesamt		4.126	57,9%
Durchgeführte Interviews (Netto-Stichprobe)		3.000	42,1%

5. Erfassung von Ausländern

Von 3.000 realisierten Interviews wurden nur 63 mit Befragten durchgeführt, die keine deutsche Staatsbürgerschaft haben, das sind 2,1% der Stichprobe. Personen mit Migrationshintergrund werden dagegen wesentlich breiter erfasst. Personen ohne deutsche Staatsbürgerschaft werden in telefonischen Umfragen, insbesondere mit anspruchsvollen sozial- oder politikwissenschaftlichen Themen regelmäßig unterproportional abgebildet. Z.B. erfasste der ebenfalls von uns durchgeführte Freiwilligensurvey 1999 trotz seiner ca. 15.000 Befragten nur einen Ausländeranteil von 2.6%. Es besteht für diesen Personenkreis eine ganze Reihe von Barrieren beim Zugang zu Telefonumfragen, die sich wechselseitig verstärken können.

- Es gibt sprachliche Schwierigkeiten.
- Das Interesse und das Verständnis für die Themen ist eher niedrig, dazu trägt oft ein niedriger Bildungsstand bei.
- Die telefonische Erreichbarkeit ist verringert.
- Ausländer gehören überproportional den Unterschichten an, die weniger befragungsbereit sind.

Die wesentlichen Faktoren sind dabei wohl Barrieren aufgrund der Sprache und des Umfrage-themas. Im Moment gibt es noch keine hinreichende Möglichkeit, diese Probleme für Telefonbe-fragungen, insbesondere sozial- oder politikwissenschaftlichen Art, zu lösen. Dazu kommt beim GMF-Survey die besonders anspruchsvolle Thematik, z.B. im Vergleich zum Freiwilligensurvey. Und obwohl Ausländer bei den ausländerbezogenen Fragen herausgefiltert werden, sind sie oft bei gruppenbezogenen Fragen latent die Betroffenen, so dass ihnen die Beantwortung der Fra-gen unangenehmer sein kann als Deutschen.

6. Die Bruttodaten

Die Daten der 3.000 realisierten Interviews wurden der Forschungsgruppe als „Nettodaten“ übergeben. Zusätzlich erhält die Forschungsgruppe – als gewünschte Sonderleistung – auch die Bruttostichprobe in Form eines Datensatzes, genannt „Bruttodaten“.

Die Bruttodaten dokumentieren die im vorigen Abschnitt dargestellte Bruttostichprobe nach me-thodischen Gesichtspunkten. Die Zahlen der Übersicht „Rücklaufmeldung“ sind unmittelbar in den Bruttodaten nachvollziehbar. Die Fallzahl beträgt $n = 13.450$, entspricht also dem Brutto-ansatz der Stichprobe.

Grundlage für die Erstellung des Bruttodatensatzes ist der sog. „Account File“, in dem während der laufenden Feldarbeit jeder Kontaktversuch bei jeder eingesetzten Telefonnummer automa-tisch dokumentiert wird. Im vorliegenden Projekt ist für die 13.450 eingesetzten Telefonnum-mern etwa das Dreifache an Kontakten dokumentiert.

Die Bruttodaten enthalten folgende Informationen:

Idadr ID-Nummer der jeweiligen Telefonnummer im ursprünglichen Adressenpool, der in ITMS für das Projekt bereitgestellt wurde (vgl. zur Erläuterung Kapitel 1). Da diese nicht

alle zum Einsatz kommen, sind die Nummern in der Abfolge nicht fortlaufend. Die Variable Idadr ist für die realisierten Interviews auch in den Nettodaten überspielt. Über diese Variable können beide Datensätze bei Bedarf verknüpft werden.

wieoft Zahl der Kontaktversuche bei jeder einzelnen Telefonnummer

ausfall Ergebnis des letzten Kontakts bzw. Kontaktversuchs

Die folgende Übersicht zeigt die vollständige Liste der Ausfallcodes. Die Verteilung ist im vorigen Abschnitt in der Übersicht „Rücklaufmeldung“ dargestellt.

Übersicht: Codes für das Kontaktergebnis (Variable „ausfall“)

Codes

2	'vollständiges Interview'
10	'Teilnehmer nimmt nicht ab'
11	'Anschluss besetzt'
15	'Maximale Kontaktzahl erreicht'
16	'Max. Anzahl – nimmt nicht ab'
17	'Max. Anzahl – besetzt'
19	'privater Anrufbeantworter'
20	'Abbruch Interview, Fortsetzung möglich – Termin'
21	'ZP vorübergehend krank – T'
24	'KP momentan beschäftigt/nicht erreichbar – Termin'
25	'ZP momentan beschäftigt – T'
30	'Firmen-Anrufbeantworter'
31	'Falsche Telefonnummer'
35	'Fax/Modem'
37	'Kein Privathaushalt'
39	'Doppeladresse'
51	'KP: keine Auskunft: Projekt/Thema'
53	'KP: keine Auskunft: sonst. Grund'
62	'ZP: keine Auskunft: Projekt/Thema'
63	'ZP: keine Auskunft: sonst. Grund'
70	'HH: keine Verständigung möglich'
71	„ZP: keine Verständigung möglich'
72	'Kein Termin möglich zu I-Zeiten'
73	'ZP geistig/körperlich n. i. d. Lage'
82	'AD: während RVA aufgelegt'
84	'AD: Teiln. nimmt nicht ab'
85	'AD: Anschluss besetzt'
87	'AD: gassenbesetzt'
88	'AD: Falsche Telefonnummer'
89	'AD: Falsche Telefonnummer (9)'
90	'Abbruch Interview mit Schreiben Daten'
98	'Externer nicht (mehr) in Leitung'

Erläuterungen:

KP heißt Kontaktperson und ist die Person, die zuerst am Telefon ist.

ZP heißt Zielperson und ist die Person, an die der Fragebogen sich richtet, hier über den Schwedenschlüssel ausgewählt die nächstälteste im Haushalt aller Personen ab 16 Jahren).

AD heißt Autodialer, ist unser Telefoncomputer.

RVA heißt „recorded voice announcement“ und ist die Sprachansage, die der Dialer von sich gibt, wenn er gerade eine Verbindung hergestellt hat, aber kein Interviewer frei ist, das Gespräch zu übernehmen.

Falsche Telefonnummer (9) ist eine spezielle Art von Info, die der Dialer über die ISDN-Leitung erhält, die für die Studie als „ungültige Telefonnummer“ interpretiert wird. Im Unterschied dazu gibt es noch die ohne (9) über der AD erkannte „ungültige Telefonnummer“, die „kein Anschluss unter dieser Nummer ...“ entspricht.

Abbruch Interview ist das Ergebnis eines Interviews, bei dem der Befragte mitten im Interview keine Lust mehr hat und verweigert.

Externer nicht (mehr) in der Leitung ist der Ausfallgrund, den der Interviewer eingibt, wenn er vom Dialer ein Gespräch auf den Kopfhörer durchgestellt bekommt, aber der Befragte vorher aufgelegt hat (es handelt sich da beim Durchstellen um 1–2 Sekunden).

Gassenbesetzt: Vermittlungsknoten der Telekom ist überlastet, es ist ein Besetzt-Zeichen zu hören.

Doppeladresse: ZP gibt an, schon einmal zu diesem Thema befragt worden zu sein.

Falsche Telefon-Nummer: Telefonnummer existiert nicht, was durch das ITMS-System leicht passieren kann (siehe Anhang).

7. Gewichtung der Nettostichprobe

Wie oben dargestellt hängt die Qualität der Stichprobe im Sinne einer realitätsgetreuen Abbildung der Grundgesamtheit davon ab, ob sich Ausfälle und realisierte Interviews nach untersuchungsrelevanten Merkmalen unterscheiden. Dies kann allerdings nicht direkt untersucht werden, da über die Ausfälle keine inhaltlichen Informationen vorliegen. Die Qualität der realisierten Nettostichprobe kann daher nur über einen Vergleich der Stichprobenstruktur mit verfügbaren Referenzstatistiken eingeschätzt werden.

NFO Infratest nimmt diesen Strukturvergleich im Zuge einer standardmäßig durchgeführten Gewichtung der Stichprobe vor. Referenzstatistik ist im Wesentlichen die aktuellste verfügbare *Bevölkerungsfortschreibung des Statistischen Bundesamts*. Damit können folgende Strukturverteilungen abgeglichen werden:

- (1) Regionale Verteilungen:
Verteilung über Bundesländer sowie – jeweils innerhalb eines Bundeslandes – über Gemeindegrößenklassen nach BIK-Typ.
- (2) Demographische Verteilungen:
Verteilung nach 7 Altersgruppen, getrennt für Männer und Frauen und bezogen auf die demographische Struktur innerhalb jedes Bundeslandes.

Bei Projekten mit besonderem Qualitätsanspruch können weitere Strukturmerkmale hinzugenommen werden, sofern entsprechende Referenzstatistiken verfügbar sind. Im vorliegenden Projekt schien es wichtig, die Struktur der Stichprobe auch nach Bildungsniveaus zu kontrollieren, da die meisten der untersuchten Einstellungen mit dem Bildungsniveau der Befragten variieren.

Als Gewichtungsmerkmal dient die *Schulbildung*, als Referenzstatistik der *Mikrozensus*. Es werden drei Bildungsniveaus definiert, die im Datensatz der Studie ebenso wie im Mikrozensus in gleicher Weise abgrenzbar sind.³ Der Strukturabgleich erfolgt innerhalb von vier Teilgruppen, nämlich Ost- und Westdeutschland und hier jeweils nach zwei Altersgruppen (bis zu 50/über 50 Jahre).⁴

Das Gewichtungsmodell, wie es für diese Studie verwendet wurde, umfasst demnach vier aufeinanderfolgende Schritte, wobei jeder Gewichtungsschritt auf dem vorangehenden aufbaut:

³ Ein gewisses Problem besteht darin, dass der Mikrozensus den Abschluss einer Polytechnischen Oberschule (DDR) nicht nach dem Abschluss der 8. Klasse und der 10. Klasse differenziert. Der Abschluss der POS muss daher durchgängig dem „mittleren Niveau“ zugeordnet werden, obwohl der Abschluss nach der 8. Klasse richtigerweise dem „unteren Niveau“ (Hauptschulabschluss) zuzuordnen wäre. Die quantitative Bedeutung dieses Problems ist allerdings begrenzt.

⁴ Die Beantwortung der Fragen zur Schulbildung ist im Mikrozensus für Personen bis zu 50 Jahren obligatorisch, für Personen über 50 Jahren dagegen freiwillig. In der Bildungsgewichtung wird dies dadurch berücksichtigt, dass alle Personen ohne Angaben zum Schulabschluss (einschließlich Schüler) in diesen Gewichtungsschritt nicht einbezogen werden.

- (1) Regionalgewichtung
- (2) Umwandlung von der haushalts- in die personenproportionale Stichprobe⁵
- (3) Demografiegewichtung
- (4) Bildungsgewichtung

Im Ergebnis stimmen die Merkmalsverteilungen in der Stichprobe mit der Bevölkerungsstruktur lt. amtlicher Statistik überein. Das Gewichtungsverfahren ist 2003 (außer der Aktualisierung der zu Grunde liegenden Referenzdaten der öffentlichen Statistik) dasselbe wie 2002.

Der Gewichtungsfaktor, der bei Auswertungen für Deutschland gesamt zu verwenden ist, ist als Variable „gewbpp“ im Datensatz vorhanden. Sofern Auswertungen für Westdeutschland und Ostdeutschland getrennt vorgenommen werden, ist besser die Variable „gewbp“ zu verwenden, die die disproportionale Anlage der Stichprobe beibehält. Der Faktor „gewbpp“ dagegen nimmt eine Proportionalisierung des Ost-West-Verhältnisses entsprechend der Verteilung in der Grundgesamtheit vor.

⁵ Vgl. dazu oben Kapitel 1.5.