

**Gruppenbezogene Menschenfeindlichkeit**

**Methodenbericht  
zur repräsentativen Befragung  
Mai/Juni 2002**

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

Infratest Sozialforschung

Studienleitung:  
Bernhard von Rosenblatt  
Dr. Thomas Gensicke  
Sabine Ohder

München, 29. Juli 2002  
Mb10-22674

## Inhaltsverzeichnis

<b>1.</b>	<b>Methodische Anlage telefonischer-Bevölkerungsumfragen von NFO Infratest</b>	<b>4</b>
1.1	Überblick	4
1.2	Grundgesamtheit und Hochrechnungsrahmen	4
1.3	Stichprobenanlage und Stichprobensteuerung	5
1.4	ADM-Standard des Random-Digit-Dialling	6
1.5	Haushalts- und Personenstichprobe (Umwandlung)	6
<b>2.</b>	<b>Überarbeitung des CATI-Fragebogens nach dem Pretest</b>	<b>7</b>
2.1	Bearbeiten von Fragetexten	7
2.2	Platzänderung von Variablen	8
2.3	Einfügen zusätzlicher Variablen / Löschen von Variablen	8
2.4	Änderung von Variablennamen	9
2.5	Filter	9
2.6	Behandlung von "keine Angabe" in Filterfragen	9
2.7	Skalenänderungen	10
2.8	Eingangstext / Ausstiegstext	10
<b>3.</b>	<b>Interviewerschulung und Durchführung der Interviewer-Interviews</b>	<b>11</b>
<b>4.</b>	<b>Durchführung der Befragung</b>	<b>12</b>
4.1	Zeitliche Verteilung der realisierten Interviews	12
4.2	Interviewdauer	14
4.3	Unterbrechungen und Abbrüche von Interviews	15
4.4	Wiederbefragungsbereitschaft	16
<b>5.</b>	<b>Von der Brutto- zur Nettostichprobe: Stichprobenausschöpfung und Stichprobengewichtung</b>	<b>17</b>
5.1	Bruttostichprobe und Stichprobenausschöpfung	17
5.2	Die Bruttodaten	21
5.3	Gewichtung der Nettostichprobe	24
<b>6.</b>	<b>Datenprüfung und -aufbereitung</b>	<b>27</b>

<b>7.</b>	<b>Schlussfolgerungen für die zweite Welle</b>	<b>28</b>
7.1	Interviewdauer und Pufferfragen	28
7.2	Inhaltliche Änderungen	28

## 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 2001 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.<sup>1</sup>

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 3000 befragten Personen. Die Stichprobenanlage ist disproportional im Sinne eines überproportionalen Anteils der ostdeutschen Bevölkerung. Zu befragen waren 2000 Personen in den alten Bundesländern (einschließlich Berlin) und 1000 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 „deutschsprachige Personen ab 16 Jahren in Telefonhaushalten“.

---

<sup>1</sup> 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

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 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-Dialling 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 zeitlichen 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 vorneherein auszuschalten, wird die Stichprobe nach einem Verfahren der "dynamischen Repräsentativität" bezüglich der Besetzung der Zellen des Multistratifikationstableaus optimiert, so dass sich für jedes Stundenintervall vorgabenproportionale Teilstichproben ergeben.

### 1.4 ADM-Standard des Random-Digit-Dialling

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 Interviewens (CATI) durchgeführt. Der CATI-Fragebogen wurde im Rahmen des Pretests erprobt, worüber ein gesonderter Bericht erstellt wurde (siehe E-Mail vom 7.5.). Nach Durchführung des Pretests vom 24.4. bis 26.4.02 wurden zahlreiche Änderungen im Fragebogen vorgenommen. Ein Originalfragebogen (prt-Format) mit Filterführung wurde bereits übermittelt (siehe E-Mail vom 16.5.), ein Fragebogen im Word-Format mit Variablennamen und Filterhinweisen befindet sich im Anhang zu diesem Dokument. Im Folgenden wird kurz beschrieben, welche Änderungen des Fragebogens nach dem Pretest vorgenommen wurden.

### 2.1 Bearbeiten von Fragetexten

Im Rahmen der Überarbeitung des Fragebogens soll zunächst auf die Bearbeitung von Fragetexten eingegangen werden.

Eine doppelte Verneinung wurde durch eine einfachere Formulierung ersetzt, um die Beantwortung der Frage für Zielpersonen zu erleichtern: "Nur wer etwas für die Gesellschaft leistet, soll auch Anspruch auf Unterstützung haben," (*ui01w1*) statt "Wer nichts für die Gesellschaft leistet, soll auch keinen Anspruch auf Unterstützung haben".

Außerdem wurden Redundanzen beseitigt: "Von den Menschen in meinem privaten Umfeld erfahre ich Achtung und Respekt" (*st\_regr*) sowie "In meinem privaten Umfeld werde ich immer mit großem Respekt behandelt" ähnelten sich sehr, worauf hin auf das erste Statement verzichtet wurde und stattdessen das zweite in leicht abgeänderter Version im Fragebogen enthalten blieb "In meinen privaten Umfeld gibt es genügend Menschen, die mir Anerkennung entgegenbringen" (*sa01w1*).

Dasselbe gilt für die Statements "Mir fällt immer etwas ein, wenn ich in der Klemme stecke" (*se\_pole*) und "Ich kann mir immer selbst helfen, wenn ein Problem auftaucht" (*sw01w1*). Hier wurde das erste Statement entfernt.

Des Weiteren wurden Eingangstexte und Statements leicht verändert bzw. verkürzt, sodass die Fragestellung bis zum Beginn der Skala abgeschlossen war und nicht über die Skala hinausging (vgl. die Frage zu wirtschaftlichen Lage in der Bundesrepublik im Vergleich zur wirtschaftlichen Lage in der eigenen Gegend = *rd03w1*).

## 2.2 Platzänderungen von Variablen

Die Variable zur Beurteilung des Lebensstandards der Zielperson (*wa01w1*) und die Variable, wie viel sich die ZP leisten kann (*wa03w1*) waren ursprünglich hintereinander platziert. Sie wurden wunschgemäß getrennt, sodass *wa03w1* schließlich vier Fragen hinter der erstgenannten Variable ihren Platz fand.

Die Fragen nach der politischen Einstellung *pp01w1*, *po02w1*, *sg01w1* sowie die Frage danach, wie stolz man sei, Deutsche/r bzw. Europäer/in zu sein, waren ursprünglich hinter den Blöcken zur Fremdenfeindlichkeit, dem Antisemitismus und Rassismus platziert. Im Pretest wurde jedoch nur eine geringe Anzahl von Personen ermittelt, die angaben, stolz Deutsche/r zu sein bzw. die sich in der Parteienlandschaft rechts einordnen wollten.

Infratest befürchtete Ausstrahlungseffekte und verschob die Fragen zum Nationalstolz, der politischen Einstellung sowie zur politischen Beteiligung (*ii04w1*, *ii05w1*, *pp01w1*, *po01w1*, *po02w1*, *sg01w1*) an den Anfang des Fragebogens. Tatsächlich stieg die Anzahl derer, die "eher stolz" bzw. "sehr stolz" darauf sind, Deutsche/r zu sein um knapp 10 Prozentpunkte im Vergleich zum Pretest. Die Anzahl derer, die "eher stolz" bzw. "sehr stolz" darauf sind, Europäer/in zu sein, stieg sogar um 15 Prozentpunkte.

Auch bei der politischen Einstellung (*po01w1*) ergaben sich Änderungen. Während sich im Pretest nur knapp 7 % der Befragten als "eher rechts" oder "rechts" einstufen, waren es in der Hauptstudie immerhin 11%. Jedoch ist noch immer eine Linksschiefe zu erkennen, wenn auch in einer etwas abgeschwächteren Form als im Pretest.

Es bleibt festzuhalten, dass es durchaus Sinn macht, Einstellungsfragen, bei denen es leicht zu Ausstrahlungseffekten kommt, an den Anfang eines Fragebogens zu stellen, um validere Ergebnisse zu erhalten.

## 2.3 Einfügen zusätzlicher Variablen / Löschen von Variablen

Zusätzlich wurden die Fragen *km01w1* ("Wie viele Ausländer leben in Ihrer Nachbarschaft?") wie auch die Frage *km02w1* ("Wie viele Ausländer gibt es an Ihrem Arbeitsplatz?") aufgenommen. An die zweite Frage schließt sich wiederum eine neue Frage *ka02w1* (Kontakt zu Ausländern am Arbeitsplatz) an.

Dafür wurde auf die Frage nach der Anzahl der Freunde/Bekannteten, die Homosexuelle sind (cho\_frie), verzichtet. Weiter wurden Redundanzen beseitigt, die weiter vorne bereits erwähnt wurden.

Im Datensatz enthalten ist auch eine Dummy-Variable Ost/West, mit der getrennt nach Ost/West ausgewertet werden kann.

## 2.4 Änderung von Variablennamen

Gemäß der Vorgaben aus Bielefeld wurden sämtliche Variablen von Infratest benannt. Nach dem Pretest verständigte sich die Forschergruppe auf eine andere Benennung sämtlicher Variablen. Die Umsetzung wurde von Infratest vorgenommen.

## 2.5 Filter

Wie im Pretest-Bericht angeregt, wurde die Frage, ob man auch Ausländer danach befragen sollte, wie sehr sie sich als Deutsche fühlten, zwischen Bielefeld und München diskutiert. Man kam zu dem Schluss, dass dies durchaus sinnvoll ist, woraufhin ein Ausländerfilter nur für die Fragen Q055 (Stolz auf dt. Geschichte, Demokratie, soziale Sicherheit in Deutschland) und Frage Q056, 1.Statement (Stolz, Deutsche/r zu sein) eingefügt wurde. Außerdem blieb der Filter bestehen, wonach nur Deutsche gefragt wurden, wie sehr sie sich als Westdeutsche bzw. Ostdeutsche fühlen.

Des Weiteren wurden die Fragen zur Fremdenfeindlichkeit Q043A/B/C mit zwei Alternativblöcken für drei verschiedene Personengruppen formuliert. Frage Q043A ging nur an Deutsche, Q043B ging an alle Befragten mit anderer Staatsangehörigkeit außer an Polen, Rumänen und Bürger der ehemaligen Sowjetunion. Q043C wurde nur der letztgenannten Personengruppe gestellt.

## 2.6 Behandlung von "keine Angabe" in Filterfragen

Grundsätzlich werden Zielpersonen, die keine Angabe zu einer Frage geben, zu der nachfolgenden Filterfrage nicht befragt. Beispiel Staatsangehörigkeit: Wenn ein Befragter keine Angabe zu seiner Staatsangehörigkeit macht, kann er keiner Staatsangehörigkeit eindeutig

zugeordnet werden und wird zu allen folgenden Fragen, die nur an bestimmte Staatsangehörige gehen sollen, nicht befragt. Ebenso verhält es sich mit den Fragen zur Erwerbstätigkeit.<sup>2</sup>

## 2.7 Skalenänderungen

Die Skalenausprägung "sehr oft", "oft", "eher selten" oder "nie" wurde auf Wunsch der Forschergruppe ersetzt durch "häufig", "manchmal", "selten" bzw. "nie". Außerdem wurde bei der Skala "sehr viele", "eher viele", "eher wenige" "fast keine", "gar keine" die Ausprägung "fast keine" ersatzlos gestrichen.

Während im Pretest einige Ausprägungen der Variablen Schulbildung (zu01w1) zusammengefasst abgefragt wurden (Bsp. "keinen Schulabschluss", "polytechnische Oberschule nach der 8. Klasse" bildete eine Ausprägung) wurden sämtliche Ausprägungen einzeln als Antwortmöglichkeiten angeboten. Dies ermöglicht eine genauere Einordnung der Zielpersonen in die einzelnen Bildungsschichten, auch getrennt nach Ost/West.

## 2.8 Eingangstext / Ausstiegstext

Im anfangs gewählten Einstiegstext "...darin soll ermittelt werden, welche Meinungen es in der Bevölkerung zum Zusammenleben von Menschen gibt" vermuteten wir einen möglichen Grund für die geringe Teilnahmebereitschaft am Interview. Hintergrund war unsere Annahme, dass die Kontaktperson bzw. Zielperson Fragen über ihr Privatleben hinter der Formulierung vermutet. So wurde der Einstiegstext umgeändert in "...darin soll ermittelt werden, welche Meinungen es in der Bevölkerung zum sozialen Zusammenleben in Deutschland gibt". Diese abgemilderte Formulierung hat sicherlich dazu beigetragen, dass die Ausschöpfung gestiegen ist.

Der Ausstiegstext wurde dahingehend geändert, dass nicht zwei Reisen im Wert von je 2.000€, sondern zwei Reisen im Wert von je 1.000€ verlost wurden, worauf hin sich die Wiederbefragungsbereitschaft etwas senkte, was unter 4.4 behandelt wird.

---

<sup>2</sup> Diese Konvention gilt nicht für die Variablen ka02w1 und partw1. Hier macht es auch Sinn, diejenigen ZP zu befragen, die keine Angabe gemacht haben, da die eindeutige Zuordnung nicht in dem Maße eine Rolle spielt wie bei der Staatsangehörigkeit.

## 3. Interviewerschulung und Durchführung der Interviewer-Interviews

Wie im Pretest wurde auch in der Hauptstudie eine intensive Interviewerschulung betrieben. Das bedeutet, dass die Interviewer von einer Mitarbeiterin von Infratest Sozialforschung zunächst 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. Diese intensive Schulung gewährleistete einen sicheren Umgang der Interviewer mit dem Fragebogen.

Im Anschluss an die Schulung haben sich jeweils 54 Interviewer gegenseitig interviewt. Hintergrund des gegenseitigen Befragens ist zum einen die Frage, inwieweit die persönliche Einstellung des jeweiligen Interviewers das Antwortverhalten der von ihm befragten Zielpersonen beeinflusst. Andererseits stellte das Interviewen für die Interviewer ein zusätzliches Training mit dem Fragebogen dar. Die Ergebnisse liegen als Datensatz vor.

Die Zielpersonenbefragung wurde von Telefonstudios in München und Parchim durchgeführt. Die Auswahl zweier Studios dient einerseits der Qualität, da die unterschiedliche Interviewerstruktur der Studios (Alter, Geschlecht, Einstellungen) ausgeglichen wird, und andererseits bot dieses Vorgehen die Möglichkeit, dass sich die Interviewer gegenseitig telefonisch interviewen konnten.

Im Rahmen der Hauptstudie befragten sich 108 Interviewer gegenseitig, verteilt über drei Schulungen, die am 16.5., 23.5. und 11.6. stattfanden. Die Münchner Interviewer wurden persönlich von einer Mitarbeiterin von Infratest Sozialforschung eingewiesen, die Parchimer Interviewer wurden telefonisch eingewiesen.

In Tabelle 1 ist, in Klassen eingeteilt, zu sehen, wie viele Interviews von wie vielen Interviewern durchgeführt wurden. Durchschnittlich machte jeder Interviewer 27 Interviews, wobei manche Interviewer mehr Interviews realisieren konnten, andere weniger. Dies liegt daran, dass manche Interviewer für die komplette Zeit der Erhebung zur Verfügung standen, andere dagegen nur zeitweise bzw. erst zu einem späteren Zeitpunkt mit ihrer Mitarbeit an der Studie begannen.

Tabelle 1: Anzahl der durchgeführten Interviews pro Interviewer<sup>3</sup>

	Häufigkeit	Prozent	Kumulierte Prozente
1-10 Interviews	571	19,0	19,0
11-20 Interviews	687	22,9	41,9
21-30 Interviews	899	30,0	71,9
31-40 Interviews	456	15,2	87,1
41-50 Interviews	319	10,6	97,7
51-64 Interviews	69	2,3	100,0
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Als Fazit lässt sich sagen, dass die Interviewerschulung sich bewährt hat und einen großen Beitrag zum Gelingen dieses Projektes geleistet hat.

## 4. Durchführung der Zielpersonen-Interviews

### 4.1 Zeitliche Verteilung der realisierten Interviews

Insgesamt wurden 3000 Interviews in der Zeit vom 21.5. bis 29.6. realisiert. Zwei nachfolgende Tabellen verdeutlichen den Verlauf über die Kalenderwochen sowie über die Wochentage.

In der Wochenverteilung (Tab. 2) macht sich die "schrittweise" Rekrutierung der Interviewer an der steigenden Anzahl von Interviews bemerkbar. In den ersten drei Wochen war erst ein Drittel der Interviewer geschult, entsprechend wenig Interviews wurden durchgeführt. In der 5. und 6. Woche, also zu einem Zeitpunkt, an dem alle Interviewer geschult waren und sich im Einsatz befanden, wurden die meisten Interviews gemacht.

<sup>3</sup> Irrtümlicherweise machten jedoch zwei Interviewer, die nicht die Schulung durchliefen, insgesamt drei Interviews. Es handelt sich um die Interviewer mit der Interviewernummer 98338 (zwei Interviews) und 97821 (ein Interview). Diese drei Interviews können zur Untersuchung von Interviewereinflüssen nicht verwendet werden. Alle sonstigen Interviews können den jeweiligen Interviewern und ihren Daten zugeordnet werden. Die Interviewer mit den Nummern 6753, 9272, 9328, 9434 und 9930 wurden bereits im Rahmen des Pretest geschult, weswegen Sie deren Daten im Interviewer-Pretest-Datensatz finden. Insgesamt nahmen also 113 Interviewer an der Hauptstudie teil.

Tabelle 2: Wochenverteilung der Interviews

	Häufigkeit	Prozent	Kumulierte Prozente
1.Woche	260	8,7	8,7
2.Woche	348	11,6	20,3
3.Woche	376	12,5	32,8
4.Woche	496	16,5	49,3
5.Woche	1002	33,4	82,7
6.Woche	518	17,3	<b>100,0</b>
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Aus Tabelle 3 ist ersichtlich, dass montags besonders viele Interviews geführt wurden, nämlich knapp 24%, während samstags und sonntags nur sehr wenige geführt wurden. Die restlichen Interviews verteilen sich relativ gleichmäßig über die anderen Wochentage.

Tabelle 3: Tagesverteilung der Interviews

	Häufigkeit	Prozent	Kumulierte Prozente
Montag	712	23,7	23,7
Dienstag	491	16,4	40,1
Mittwoch	508	16,9	57,0
Donnerstag	526	17,5	74,6
Freitag	582	19,4	94,0
Samstag	146	4,9	98,8
Sonntag	35	1,2	<b>100,0</b>
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Grundsätzlich wurde mit unseren Telefonstudios die Vereinbarung getroffen, ab 17 Uhr mit dem Telefonieren zu beginnen. Der Grund hierfür liegt in der besseren Erreichbarkeit bestimmter Personengruppen, besonders der Berufstätigen, die tagsüber nur schwer zu Hause erreichbar sind. Kontaktversuche und Interviews wurden bis 21 Uhr gemacht, von 21 Uhr bis 21.30 Uhr wurden lediglich Termine mit Zielpersonen für die Durchführung eines Interviews gemacht, da das Führen eines Interviews zu später Stunde als nicht zumutbar betrachtet wird.

Die nachfolgende Tabelle macht dies deutlich, wobei wir darauf hinweisen möchten, dass sich die Anfangszeiten auf das jeweils zuletzt geführte Teilinterview beziehen. Unser Programm überschreibt jede Anfangszeit eines Interviews, sollte das Interview unterbrochen werden und zu einem anderen Zeitpunkt wieder aufgenommen werden. Da aber lediglich 3,3% der Interviews unterbrochen wurden und später zu Ende geführt wurden gibt diese Variable *anfztnm* in großem Umfang Auskunft über die Startzeiten.

Tabelle 4: Beginn des Interviews (Stundenzeiträume)

	Häufigkeit	Prozent	Kumulierte Prozente
10 Uhr	20	,7	,7
11 Uhr	16	,5	1,2
12 Uhr	17	,6	1,8
13 Uhr	11	,4	2,1
14 Uhr	5	,2	2,3
15 Uhr	29	1,0	3,3
16 Uhr	48	1,6	4,9
17 Uhr	688	22,9	27,8
18 Uhr	785	26,2	54,0
19 Uhr	809	27,0	80,9
20 Uhr	568	18,9	99,9
21 Uhr	4	,1	100,0
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Die überwiegende Anzahl von Interviews begannen zwischen 17 und 21 Uhr, knapp 5% wurden zwischen 10 und 16 Uhr durchgeführt. Dies liegt daran, dass individuelle Terminvereinbarungen gemacht wurden, um auch solche Zielpersonen in die Stichprobe aufzunehmen, die zu den üblichen Interviewzeiten nicht zu erreichen waren.

## 4.2 Interviewdauer

Die vereinbarte Interviewdauer betrug 30 Minuten. Bereits im Pretest deutete sich an, dass dieses Zeitlimit überschritten werden würde, woraufhin einige Änderungen vorgenommen wurden. Dennoch errechnet sich für die Nettobefragungsdauer, also die Befragungsdauer von der ersten Frage bis zur letzten Frage, eine Zeitdauer von 32,33 Minuten (Tabelle 5). Interessant ist, dass die Spannbreite zwischen kürzestem und längstem Interview sehr groß ist. Das kürzeste Interview dauerte lediglich 15 Minuten, wogegen das längste 106 Minuten dauerte. Die Tatsache, dass zwischen kürzestem und längstem Interview eine extrem große Spannbreite liegt, 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.

Tabelle 5: Minimum/Maximum-Werte der Nettobefragungsdauer

	Minimum	Maximum	Mittelwert
<b>Minuten- Nettozeit</b>	15	106	33,40

Die Konsequenz war, dass die Befragten auswichen, versuchten, sich für ihre Antwort zu rechtfertigen, und dadurch die Befragungsdauer verlängerten. Besonders macht dies Tabelle 6 deutlich, die Interviews über mehrere Zeitabschnitte verteilt darstellt: 49% der Interviews dauerten mehr als 30 Minuten.

Tabelle 6: Verteilung der Interviews über die Zeitdauer

	Häufigkeit	Prozent	Kumulierte Prozente
15 bis 20 Minuten	43	1,4	1,4
21 bis 30 Minuten	1485	49,5	50,9
31 bis 40 Minuten	1032	34,4	85,3
41 bis 50 Minuten	308	10,3	95,6
51 bis 60 Minuten	88	2,9	98,5
61 bis 70 Minuten	27	,9	99,4
71 bis 80 Minuten	10	,3	99,8
81 bis 90 Minuten	4	,1	99,9
91 bis 106 Minuten	3	,1	<b>100,0</b>
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Für die Bruttobefragungsdauer, also die Zeitdauer von Kontaktbeginn bis zum Ende des kompletten Interviews, errechneten wir 35,88 Minuten.

## 4.3 Unterbrechungen und Abbrüche von Interviews

Bei 3,3% der Interviews kam es zu Unterbrechungen des Interviews durch die Zielpersonen und einer Fortführung zu einem späteren Termin. Ein Grund hierfür kann sein, dass das Interview länger dauerte als zunächst angenommen und die Zielpersonen das Interview nur zu einem späteren Zeitpunkt beenden konnten.

Es ist weiter denkbar, dass die Zielperson sich zu sehr beansprucht fühlte und erst nach einer Pause in der Lage war, das Interview zu Ende zu führen. Tabelle 7 zeigt die Interview-Verteilung über die Anzahl der Teilinterviews. Bei 1% der Interviews bedurfte es sogar drei Anläufe und mehr bis zum vollständigen Interview.

Tabelle 7: Anzahl der Teilinterviews

Anzahl der Teil-Interviews	Häufigkeit	Prozent	Kumulierte Prozente
1	2900	96,7	96,7
2	63	2,1	98,8
3	22	,7	99,5
4	9	,3	99,8
5	4	,1	99,9
6	1	,0	100,0
7	1	,0	100,0
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	

Außerdem registrierten wir 261 abgebrochene Interviews (siehe auch Bruttodaten), was einer Abbruchquote von 8,7% der realisierten Interviews entspricht. Diese Abbruchquote liegt um das 4-fache über der üblichen Abbruchquote, was deutlich macht, dass anfangs befragungsbereite Zielpersonen überdurchschnittlich oft das Interview abbrechen. Bis zum Beginn der Staatsangehörigkeitsfragen (Q011A) hatten bereits 79 Personen das Interview abgebrochen. Es handelt sich dabei um Personen, die sich nur schwer zu dem Interview überreden ließen und bald doch nicht mehr mitmachen wollten. Diese anfängliche Abbruchquote von 2,6% entspricht etwa dem durchschnittlichen Abbruchwert und erfasst Personen, die von Anfang an nicht sonderlich bereitwillig waren, das Interview zu führen.

Je länger die Befragung dauert, desto mehr Zielpersonen beenden jedoch das Interview vorzeitig. Dies ist ein eindeutiges Indiz dafür, dass die Länge und Schwierigkeit des Interviews zu Ermüdungseffekten führen. Vorschläge zur Vermeidung von Ermüdungseffekten werden unter Punkt X Schlussfolgerungen zu Verbesserungen gemacht.

## 4.4 Wiederbefragungsbereitschaft

Die Wiederbefragungsbereitschaft wurde durch eine Frage (QBEFRA) am Ende des Fragebogens ermittelt (siehe Seite 40 in diesem Dokument). Diese Frage ist insofern wichtig, weil ein Pool von wiederbefragungsbereiten Personen für das Panel in den nächsten zwei Jahren rekrutiert werden muss.

Die im Rahmen des Pretests erhobene Wiederbefragungsbereitschaftsquote war mit 97% extrem hoch. Wie bereits erwähnt sank die Wiederbefragungsbereitschaft in der Hauptuntersuchung um 10 Prozentpunkte von 97% auf 87%. Diese nicht sehr gesunkene Wiederbefragungsbereitschaft ist angesichts der Länge und Schwierigkeit des Interviews ein gutes Ergebnis und ermöglicht die Stichprobenziehung im Rahmen des Panel. Die Fallzahl des Panels im zweiten Jahr soll 1300 Befragte, im dritten Jahr 900 Befragte betragen.

## 5. Von der Brutto- zur Nettostichprobe: Stichprobenausschöpfung und Stichprobengewichtung

### 5.1 Bruttostichprobe und 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 (bereinigten) Brutto, der sog. „Stichprobenausschöpfung“.

Die folgende Übersicht **„Rücklaufmeldung“** zeigt den Bruttoansatz der Stichprobe, die verschiedenen Arten von Ausfällen und die Berechnung der Stichprobenausschöpfung.<sup>4</sup>

Wiedergegeben ist das Ergebnis des letzten Kontaktversuchs mit dem jeweiligen Haushalt bzw. der jeweiligen Zielperson. Auf die Zahl der Kontaktversuche wird im folgenden Kapitel 5.2 eingegangen.

Die Ausgangsstichprobe („Bruttoansatz“) umfasste 14.225 Telefonnummern. Nach Abschluss der Feldarbeit sind über 11.000 hiervon 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 5.000 der eingesetzten Telefonnummern des Bruttoansatzes zu. Es verbleiben 8.307 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:

- 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

---

<sup>4</sup> Die Ausfallgründe sind teilweise zusammengefasst. Die in Klammern genannten Code-Nummern zeigen die Ausfallkategorien an, wie sie im Bruttoband dokumentiert sind (vgl. Abschnitt 5.2).

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 7.393 Telefonnummern. Im Laufe der Feldarbeit wurden hiervon 4.393 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. Dies betrifft im vorliegenden Projekt 1.015 Fälle, das sind 13,7% des bereinigten Brutto. 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

Dies betrifft  $n = 2.251$  Fälle, das sind 30,4% des bereinigten Brutto. Damit ist dieser Ausfallgrund 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. Insgesamt waren 11,6% der Ausfälle Verweigerungen von Seiten der Zielpersonen.

(4) Abbrüche des Interviews

In 261 Fällen wurde das Interview begonnen, dann aber vor Abschluss abgebrochen. Die Abbruchquote ist mit 3,5% deutlich höher als in den meisten – meist kürzeren – Telefonbefragungen. Hierauf wurde in Kapitel 2 bereits eingegangen.

Um 3.000 Interviews zu realisieren, oder eine bereinigte Bruttostichprobe von 7.393 Telefonhaushalten benötigt. Die Stichprobenausschöpfung beträgt demnach 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.

## Rücklaufmeldung

**Projekt:** 08.10.22674  
**Infratel-Nr.:** 2267401  
**Projektleiter:** Dr. Gensicke / B.v.Rosenblatt  
**Studienbereich:** SoFo  
**Feldzeit:** 21. 05. bis 28. 06. 2002  
**Zielgruppe:** Bev. ab 16 Jahren  
**Befragungsgebiet:** Bundesrepublik  
**Auswahlverfahren:** ADM-Sample  
**Telefon-Studio:** Parchim, München

### *Ergebnis des letzten Kontaktversuchs*

<b>Bruttoansatz</b>	<b>14225</b>	<b>100,0%</b>	<b>14225</b>
Falsche Telefonnummer (31, 88, 89)		33,7%	4788
Fax / Modem (35)		4,5%	640
Doppeladresse (39)		0,1%	14
Falsche Regionalzelle (42)		-	0
Regionalzelle nicht zu ermitteln (43, 65)		-	0
Maximale Kontaktzahl (>12) erreicht (15, 16, 17, 18)		3,3%	476
<b>Gesamt: Nicht verwendbare Tel-Nr.</b>		<b>41,6%</b>	<b>5919</b>
<b>Verwendbares Brutto</b>	<b>8307</b>	<b>100,0%</b>	<b>8307</b>
Zelle voll (49)		-	0
Kein Privathaushalt (30, 37)		5,6%	466
Keine Pers. i.d. Zielgruppe (41)		-	0
im HH keine Verständigung (70)		3,4%	279
mit ZP keine Verständigung (71)		0,8%	69
ZP nicht in der Lage (73)		1,2%	100
<b>Neutrale Ausfälle gesamt</b>	<b>914</b>	<b>10,9%</b>	<b>914</b>
<b>Bereinigtes Brutto</b>	<b>7393</b>	<b>100,0%</b>	<b>7393</b>
Teilnehmer nimmt nicht ab (10, 82, 84)		7,6%	561
Besetzt (11, 85, 87)		0,1%	10
priv. Anrufbeantworter (19)		6,0%	444
Abbruch Interview (20, 26, 90)		3,5%	261
ZP vorübergehend krank (21)		0,0%	3
KP beschäftigt (24)		2,4%	179
ZP beschäftigt (25)		2,6%	192
KP verweigert Auskunft (51, 53, 54)		28,0%	2072
ZP verweigert Auskunft (62, 63, 64)		5,7%	424
Kein Termin möglich (72)		3,3%	247
Sonstige Ausfälle		-	0
<b>Ausfälle gesamt</b>	<b>4393</b>	<b>59,4%</b>	<b>4393</b>
<b>Durchgeführte Interviews</b>	<b>3000</b>	<b>40,6%</b>	<b>3000</b>

## 5.2 Die Bruttodaten

Die Daten der 3.000 realisierten Interviews werden 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 methodischen Gesichtspunkten. Die Zahlen der Übersicht „Rücklaufmeldung“ sind unmittelbar in den Bruttodaten nachvollziehbar. Die Fallzahl beträgt  $n = 14.225$ , entspricht also dem Bruttoansatz der Stichprobe.

Grundlage für die Erstellung des Bruttodatensatzes der sog. „Account File“, in dem während der laufenden Feldarbeit jeder Kontaktversuch bei jeder eingesetzten Telefonnummer automatisch dokumentiert wird. Im vorliegenden Projekt sind für die 14.225 eingesetzten Telefonnummern insgesamt 33.290 Kontakte dokumentiert.

Die Zahl der Kontaktversuche bei jeder einzelnen Telefonnummer sind in den Bruttodaten in der Variable „wie oft“ angegeben. Die Werte reichen von 1 Kontakt bis zu einem Maximum von 28 Kontakten. Die folgende Tabelle zeigt die prozentuale Verteilung getrennt für Ausfälle und realisierte Interviews.

Tabelle 8:  
Anzahl Kontaktversuche pro Telefonnummer

	Ausfälle %	Interviews %
1 Kontakt	67,4	35,1
2 Kontakte	10,7	25,0
3 Kontakte	5,8	14,0
4- 6 Kontakte	8,3	19,2
7- 9 Kontakte	6,2	4,8
10-12 Kontakte	1,5	1,6
13 und mehr Kontakte	0,1	0,4
	100,0	100,0
Basis	$n = 11.225$	$n = 3.000$

Die Ausfälle werden zu Ausfällen häufig schon beim ersten Kontaktversuch. Insbesondere gilt das natürlich für die „nicht verwendbaren“ Telefonnummern. Bei immerhin 3.660 Telefonnummern wurden mehrere Kontaktversuche unternommen, bis die Situation (negativ) geklärt war.

Anders stellt sich die Verteilung bei den realisierten Interviews dar. Nur in einem Drittel der Fälle konnte das Interview gleich beim ersten Kontakt realisiert werden. In der Mehrzahl der Fälle werden 2 – 6 Kontaktversuche unternommen, bis das Interview realisiert ist. In 7% der Fälle waren noch mehr Kontakte bzw. Kontaktversuche erforderlich, das Maximum liegt bei 22 dokumentierten Kontakten.

Die Bruttodaten enthalten drei weitere Informationen:

**ldadr** ID-Nummer der jeweiligen Telefonnummer im ursprünglichen Adressenpool, der in ITMS für das Projekt bereitgestellt wurde (vgl. zur Erläuterung Kapitel1). Da diese nicht alle zum Einsatz kommen, sind die Nummern in der Abfolge nicht fortlaufend. Die Variable **ldadr** ist für die realisierten Interviews auch in den Nettodaten überspielt. Über diese Variable können beide Datensätze bei Bedarf verknüpft werden.

**datum** Tag des letzten Kontakts bzw. Kontaktversuchs

**ausfall** Ergebnis des letzten Kontakts bzw. Kontaktversuchs

Die folgende Übersicht zeigt die vollständige Liste der Ausfallcodes. Die Verteilung ist im vorigen Abschnitt 5.1 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'
```

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 der Schwedenschlüssel ausgewählt die nächst älteste 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 Interviewer 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 durchstellt kriegt, 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).

## 5.3 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 Gemeindegrößeklassen 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 hinzugekommen 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.<sup>5</sup> Der Strukturabgleich folgt innerhalb von vier Teilgruppen, nämlich Ost- und Westdeutschland und hier jeweils nach zwei Altersgruppen (bis zu 50/über 50 Jahre).<sup>6</sup>

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

---

<sup>5</sup> 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.

<sup>6</sup> 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 Stichprobe<sup>7</sup>
- (3) Demografiegewichtung
- (4) Bildungsgewichtung

Im Ergebnis sollten die Merkmalsverteilungen in der Stichprobe mit der Bevölkerungsstruktur lt. amtlicher Statistik übereinstimmen.<sup>8</sup>

Dass dies tatsächlich der Fall ist, zeigt die folgende Tabelle 9.

Tabelle 9

**Strukturdaten der Stichprobe im Vergleich mit Referenzdaten der amtlichen Statistik**

Strukturmerkmale	Verteilungen in der Stichprobe nach Gewichtung		Amtliche Statistik <sup>1</sup>
	Fallzahl	%	%
<b>Bundesländer</b>			
Schleswig-Holstein	105	3,5	3,5
Hamburg	60	2,0	2,0
Niedersachsen	291	9,7	9,7
Bremen	24	,8	,8
NRW	640	21,3	21,3
Hessen	216	7,2	7,2
Reinland-Pfalz	149	5,0	4,9
Baden-Württemberg	361	12,0	12,1
Bayern	438	14,6	14,6
Saarland	40	1,3	1,3
Berlin ges	122	4,1	4,0
Brandenburg	103	3,4	3,4
Mecklenburg-Vorpommern	70	2,3	2,4
Sachsen	178	5,9	5,9
Sachsen-Anhalt	105	3,5	3,5
Thüringen	98	3,3	3,3
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	<b>100,0</b>

1) Quelle: Bevölkerungsfortschreibung zum 31.12.2000. Für die Bildungsniveaus: Mikrozensus 1997.

<sup>7</sup> Vgl. dazu oben Kapitel 1.4

<sup>8</sup> Mit gewissen Abweichungstoleranzen wegen des Zusammenhangs der Strukturmerkmale untereinander.

## Strukturdaten der Stichprobe im Vergleich mit Referenzdaten der amtlichen Statistik

Strukturmerkmale	Verteilungen in der Stichprobe nach Gewichtung		Amtliche Statistik <sup>1</sup>
	Fallzahl	%	%
<b>Geschlecht</b>			
männlich	1434	47,8	47,8
weiblich	1566	52,2	52,2
<b>Gesamt</b>	<b>3000</b>	<b>100,0</b>	<b>100,0</b>
<b>Altersgruppen</b>			
16-19	157	5,2	5,3
20-29	388	13,0	12,8
30-39	596	19,9	19,8
40-49	533	17,8	17,7
50-59	428	14,3	14,4
60-69	457	15,3	15,1
70-99	432	14,5	14,9
<b>Gesamt</b>	<b>2992</b>	<b>100,0</b>	<b>100,0</b>
Fehlend	System	8	
<b>Gesamt</b>	<b>3000</b>		
<b>Bildungsniveau</b>			
Gültig			
Untere Schulbildung	1558	54,7	54,1
Mittlere Schulbildung	739	25,9	26,7
Höhere Schulbildung	552	19,4	19,1
<b>Gesamt</b>	<b>2849</b>	<b>100,0</b>	<b>100,0</b>
Schüler, KA	151		
<b>Gesamt</b>	<b>3000</b>		

1) Quelle: Bevölkerungsfortschreibung zum 31.12.2000. Für die Bildungsniveaus: Mikrozensus 1997.

Der Gewichtungsfaktor, der bei Auswertungen für Deutschland gesamt zu verwenden ist, ist als Variable „gewb2“ im Datensatz vorhanden. Sofern Auswertungen für Westdeutschland und Ostdeutschland getrennt vorgenommen werden, ist besser die Variable „gewb1“ zu verwenden, die die disproportionale Anlage der Stichprobe beibehält. Der Faktor gewb2 dagegen nimmt eine Proportionalisierung des Ost-Westverhältnisses entsprechend der Verteilung in der Grundgesamtheit vor.

### 6. Datenprüfung / Datenaufbereitung

Vor Beginn der Hauptstudie wurde der Fragebogen von Infratest gründlich getestet, um mögliche Programmierfehler zu entdecken und gegebenenfalls zu korrigieren. Kurz nach Beginn der Befragung wurden die ersten Daten nochmals auf Fehlerfreiheit hin überprüft. Es fanden sich jedoch keine technische Fehler.

Im Rahmen der Datenaufbereitung wurden die im ASCII-Format erhobenen Daten zunächst in SPSS umgewandelt. Es folgte die Labelung und Codierung der Variablen. Um die teils komplizierte Filterführung verständlicher zu machen, wurden filterbedingte Missing-Werte mit dem Wert 66 versehen und mit inhaltlichen Stichworten gelabelt.

Im Bildungsblock fand eine Datenbereinigung statt: Auszubildende, die angaben, nicht erwerbstätig zu sein, obwohl sie nach unserem Verständnis zu den Erwerbstätigen zählen, wurden durch die Filterführung zunächst in den Nicht-Erwerbstätigen-Block geführt. Sobald sich herausstellte, dass sie Auszubildende sind und somit als erwerbstätig gelten, führte sie der Filter zurück in den Erwerbstätigen-Block und sie beantworteten die dortigen Fragen. Diese Filterführung führte zu Konfusion, und so wurden die Daten dieser elf Zielpersonen in Abstimmung mit der Forschergruppe bereinigt. Dies bedeutet, dass die Angaben, die die Auszubildenden als Nicht-Erwerbstätige machten, gelöscht wurden.

Weiter wurden von Infratest Angaben zur Brutto- wie zur Netto-Interviewdauer an die Daten angespielt, sowie zur Startzeit (*anzth/anztm*) des Interviews. Da die Startzeit den Zeitpunkt des letzten Teilinterviews meint, haben wir zusätzlich eine Variable *teilint* hinzugefügt, die die Anzahl der Teilinterviews angibt. Zur Entstehung von Teilinterviews siehe Punkt 4.3.

Infratest übermittelte insgesamt drei SPSS-Files: Die Interviewer-Daten (n=108), die Bruttodaten (n=14.225) und die Nettodaten der Zielpersoneninterviews (n=3000).

## 7 Schlussfolgerungen für mögliche Verbesserungen in der zweiten Welle

### 7.1 Interviewdauer und Pufferfragen

Als Hauptproblem ist die Länge des Interviews zu nennen. Auf der Grundlage der Auswertungsergebnisse der ersten Welle sollte daher geprüft werden, ob und an welchen Stellen eine Kürzung des Fragebogens inhaltlich vertretbar erscheint.

Um Ermüdungseffekten im Interviewverlauf entgegenzuwirken, könnte darüber hinaus geprüft werden, ob sogenannte "Pufferfragen" in den Interviewverlauf eingefügt werden sollten. Dies sind Fragen, die etwas vom Thema abweichen und die Interviewsituation auflockern. Natürlich ist dies nur möglich, wenn an anderen Stellen eine Kürzung des Fragenprogramms erreicht würde.

### 7.2 Inhaltliche Änderungen

Wir schlagen vor, der Frage km02w1 ("Wie viele Ausländer gibt es an Ihrem Arbeitsplatz?") eine Frage vorzuschalten, die die Erwerbstätigkeit abfragt. Es hat sich in der Hauptstudie gezeigt, dass bei der Frage km02w1 weit weniger Personen (n=813) angaben, nicht berufstätig zu sein ("trifft nicht zu"), als dies bei der direkten Frage nach der Erwerbstätigkeit zu00w1 ("Sind Sie zur Zeit erwerbstätig?") der Fall ist (n=1.234).

Die folgende Anmerkung betrifft den Fragenblock Q043A/B/C: Im Fragenblock A werden lediglich Befragte mit deutscher Staatsangehörigkeit befragt, im Fragenblock B werden alle Ausländer außer Russen, Polen, Rumänen und Staatsangehörige der ehemaligen Sowjetstaaten befragt. Mit diesem Filter möchte man die Aussiedler aus den ehemaligen Ostblockstaaten ausschließen und die sonstigen Ausländer zu den Aussiedlern befragen. In Frage Q043C werden Russen, Polen, Rumänen und Staatsangehörige der ehemaligen Sowjetstaaten zu Türken in Deutschland befragt.

Für Infratest ist dieses Vorgehen problematisch, da deutschstämmige Aussiedler aus den ehemaligen Ostblockländern, nachdem sie nach Deutschland eingereist sind, die deutsche Staatsangehörigkeit erhalten haben und nun nicht mehr als Aussiedler identifizierbar sind. Die Wahrscheinlichkeit ist groß, dass in der ersten Welle Polen, Rumänen, Russen und Staatsangehörige der ehemaligen Sowjetstaaten, die keinerlei deutsche Wurzeln haben, in den Fragenblock C geleitet wurden. Andererseits wurden Aussiedler zu Ausländern befragt (Block A).

Um die Aussiedler von nun an eindeutig von den übrigen Ausländern zu trennen, wäre hier eine Frage zum Geburtsort sinnvoll.

In Frage zu01w1 wurden die Befragten nach ihrer formalen Schulbildung befragt. Nicht erfasst wurden allerdings die Schüler, die noch keinen Schulabschluss haben und sich nach unseren Untersuchungen in diejenige Kategorie einordnen, die sie anstreben (Bsp.: Gymnasiasten ordneten sich der Kategorie Abitur/Fachhochschulreife zu). Wir schlagen vor, diese Frage um die Kategorie "Schüler/Schülerin" zu erweitern.

Ebenso empfiehlt es sich, in Frage zu02w1 ("Haben Sie eine abgeschlossene Berufsausbildung?") die Antwortmöglichkeit "Auszubildende/Auszubildender" aufzunehmen. Damit vermeidet man, dass die Auszubildenden zunächst in den Nicht-Erwerbstätigenblock geleitet werden, bei zu05n1a identifiziert und in den Erwerbstätigen-Block zurückgeschickt werden.

Für die Forschergruppe von Interesse könnte die Aufnahme einer Einkommensvariable sein. Mit der Information über die Einkommen der Befragten wäre es möglich, Zusammenhänge zwischen der Höhe des Einkommens und Einstellungen zu erforschen.

Da die Aufnahme zusätzlicher Fragen in den Fragebogen nur mit einem Verzicht auf andere Fragen einhergehen kann, möchten wir einen Vorschlag zur Eliminierung von Variablen machen: Wie die Befragung ergeben hat, gibt es nur sehr wenig nicht-erwerbstätige Zielpersonen als Haushaltsvorstände, die wiederum keine Rentner sind (n=30). Hier stellt sich die Frage, ob die Forschergruppe auf diese Variable verzichten kann.