

## **BIBB-Übergangsstudie 2011**

- Methodenbericht -

21. Dezember 2011

## Inhaltsverzeichnis

1.	Hintergrund und Zielsetzung der BIBB-Übergangsstudie 2011 .....	3
2.	Grundgesamtheit und Stichprobenziehung .....	3
3.	Ausschöpfung und Response Rates .....	5
4.	Erhebungsinstrument und -methode.....	9
5.	Durchführung der Interviews .....	11
5.1	Feldzeit.....	11
5.2	Interviewdauer .....	11
5.3	Anzahl der Spells .....	13
5.4	Art der Spells .....	15
5.5	Feedback der Interviewer .....	18
6.	Datenaufbereitung und -prüfung .....	19
6.1	Datensätze.....	19
6.2	Datenprüfung.....	20
6.3	Antwortverweigerungen und fehlende Angaben.....	21
6.4	Prüfung der Qualität der Interviews durch ALBERTA .....	22
6.5	Selektivitätsanalyse .....	23
7.	Repräsentativität und Gewichtung .....	25
7.1	Designgewichtung .....	31
7.2	Anpassungsgewichtung.....	32
8.	Gesamtbewertung .....	36
	Anhang: .....	37
	Finale Fragebogenversion .....	37

## **1. Hintergrund und Zielsetzung der BIBB-Übergangsstudie 2011**

Im Rahmen seiner Forschungstätigkeiten hat das Bundesinstitut für Berufsbildung (BIBB) eine neue Übergangsstudie initiiert, die zuletzt im Jahr 2006 durchgeführt wurde. Ziel der Übergangsstudie 2011 ist es unter anderem festzustellen, wie sich der einsetzende Wandel des Ausbildungsmarkts von einem Anbieter- zu einem Nachfragermarkt auf die Bildungs- und Berufswege der heutigen Generation junger Menschen und insbesondere auf die Ausbildungschancen von Jugendlichen mit schwächeren schulischen Voraussetzungen und Jugendlichen mit Migrationshintergrund auswirkt.

Die GfK ist in diesem Zusammenhang vom BIBB beauftragt worden, eine repräsentative Befragung von Jugendlichen bzw. jungen Erwachsenen in Deutschland der Geburtsjahrgänge 1987 bis 1993 hinsichtlich ihrer individuellen Bildungs- und Berufsbiografien durchzuführen. Dabei sollte der Werdegang im Hinblick auf berufliche Ausbildung bzw. Erwerbstätigkeit näher beleuchtet und möglichst konkret offengelegt werden, welche Wege oder auch „Umwege“ die jungen Menschen auf ihrem Weg zu einer Berufsausbildung bzw. Erwerbstätigkeit bisher genommen haben.

Wie schon bei der Vorgängeruntersuchung so sollte auch bei der BIBB-Übergangsstudie 2011 die Befragung mittels computerunterstützter Telefoninterviews (CATI) erfolgen. Da sich bei der BIBB-Übergangsstudie 2006 jedoch gezeigt hatte, dass junge Erwachsene über einen Festnetzanschluss schlecht zu erreichen waren, beauftragte das BIBB die GfK eine reine Mobilfunkbefragung durchzuführen.

## **2. Grundgesamtheit und Stichprobenziehung**

Die Grundgesamtheit der vorliegenden Untersuchung besteht aus Jugendlichen und jungen Erwachsenen, die in den Jahren 1987 bis 1993 geboren wurden, d.h. zum Zeitpunkt der Erhebungen zwischen 18 und 24 Jahre alt waren. Dabei war es unerheblich, ob die betroffenen Personen zum Zeitpunkt der Befragung die allgemeinbildende Schule bereits verlassen hatten oder nicht. Die Grundgesamtheit umfasst somit Schulabgänger, Schulabsolventen und auch Schüler. Nach Angaben der aktuell verfügbaren amtlichen Statistik, dem Mikrozensus 2010, sind dies etwa 6,5 Mio. Personen.

Bei der Stichprobengewinnung der Vorgängerstudie aus dem Jahr 2006 wurden ausschließlich Festnetzanschlüsse berücksichtigt, die seinerzeit in über 95 % aller Haushalte vorhanden waren. Dabei zeigte sich jedoch, dass ein Teil der jungen Menschen aufgrund ihrer hohen Mobilität über

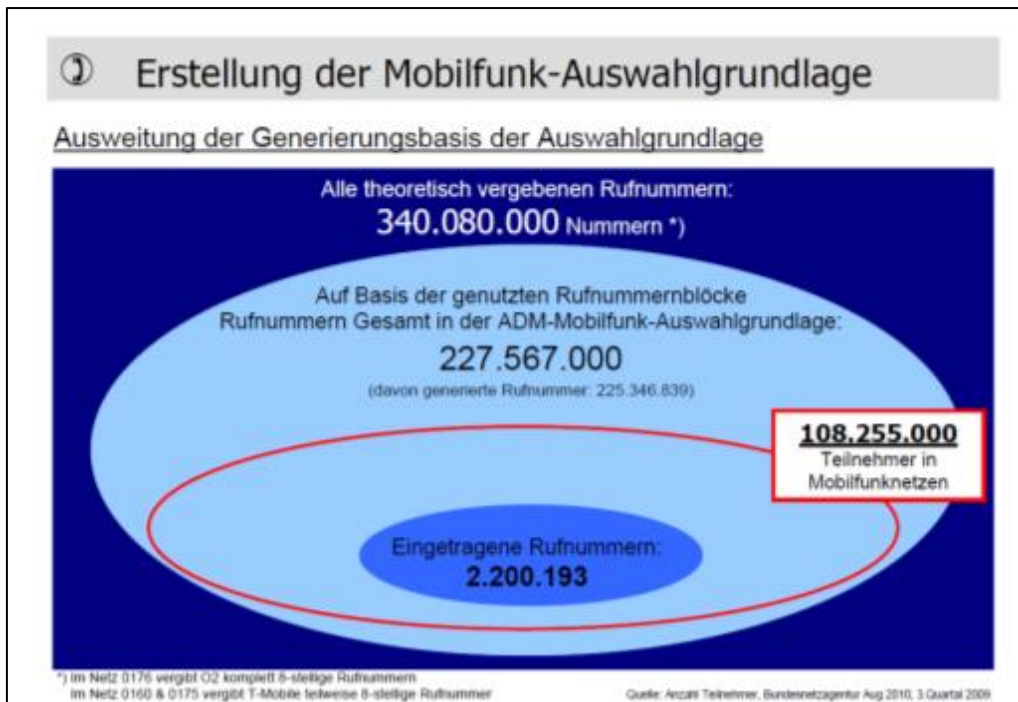
das Festnetz nur noch schwer zu erreichen war; dies betraf insbesondere die Altersgruppe der über 20-Jährigen.

Aus diesen Gründen beauftragte das BIBB die GfK sowohl die Stichprobengewinnung als auch die Durchführung der Interviews bei der BIBB-Übergangsstudie 2011 allein über das Mobilfunknetz vorzunehmen. Grundlagenuntersuchungen der GfK stützen dieses Vorgehen. So ist der Anteil der Personen, die auf einen Festnetzanschluss verzichten und nur noch über ein Mobiltelefon verfügen, weiter gestiegen: Gerade in der relevanten Zielgruppe der 18- bis 24-Jährigen zeigen verschiedene Grundlagenstudien der GfK, dass über ein Viertel heute zu den sogenannten „Nur-Mobilfunk-Telefonierern“ zählen, und damit über einen Festnetzanschluss nicht erreicht werden können. Ferner nutzt die Zielgruppe, selbst wenn im Haushalt ein Festnetzanschluss vorhanden ist, in der Regel zusätzlich ein eigenes Mobiltelefon.

Eine optimale Auswahlgesamtheit stellte hierfür die **„ADM Auswahlgrundlage Mobilfunk“** dar. Bei dieser werden für alle deutschen Mobilfunknetze die theoretisch vergebenen Rufnummern (laut Bundesnetzagentur) in Form von 100.000er-Blöcken gelistet. Um Überschätzungen zu vermeiden, werden allerdings nur solche Blöcke generiert, die auch tatsächlich Einträge (im Telefonbuch oder Internet) aufweisen, und die nicht für Mailboxen, technische Sperrungen oder Sonderservices reserviert sind. Im Ergebnis werden bei dieser Vorgehensweise rund 67 % aller möglichen Blöcke berücksichtigt.

Insgesamt ergeben sich durch diese Art der Generierung ca. 227 Mio. Handy-Nummern, die 340 Mio. theoretisch vergebenen Rufnummern gegenüberstehen. Demgegenüber steht, dass nur etwa 2,2 Mio. Rufnummern eingetragen sind und aus den entsprechenden Verzeichnissen gewonnen werden können.

Abbildung 1: Erstellung der Mobilfunk-Auswahlgrundlage



Aus der ADM Mobilfunk-Auswahlgrundlage der theoretisch vorhandenen Mobilfunknummern wurde eine echte Zufallsstichprobe gezogen. Dabei war gewährleistet, dass Nummern aus allen Nummernbereichen proportional enthalten sind.

### 3. Ausschöpfung und Response Rates

Insgesamt wurden in der Haupterhebung 5.579 Interviews durchgeführt – plus 503 Kurzinterviews mit Verweigerern zur weiteren Untersuchung von systematischen Ausfällen (siehe hierzu Kap. 6.5 „Selektivitätsanalyse“). Um diese Nettostichprobe zu erreichen wurden insgesamt 1.982.326 Mobilrufnummern eingesetzt. Damit war der Bedarf an erforderlichen Rufnummern mehr als sechsmal so hoch wie im Vorfeld geschätzt wurde.

Ursprünglich wurde damit gerechnet, dass nur etwa 310.000 Nummern kontaktiert werden müssten, um 5.500 Interviews zu realisieren. Die wesentlichen Gründe für die deutlich höhere Anzahl an benötigten Kontakten waren die Folgenden:

- Deutlich höhere Anzahl an aktivierter Mailboxen bzw. Ansagentexten
- Nicht zugewiesene Mobilfunknummern
- Nummer out of order

- Nicht spezifizierbare Fehlkontakte

Allein auf diese vier Gründe entfielen rund 1,56 Mio. Ausfälle.

Für die Bestimmung der Ausschöpfungsquote (Response rate) wurde die Systematik der American Association for Public Opinion Research (AAPOR) herangezogen. Der weitaus größte Teil der Fehlkontakte entfiel dabei auf stichprobenneutrale Ausfälle (SN) mit 1.804.200 Rufnummern. Das sind 91% der insgesamt eingesetzten Nummern. Bei den wesentlichen Ausfallgründen handelt es sich darum, dass die Nummer entweder nicht vergeben bzw. außer Betrieb/gestört gewesen ist. Bei mehr als 426 Tsd. Nummern wurde ausschließlich eine Mailbox erreicht.

Fasst man die Top 3-Fehlkontakte der stichprobeneutralen Ausfälle (number out of order, answering machine, unassigned number) zusammen, so entfallen auf diese 70% aller Rufnummern.

Neben den stichprobenneutralen Ausfällen wurden entsprechend AAPOR im Wesentlichen drei Arten von stichprobenrelevanten Fehlkontakten unterschieden: „Unknown Handheld“ (UH)<sup>1</sup>, „Unknown Others“ (UO) und „Non-Contacts“ (NC), die neben der Anzahl durchgeführter Interviews (I) und unterbrochenen bzw. abgebrochenen Interviews (P) für die Berechnung der Response Rates erforderlich sind.

Bei „Unknown Handheld“ (UH) handelt es sich um Fehlkontakte, die eher technischer Art sind und im Wesentlichen auf die Erhebungsmethode Mobilfunkbefragung zurückzuführen sind, wie „besetzt“ oder „keine Antwort“. Bei den übrigen stichprobenrelevanten Fehlkontakten, wie Verweigerungen oder Abbrüche, konnte ein erster persönlicher Kontakt hergestellt werden. Sie werden unter den „Unknown Others“ (UO) subsumiert. Beiden Kategorien ist letztlich gemeinsam, dass nicht bekannt ist, ob diese zur Zielgruppe gehören oder nicht.

So weigerten z.B. sich mehr als 40.000 erreichte Zielpersonen, den Fragebogen zu beantworten, mehr als 11.000 Personen verweigerten grundsätzlich ihre Teilnahme. Letztere wollten unabhängig vom Befragungsinhalt nicht teilnehmen, da sie empirische Erhebungen bzw. Befragungen jeglicher Art grundsätzlich ablehnen. Erstere stehen Befragungen am Telefon nicht grundsätzlich ablehnend gegenüber, wollten aber in dieser konkreten Studie nicht teilnehmen.

---

<sup>1</sup> Nach AAPOR steht UH ursprünglich für „Unknown Households“, da hier stets von Haushaltsstichproben ausgegangen wird. Im Rahmen der Übergangsstudie 2011 mit einer Personenstichprobe basierend auf Mobilfunkanschlüssen wurde das Kürzel hier im Speziellen modifiziert in „Unknown Handhelds“.

In der nachfolgenden Tabelle 1 sind die genauen Anzahlen stichprobenneutraler und -relevanter Ausfälle aufgeführt.

Tabelle 1: Stichprobenneutrale und -relevante Ausfälle

<b>AAPOR Zuordnung der Tipcodes</b>	<b>Text</b>	<b>Anzahl</b>
	<b>eingesetzte Rufnummern</b>	<b>1.982.326</b>
I	Completed Interview	5.592
UO	TP / HH / company refuses to answer	40.195
UO	Leave questionnaire	698
UO	TP / HH generally declines to answer	11.261
NC	Binding appointment	1.384
UH	Busy	477
UH	No answer	712
P	Partial interview	748
UO	Appointment	3.551
UO	Unavailable because of age/disability	657
NC	Communication problems	9.786
UH	Dialer - busy	3.128
UH	Dialer - no answer	99.937
SN	No target person in HH / company	203.995
SN	Company number / Residential number	2.285
SN	Duplicate number	44
SN	Modem Tone/ Fax	1.033
SN	Invalid number	1.490
SN	Answering machine	426.616
SN	Change telephone number	227
SN	Not available during fieldwork	475
SN	Dialer - modem tone	47
SN	Dialer - nuisance hangup	11.346
SN	Dialer - unassigned number	307.125
SN	Dialer - incomplete number	6.552
SN	Dialer - internal dialer error	1.207
SN	Dialer - resource bottleneck	594
SN	Dialer - number out of order	528.755
SN	Dialer - number changed	189
SN	Dialer - unspecified call failure	297.321
SN	Dialer - number barred	14.899

Tabelle 1a: Zusammenfassung der Ausfälle

I	Complete Interview	5.592
P	Partial Interview (Stopped Interview)	748
NC	Non contact	11.170
UO	Unknown other	56.362
UH	Unknown Handheld	104.254
SN	stichprobenneutral	1.804.200
	<b>Gesamt</b>	<b>1.982.326</b>

In Abstimmung mit dem BIBB wurden zur Berechnung der Ausschöpfungsquoten die Indices RR3 und RR4 herangezogen, wie sie von der AAPOR empfohlen werden<sup>2</sup>:

$$RR3 = \frac{I}{(I + P) + NC + e * (UH + UO)}$$

$$RR4 = \frac{(I + P)}{(I + P) + NC + e * (UH + UO)}$$

Dabei bezeichnet e grundsätzlich die geschätzte Inzidenz der Zielgruppe in der Grundgesamtheit. Als Inzidenzrate e wurde das Ergebnis aus der Screeningfrage v101 herangezogen, die ca. 10% beträgt. Die bereits im Pretest ermittelte Inzidenz bestätigte sich damit auch in der Hauptstudie. Sie lag deutlich niedriger als im Vorfeld auf Grundlage vergleichbarer Studien und Erkenntnisse geschätzt worden war.

Auf Grundlage der oben dargestellten Berechnungsmethoden der beiden Response Rates sowie der ermittelten Inzidenz von 10% ergeben sich somit folgende Werte:

$$RR3 = 17\%$$

$$RR4 = 19\%$$

RR3 bezeichnet dabei den Anteil derjenigen Zielpersonen, die für ein vollständiges Interview gewonnen werden konnten, wobei Abbrüche (Stopped Interviews) nicht berücksichtigt werden. RR4 gibt den Anteil derjenigen Personen aus der Zielgruppe an, die ein Interview zumindest be-

<sup>2</sup> Details zur Berechnung von Ausschöpfungsquoten: vgl. The American Association for Public Opinion Research. 2011. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition. AAPOR.



gonnen haben – d.h. also einschließlich der Abbrüche (Stopped Interviews). Die Differenz von RR3 und RR4 kann als ein Qualitätskriterium für die Befragung bzw. den Fragebogen gewertet werden. Eine große Differenz – d.h. relativ viele Interviewabbrüche im Verhältnis zu den vollständigen Interviews – wäre ein Indiz dafür, dass z.B. der Fragebogen für den Befragten zu lange, zu schwierig, oder auch zu indiskret ist, so dass die Teilnahme abgebrochen wird.

Die in der Übergangsstudie 2011 vorliegende Differenz von zwei Prozentpunkten kann als gering angesehen werden und spricht für eine hohe Qualität des Untersuchungsdesigns und speziell des Fragebogens.

#### **4. Erhebungsinstrument und -methode**

Sämtliche Interviews wurden telefonisch per CATI (Computer Assisted Telephone Interview) durchgeführt. In zwei Fällen wurde telefonisch über das Mobilfunktelefon der Zielperson eine Terminvereinbarung über ein Interview über das Festnetz getroffen. Ein Auto-Dialing System wird dabei von der GfK standardmäßig eingesetzt, um insbesondere Terminierungen von Interviews und deren Realisierung optimal zu gewährleisten und nicht existente Rufnummern technisch zu ermitteln. Die CATI-Erhebung erlaubt es außerdem bereits während der Befragung, automatisiert Plausibilitätskontrollen vorzunehmen, die z.B. Tippfehler bei der Eingabe unrealistischer numerischer Werte bereits im Interview vermeidet. Zudem wird über programmierte Filterregelungen der korrekte Ablauf des Interviews sichergestellt.

Für die Übergangsstudie wurde vom BIBB ein Erhebungsinstrumentarium vorgelegt, das im Wesentlichen auf der letzten Erhebung 2006 basierte, jedoch gekürzt und um aktuelle Fragestellungen ergänzt sowie an die nun erstmals vorgenommene Mobilfunkbefragung angepasst wurde. In enger Abstimmung mit dem BIBB wurde das Erhebungsinstrument durch die GfK als CATI Skript programmiert, im Rahmen eines Pretests überprüft und optimiert. Der Pretest basierte auf einer Stichprobe von 50 Zielpersonen aus dem Bundesgebiet. Die Ergebnisse und Implikationen für die Hauptstudie wurden in einem separaten Pretest-Bericht dem BIBB zur Verfügung gestellt.

Für die erfolgreiche Durchführung einer telefonischen Befragung spielt die Screeningphase zu Beginn des Interviews eine entscheidende Rolle. Im Rahmen der BIBB-Übergangsstudie 2011 musste als Kriterium für die Bestimmung der Zielpersonen das Alter ermittelt werden

Personen, die in der Screeningphase als Zielperson identifiziert wurden, aber nicht an der Befragung teilnehmen wollten, wurden gebeten, wenigstens im Rahmen eines Kurzinterviews einige soziodemografische Angaben zu ihrer Person zu machen, wie Alter, Geschlecht, Schulabschluss und Staatsangehörigkeit. Damit konnte im Nachhinein geprüft werden, ob sich die Verweigerer im Hin-

blick auf die soziodemografischen Merkmale systematisch von den Personen, die das vollständige Interview durchgeführt haben, unterscheiden. Insgesamt konnten 503 Personen für dieses Kurzinterview gewonnen werden. Die betreffenden Angaben wurden im Datensatz abgelegt.

Der im Anhang dieses Berichts beigefügte Fragebogen untergliederte sich in fünf Hauptbereiche:

- Screening der Zielperson und Kurzinterview für Verweigerer
- Eingangsfragen und Fragen zum Grundschulbesuch
- Retrospektive Abfrage der Ausbildungs- und Berufsbiografie im Längsschnitt
- Angaben zur Zielperson
- Angaben zu den Eltern der Zielperson

Zentraler Bestandteil des Fragebogens ist die Längsschnittanalyse der Ausbildungs- und Berufsbiografie der Jugendlichen und jungen Erwachsenen. Vorrangiges Ziel war es, alle diesbezüglichen Aktivitäten der Zielpersonen möglichst vollständig und lückenlos zu erfassen. Um dies zu gewährleisten, wurden den Befragten alle erfassten Aktivitäten am Ende der retrospektiven Abfrage nochmals vorgelesen. Zusätzlich ermittelte eine Prüfroutine, ob zwischen den einzelnen Aktivitäten zeitliche Lücken von mehr als zwei Monaten vorhanden waren oder Aktivitäten ganz vergessen wurden. Zur Schließung dieser Lücken sowie zur Ergänzung zunächst vergessener Aktivitäten wurden die Probanden noch einmal in die Längsschnittabfrage der Ausbildungs- und Berufsbiografie zurückgeleitet. Erst nachdem alle Aktivitäten lückenlos erfasst waren, wurde das Interview mit den Angaben zur Zielperson und zu deren Eltern fortgeführt und abgeschlossen.

Die Abfrage der Ausbildungs- und Berufsbiografie geschah mit Hilfe sogenannter Spells, die alle möglichen Aktivitäten umfassten. In Tabelle 2 sind die Spells aufgeführt.

Tabelle 2: Im Rahmen der BIBB-Übergangsstudie relevante Spells

	Schulbesuch
	Berufliche Ausbildung, Lehre
	Studium
	Berufsvorbereitung, Berufsgrundbildung, Berufsfachschule, die nicht zu einem Berufsabschluss führt
	Einstiegsqualifizierung, Praktikum und Teilqualifizierung
	Erwerbstätigkeit, Jobben

	Langfristige Fortbildung oder Umschulung
	Maßnahmen der Bundesagentur für Arbeit (BA)
	Wehrdienst, Zivildienst, Freiwilliges Soziales oder Ökologisches Jahr
	Sonstiges oder zu Hause (z.B. Suche nach Ausbildungs- oder Arbeitsplatz, Kinderbetreuung etc.)

## 5. Durchführung der Interviews

### 5.1 Feldzeit

Die Haupterhebung startete am Montag, 4. Juli 2011 und erstreckte sich über insgesamt 12 Wochen. Die letzten Interviews wurden am Freitag, 23. September 2011 durchgeführt. Die Feldzeit überlagerte somit komplett die Sommerferienzeiten sämtlicher Bundesländer.

Durchschnittlich wurden pro Tag etwa 80 Interviews durchgeführt, wobei es je nach Terminvereinbarungen auch teilweise zu weniger als 20 bzw. bis zu 200 Interviews pro Tag kam. Für die Durchführung der Interviews wurden primär nur Werktage von Montag bis Samstag genutzt. Dabei wurde montags bis freitags jeweils von 16.00 Uhr bis 21.00 Uhr und samstags von 10.00 Uhr bis 17.00 Uhr telefoniert. Auf Wunsch von Zielpersonen wurden aber grundsätzlich auch Termine außerhalb der regulären Interviewzeiten vereinbart und abtelefoniert - in Einzelfällen wurden auch an Sonntagen Interviewtermine wahrgenommen.

### 5.2 Interviewdauer

Die durchschnittliche Dauer der 5.579 auswertbaren Interviews lag bei 30,2 Minuten (Median = 28,7 min). Damit wurde die vorab veranschlagte Dauer von 30 Minuten fast punktgenau erreicht. In der Vorgängerstudie von 2006 betrug die Interviewdauer noch 34,4 Minuten.

Das längste Interview dauerte 89,9 Minuten (2006: 107,0 Minuten), das kürzeste gut 9 Minuten (2006: 11,6 Minuten). Die Standardabweichung beträgt 10,2 Minuten (2006: 11,9 Minuten).

Tabelle 3: Interviewdauer

Anzahl auswertbarer Interviews	5.579
Interviewdauer (Durchschnitt)	30,2 Minuten
Kürzestes Interview	9,1 Minuten
Längstes Interview	89,9 Minuten
Standardabweichung	10,2 Minuten

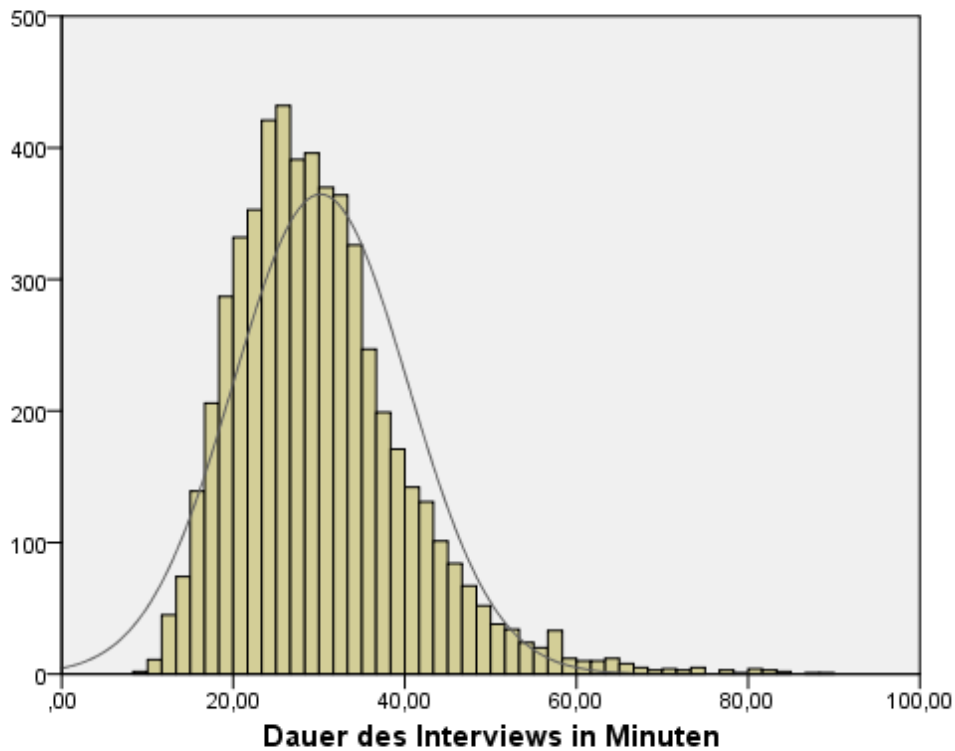
Basis: 5.579 auswertbare Interviews

Die relativ hohe Varianz der Länge der einzelnen Interviews hat vor allem zwei Gründe:

- 1) Die Zielpersonen haben durchaus unterschiedliche Bildungs- und Berufsbiografien. Während z.B. Schüler einer allgemeinbildenden Schule jeweils nur einen einzigen Spell durchlaufen mussten, haben Erwerbstätige mit Lehre, Wehrdienst und mehreren Schulformwechseln entsprechend mehrere Spells zu beantworten.
- 2) Eine Reihe von Probanden musste, unabhängig vom eigentlichen Untersuchungsinhalt, generell von der Seriosität der Befragung zu Beginn des Interviews überzeugt werden. Diese Probanden hatten z.B. Fragen zum Datenschutz, zur Seriosität der Untersuchung bzw. des Instituts, und/oder verbanden mit dem Anruf zunächst unseriöse Marketing-/Vertriebsaktionen, worauf die Interviewer entsprechend geschult reagierten. Die Zeit für derartige Kontaktgespräche auf einer Metaebene fließen bei der Interviewdauer ebenso mit ein.

Abbildung 2 veranschaulicht die Verteilung der Gesprächsdauern pro Interview.

Abbildung 2: Verteilung der Gesprächsdauer



### 5.3 Anzahl der Spells

Wie oben ausgeführt hängt die Interviewdauer in hohem Maße davon ab, wie viele Spells ein Befragter zu durchlaufen hat. Tabelle 4 gibt die Anzahl der durchlaufenen Spells aller Aktivitäten nach der Grundschule wieder.

Tabelle 4: Anzahl der Aktivitäten nach der Grundschule

Anzahl Spells	absolut	in %
1	515	9,2
2	1.134	20,3
3	1.404	25,2
4	1.137	20,4
5	639	11,5
6	415	7,4
7	168	3,0
8	86	1,5
9	43	0,8
10	18	0,3
11	10	0,2
12	6	0,1
13	3	0,1
15	1	0,0
Gesamt	5.579	100

Insgesamt wurden bis zu 15 Spells im Anschluss an die Grundschulzeit genannt. Drei Viertel der Zielpersonen haben allerdings nur bis zu vier Spells durchlaufen.

Im arithmetischen Mittel gaben die Befragten 3,6 Spells an – der Median beträgt 3 Spells.

Bei den Top 3 der am häufigsten genannten Spells handelt es sich um den Schulbesuch, die Erwerbstätigkeit sowie die berufliche Ausbildung.

## 5.4 Art der Spells

Tabelle 5: Art der Aktivitäten

Art des Spells	absolut	in %
Schule besucht	8.695	43,5
Berufliche Ausbildung, Lehre	2.903	14,6
Studium, Berufsakademie	1.319	6,6
Berufsvorbereitung, -grundbildung oder -fachs Schule	593	3,0
Einstiegsqualifizierung	529	2,7
Erwerbstätig oder gejobbt	3.579	18,0
Langfristige Fortbildung, Umschulung in Vollzeit	52	0,3
Trainings-, Weiterbildungsmaßnahme, ABM, 1-Euro-Job o.Ä.	218	1,1
Wehrdienst, Zivildienst, Soziales oder Ökologisches Jahr	646	3,2
Sonstiges oder waren zu Hause	1.397	7,0
Gesamt	19.895	100

Insgesamt wurden von den 5.579 Teilnehmern 19.895 Spells genannt. Klarer Spitzenreiter ist der Besuch der allgemeinbildenden Schule mit 8.695 Spells (43,5%), was nicht verwundert, da dieser Spell von allen Teilnehmern mindestens einmal zu durchlaufen war. Auf Rang zwei bereits mit deutlichem Abstand folgt die Erwerbstätigkeit bzw. der Job mit 3.579 Durchläufen, dies entspricht 18% aller genannten Spells. Den dritten Platz belegt die berufliche Ausbildung/Lehre. Sie kommt auf knapp 15% der Spells. Auf diese Top 3 entfallen alleine mehr als drei Viertel aller genannten Spells.

### Korrekturen und Ergänzungen der Spells

Es war ein zentrales Anliegen der Untersuchung, die Ausbildungs- und Berufsbiografie der jungen Menschen möglichst lückenlos zu erfassen. Aus diesem Grund wurde am Ende der Aktivitätenschleifen noch einmal explizit nachgefragt, ob die gemachten Angaben vollständig waren. Hierzu wurden dem Interviewer in der Frage v3000 noch einmal alle bisher genannten Aktivitäten inkl. deren angegebenen Beginn- und Enddaten angezeigt. Das Programm signalisierte dem Interviewer, wenn in der Aktivitätenliste eine zeitliche Lücke von drei Monaten und mehr vorhanden war. Diese konnte dann gegebenenfalls ergänzt oder korrigiert werden.

Von den insgesamt vorliegenden 19.895 Spells wurden mit diesem Vorgehen 686 korrigiert (3,4%) und 708 ergänzt (3,6%). Tabelle 6 zeigt wie sich die Ergänzungen auf die einzelnen Aktivitäten verteilen.

Tabelle 6: Ergänzte Aktivitäten

<b>Ergänzte Tätigkeit</b>	<b>absolut</b>	<b>in %</b>
Erwerbstätig/gejobbt	258	36,4
Schulbesuch	231	32,6
Sonstiges oder zu Hause	88	12,4
Berufliche Ausbildung	42	5,9
Einstiegsqualifizierung	41	5,8
Berufsvorbereitung	19	2,7
Studium	16	2,3
Wehr-/Zivildienst, FSJ	7	1,0
Trainingsmaßnahme der BA	5	0,7
Langfristige Fortbildung	1	0,1
<b>Gesamt</b>	<b>708</b>	<b>100</b>



Am häufigsten wurde die Aktivität Erwerbstätigkeit ergänzt. Deutlich mehr als ein Drittel der nachträglich aufgenommenen Spells entfallen darauf. Auf Rang zwei findet sich mit einem Anteil von gut 32% der Schulbesuch. Mit deutlichem Abstand folgt mit gut 12% Sonstiges oder zu Hause. Auf den Rängen vier bis zehn sind die übrigen Aktivitäten mit Anteilen zwischen 5,9 und 0,1% zu finden. Mehr als 80% der Ergänzungen entfallen auf die Top 3-Aktivitäten (Erwerbstätigkeit, Schulbesuch und sonstiges/zu Hause).

### Jahreszeiten als Datumsangabe

Im Falle, dass sich Befragte nicht mehr an die genauen Monate für den Beginn oder das Ende einer Aktivität erinnern konnten, wurden neben der Eingabe des exakten Monats für den Beginn oder das Ende eines Spells auch Jahreszeiten, wie Sommer oder Herbst, zugelassen. Ziel dieser Vorgehensweise war es, den Anteil der fehlenden Angaben bei diesen zentralen Fragen möglichst weit zu reduzieren.

Wie die folgende Tabelle 7 zeigt, haben die Befragten auch relativ häufig von Jahreszeitangaben Gebrauch gemacht.

Tabelle 7: Angaben von Jahreszeiten

<b>Jahreszeit</b>	<b>Beginn der Aktivität</b>	<b>Ende der Aktivität</b>
Jahresanfang	98	63
Frühjahr/Ostern	117	139
Sommer	1.299	1.394
Herbst	730	159
Jahresende	62	77
<b>Gesamt</b>	<b>2.306</b>	<b>1.832</b>

Insgesamt wurde in mehr als 2.300 Fällen zum Beginn der Aktivität und in mehr als 1.800 Fällen zum Ende der Aktivität eine Jahreszeit angegeben. Das heißt, in etwa 10% der Spells wurde auf eine konkrete Monatsangabe verzichtet und stattdessen eine etwas ungenauere Angabe gemacht. Dennoch muss dies eindeutig als Informationsgewinn gesehen werden und rechtfertigt damit die Zulassung dieser Jahreszeitangaben.

Am häufigsten wurde der Sommer genannt, sowohl für den Beginn als auch für das Ende einer Aktivität. Dies ist nicht überraschend, denn die am häufigsten genannte Aktivität, der Schulbesuch, beginnt oder endet meist in den Sommermonaten.

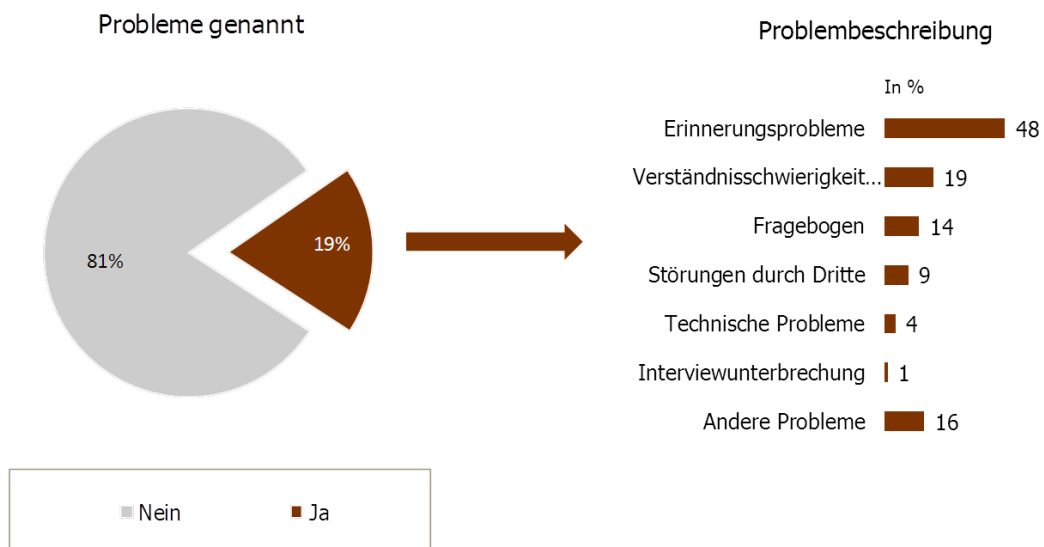
Es fällt zudem auf, dass der Herbst wesentlich häufiger als der Beginn der Aktivität genannt wurde. Wenn man berücksichtigt, dass z.B. in der Regel im Herbst das Studium oder auch eine Lehrstelle angetreten wird, ist dieses Ergebnis naheliegend.

## 5.5 Feedback der Interviewer

Nach Abschluss jedes einzelnen Interviews war jeder Interviewer aufgefordert, über etwaige Probleme während des Interviews zu berichten.

Die von den Interviewern gemachten Angaben zeigen, dass mehr als 80% der Interviews ohne jegliche Probleme verliefen. In 19% der Fälle berichteten sie allerdings von Schwierigkeiten während des Interviews zu.

Abbildung 3: Anteil und Art genannter Probleme



Basis: 5.579 Teilnehmer aus den Geburtsjahren 1987 bis 1993

Die mit Abstand größte Schwierigkeit für die Befragten lag dabei im Erinnern. In nahezu der Hälfte der betreffenden Fälle hatten die Befragten Probleme damit, sich spezifische Inhalte bzw. Zeiten ihrer Bildungs- und Berufsbiografie ins Gedächtnis zurückzurufen. Dies ist vor allem bei Personen zum Tragen gekommen, die mehrere Spells durchlaufen haben. Knapp ein Fünftel der Interviews mit Problemen betraf individuelle Verständnisschwierigkeiten, die die Befragten mit einzelnen Fragestellungen hatten. Diese waren entweder sprachlicher oder vereinzelt auch inhaltlicher Natur.

Mit der Länge und Gestaltung des Interviews gab es in 14% der Fälle, bei denen die Interviewer von Problemen berichteten, Schwierigkeiten. Mehrheitlich war der Fragebogen in den Augen der Befragten zu lang bzw. das ständige, monotone Wiederholen von Fragestellungen in den einzelnen Spells wurde kritisiert.

Bei knapp jedem Zehnten der „Problemfälle“ wurde der Befragte während des Interviews durch Dritte gestört, während es in 4% der Fälle zu technischen Problemen kam, meist technische Abbrüche wegen Funkloch oder leeren Akkus.

Insgesamt kann man feststellen, dass die hier auftretende Häufigkeit und auch die Art der genannten Probleme vergleichbar sind mit ähnlichen Untersuchungen. Im Gegensatz zu Befragungen im Festnetz sind erwartungsgemäß Probleme rein technischer Art (Funkloch, leerer Akku) aufgetreten, die in ihrem Umfang völlig akzeptabel erscheinen.

## **6. Datenaufbereitung und -prüfung**

### **6.1 Datensätze**

Vor dem Hintergrund der Tatsache, dass Auswertungen sowohl aktivitätsbezogen als auch personenbezogen vorgenommen werden sollen, wurden zwei Datensätze in SPSS erstellt. Zunächst wurden die erhobenen Daten als Ereignisdatensatz auf die Aktivitäten bezogen abgelegt. Dieser Ereignisdatensatz umfasst 20.398 Fälle, d.h. Aktivitäten, wobei die 503 Fälle der Kurzbefragung der Verweigerer („Screen outs“) mit enthalten sind. Bei Herausnahme der Screen outs sind es 19.895 Aktivitäten.

Mit Hilfe einer von der GfK erstellten Syntax, die dem BIBB zur Verfügung gestellt wurde, kann der Ereignisdatensatz in einen personenbezogenen Falldatensatz umgewandelt werden. Dieser enthält 5.579 auswertbare vollständige Interviews sowie 503 Fälle aus der Kurzbefragung der Verweigerer (insgesamt somit 6.082 Fälle).

## 6.2 Datenprüfung

Ein Kernelement im Rahmen der Qualitätssicherung stellt die nachträgliche Prüfung der erhobenen Daten dar. Durch den Einsatz von CATI, das bereits eine exakte Filterführung sicherstellt, verlagerte sich ein großer Teil des Prüfaufwands bereits auf die CATI Programmierung, so dass damit die nachträgliche Datenüberprüfung deutlich vereinfacht werden konnte.

In einem ersten Schritt wurde der Datensatz dahingehend überprüft, ob die genannten Spells jeweils korrekt abgelegt und die Filter gemäß der Vorgabe aus dem Fragebogen programmiert worden waren. Im Einzelnen wurden die Fallzahlen in den jeweiligen Spells (Fragen v2000a bis v2861bnb) sowie die Nennungen im gesamten Datensatz geprüft.

In einem weiteren Schritt wurde die korrekte Ablage der Datumsvariablen (z.B. v2105) einem Check unterzogen. Dabei zeigte sich, dass sich die überaus intensive Testphase im Rahmen des Pretests, insbesondere der Filterung, ausgezahlt hatte, da keinerlei Filterfehler gefunden wurden.

Neben der korrekten Ablage der Datumsvariablen galt im Rahmen der Plausibilisierung den Fragen nach der Anzahl der Handy-Rufnummern (v5050n/v5051n), unter der der Befragte jeweils erreichbar war, besonderes Augenmerk. Diese Fragen wurden für die Bildung des Designgewichtes herangezogen<sup>3</sup>.

In Absprache mit dem BIBB wurde folgendes Vorgehen bei der Bereinigung dieser Frage vereinbart:

- Grundsätzlich erscheinen Werte von mehr als 10 genannten Rufnummern als unplausibel.
- „Schnapszahl“-Angaben, wie 11, 22, 33 usw., werden als Erfassungsfehler betrachtet und werden jeweils entsprechend auf 1, 2, 3 usw. gesetzt.
- Alle übrigen Angaben, die den Wert 10 übersteigen, werden auf „keine Angabe“ (k.A.) gesetzt.
- Abschließend werden alle k.A.-Werte auf den Medianwert, der bei eins liegt, gesetzt.

Zusammenfassend ist zu sagen, dass sich die in der Programmierung sehr aufwändige und komplexe Überprüfung der Biografie der Befragten im Hinblick auf deren Vollständigkeit sowie der Möglichkeit der Korrektur und/oder Ergänzung als gerechtfertigt erwiesen und die Datenqualität dadurch erheblich gesteigert werden konnte.

---

<sup>3</sup> Vgl. hierzu auch Abschnitt 7.1. in diesem Report

### 6.3 Antwortverweigerungen und fehlende Angaben

Ein weiteres Kriterium für die Qualität der Stichprobe ist der Anteil der Antwortverweigerungen bzw. fehlender Antworten. Da in der vorliegenden Studie die Zeitangaben der einzelnen Aktivitäten für die Auswertung der jeweiligen Biografien von zentraler Bedeutung sind, wird die Analyse fehlender Angaben anhand dieser Variablen vorgenommen. Tabelle 8 gibt einen Überblick über die fehlenden Zeitangaben in Abhängigkeit der Aktivitäten

Es ergeben sich folgende Werte:

Tabelle 8: Anteil der fehlenden Zeitangaben nach Aktivität

<b>Aktivität</b>	<b>Verweigerungen/keine Angabe absolut</b>	<b>in % der genannten Spells</b>
Grundschule	368	6,6%
Schulbesuch	395	4,9%
Berufliche Ausbildung	112	3,6%
Studium	76	5,7%
Berufsvorbereitung	70	7,0%
Einstiegsqualifizierung	14	2,6%
Erwerbstätig/gejobbt	130	3,6%
Langfristige Fortbildung	4	3,3%
Trainingsmaßnahme der BA	12	6,7%
Wehr-/Zivildienst, FSJ	14	2,3%
Sonstiges oder zu Hause	39	2,9%

Über alle Datumsangaben hinweg ergibt sich ein Anteil fehlender Angaben von 4,8%. Dies ist vor dem Hintergrund der Tatsache, dass sich die Befragten hinsichtlich des Beginns und dem Ende einer Aktivität sowohl an den genauen Monat wie auch das Jahr erinnern mussten, ein relativ niedriger Wert, der sicherlich durch die Zulassung von Jahreszeiten als Angabe noch verringert werden konnte.

Der prozentual gesehen höchste Anteil fehlender Angaben ist im Spell „Berufsvorbereitung, Berufsbildung, Berufsfachschule, die nicht zu einem Berufsabschluss führt“ zu finden. In 7% der Fälle konnten keine Angaben gemacht werden. Auf Rang zwei sind mit 6,7% die Trainingsmaßnahmen der BA zu finden, dicht gefolgt von der Grundschule mit 6,6%.

Der niedrigste Anteil von Verweigerungen bzw. fehlender Angaben ist dagegen bei der Aktivität „Wehr-/Zivildienst, FSJ“ zu finden. Nur 2,3% der Befragten mit dieser Aktivität konnten oder wollten hinsichtlich des Datums keine Angabe machen.

Der am häufigsten genannte Spell „Besuch allgemeinbildender Schulen“ kommt auf Quote von 4,9% und liegt damit in etwa im Durchschnitt.

## 6.4 Prüfung der Qualität der Interviews durch ALBERTA

Zur weiteren Prüfung der Datenqualität der Interviews wurde das Prüfprogramm ALBERTA eingesetzt. Hierbei handelt es sich um eine Eigenentwicklung der GfK mit dem vorrangigen Ziel, Manipulationen von persönlich (dezentral) durchgeführten Interviews zu entdecken.

Es lässt sich aber auch bei allen anderen Erhebungsformen einsetzen, dann allerdings weniger um Interviewer zu überprüfen, sondern vor allem die Qualität der gemachten Angaben in einem Interview. Hierbei handelt es sich um eine theorielose Herangehensweise, die regulär aus vier Tests besteht, wovon für die Übergangsstudie drei angewandt wurden:

- Varianzprüfung über alle Interviews hinweg (Übergreifende Konsistenzprüfung)
- Interviewervarianzprüfung
- Vergleich der Varianz in einzelnen Statementbatterien (Interne Konsistenzprüfung)

(Der vierte Test, die Prüfung auf gedoppelte Interviews, ist bei CATI-Erhebungen per se irrelevant.)

Diese Prüfungen werden anhand von ausgewählten Statement-Batterien vorgenommen. Für die Analyse des vorliegenden Datensatzes wurde jeweils eine Statement-Batterie am Anfang (v201n) sowie am Ende des Fragebogens (v5033n) herangezogen.

Die Analyse zeigt, dass insgesamt nur 37 Interviews Auffälligkeiten zeigten:

- 6 Interviews weisen in beiden Statement-Batterien jeweils eine starke Streuung auf
- 7 Interviews weisen in beiden Statement-Batterien jeweils eine schwache Streuung auf
- 11 Interviews zeigen eine starke übergreifende Konsistenz
- 13 Interviews zeigen generell eine geringe Varianz

Für eine ggf. weitergehende inhaltliche Analyse durch das BIBB wurden diese auffälligen Interviews im Datensatz mit der Variable ALBERTA gekennzeichnet.

## **6.5 Selektivitätsanalyse**

Ziel der Selektivitätsanalyse ist es zu untersuchen, inwieweit sich Personen, die das Interview vollständig durchgeführt haben, von denen unterscheiden, die gleich zu Beginn des Interviews eine Beantwortung des Fragebogens verweigert haben und zumindest einige wenige Fragen zur Soziodemografie beantworten wollten.

Dieser „Kurzfragebogen“ für die Verweigerer beinhaltet Fragen zu Alter, Geschlecht, Schulabschluss und Staatsangehörigkeit. Insgesamt haben 503 Verweigerer die Bereitschaft gezeigt, zumindest diese wenigen Fragen zur Soziodemografie zu beantworten.

Der soziodemographische Vergleich der beiden Gruppen – Verweigerer vs. Teilnehmer - ist in der nachfolgenden Tabelle zu sehen.

Tabelle 9: Soziodemografie der Verweigerer und Teilnehmer im Vergleich

<b>Merkmal</b>		<b>Verweigerer</b>	<b>Teilnehmer</b>
<b>Geschlecht</b>	Männlich	58%	52%
	Weiblich	42%	48%
<b>Geburtsjahr</b>	1987	22%	17%
	1988	24%	16%
	1989	13%	15%
	1990	13%	16%
	1991	11%	12%
	1992	10%	13%
	1993	7%	12%
<b>Staatsangehörigkeit</b>	Deutsch	93%	93%
	Andere	7%	7%
<b>Bildungsabschluss</b>	Niedrig	17%	20%
	Mittel	34%	32%
	Hoch	40%	40%
	kein Abschluss/k.A.	10%	9%
<b>Anzahl (n)</b>		<b>503</b>	<b>5.579</b>

Hinsichtlich des Geschlechtes zeigt sich, dass tendenziell mehr Männer die Beantwortung des Fragebogens verweigert haben: 58% der Verweigerer sind männlich, während es bei den Teilnehmern 52% sind.

Die Betrachtung des Geburtsjahres ergibt bei den Verweigerern einen deutlich überproportionalen Anteil der beiden ältesten Jahrgänge 1987 und 1988. 22% bzw. 24% der Verweigerer sind darin zu finden, während es unter den Teilnehmern 17% respektive 16% sind. In den übrigen Altersgruppen der Geburtsjahre 1989 bis 1993 sind dagegen die Personen stärker vertreten, die das Interview vollständig absolviert haben.



Im Hinblick auf die Staatsangehörigkeit zeigen sich dagegen überhaupt keine Unterschiede zwischen Verweigerern und Teilnehmern. Und auch bei der Betrachtung des Schulabschlusses sind die Unterschiede eher weniger groß und zeigen sich in erster Linie bei Personen mit niedrigem und mittlerem Bildungsabschluss.

Tabelle 10 zeigt welche im Rahmen der beschriebenen Datenprüfungen und -analysen zusätzlichen Variablen gebildet und dem BIBB im Datensatz zugespielt wurden

Tabelle 10: Zusätzliche im Datensatz zur Verfügung gestellte Variablen

<b>Variable</b>	<b>Inhalt</b>
ALBERTA	Durch ALBERTA-Prüfung auffällig
Alter_ges	Geburtsjahr overall
Bildung	Höchster Bildungsabschluss overall kategorisiert
Completed	Verweigerer versus Teilnehmer
Comp_Case	Verweigerer versus Teilnehmer Interviews auf Fallebene
Dauer_Int_Min	Dauer des Interviews in Minuten
Geschlecht	Geschlecht overall
Staat	Staatsangehörigkeit
ostwest	Gebietsdifferenzierung Ost-/Westdeutschland

## 7. Repräsentativität und Gewichtung

In der empirischen Sozialforschung entspricht die Struktur von Stichproben fast nie der Grundgesamtheit. Häufig sind einzelne Segmente der Grundgesamtheit in der Stichprobe über- bzw. unterrepräsentiert, wobei sich die Gründe meist der vollständigen Kontrolle entziehen. Mögliche Ursachen sind sowohl Charakteristika der Feldarbeit, aber auch unbekannte Unterschiede in der Erreichbarkeit einzelner Segmente der Grundgesamtheit. Strukturverzerrungen können aber auch durch Antwortverweigerungen sowie durch zufällige Abweichungen entstehen.

Deshalb ist es in der Regel notwendig, die Stichprobe zu gewichten, um repräsentative Schätzer für die Grundgesamtheit zu erhalten. Damit kann sichergestellt werden, dass Stichprobe und Grundgesamtheit in den vordefinierten Strukturvariablen übereinstimmen. Strukturvariablen sind wichtig zur Beschreibung der Grundgesamtheit.

In die vorliegende Studie werden folgende Strukturvariablen einbezogen, die auch für die Gewichtung maßgebend sind:

- Geschlecht
- Geburtsjahrgang
- Region
- Schulabschluss

Die Angaben zu den Verteilungen der Strukturvariablen in der Grundgesamtheit wurden den Daten der Mikrozensusbefragung 2010 entnommen. Dieser Entscheidung lag die Annahme zugrunde, dass der Mikrozensus die Gegebenheiten in der Grundgesamtheit valide abbildet.

Die Verwendung der Merkmalsverteilungen *Geburtsjahrgang*, *Geschlecht* und *Wohnregion* aus dem Mikrozensus war weitgehend unproblematisch. Es wurde dabei davon ausgegangen, dass es zwischen 2010 (dem Jahr der Mikrozensus-Erhebung) und 2011 (dem Jahr der BIBB-Übergangsstudie) zu keinen bedeutenden Veränderungen in den relativen Anteilen der Merkmalsausprägungen gekommen war, sich also z.B. die Anteile der Männer und Frauen nicht verschoben hatten.

Diese Annahme konnte jedoch *nicht* in Hinblick auf den *Schulabschluss* getroffen werden. Denn gerade für die Gruppe der jungen Erwachsenen bildet der Schulabschluss ein Merkmal, das sich im Laufe der Zeit noch substantiell verändern kann. Dies gilt insbesondere für die jüngeren Jahrgänge. Ein weiteres Problem bestand darin, dass die Mikrozensusbefragung anders als die BIBB-Übergangsstudie mehr oder weniger gleichmäßig über alle Monate des Jahres durchgeführt wurde. Die BIBB-Studie fand dagegen ausschließlich im dritten Jahresquartal statt. Zudem wurde im Rahmen der Mikrozensusbefragung der höchste Schulabschluss nur für diejenigen Personen erfasst, welche nicht mehr die allgemeinbildende Schule besuchten. Somit wurden die Bildungsabschlüsse im Mikrozensus für Personen desselben Geburtsjahrgangs und desselben Bildungsziels im Jahr 2010 (z.B. Erwerb des Abiturs) nicht einheitlich erhoben. Denn die einen Probanden – in den ersten Monaten des Jahres 2010 befragt – hatten diesen Abschluss noch nicht erreicht (so dass für sie gegebenenfalls nur die Information zur aktuell besuchten allgemeinbildenden Schulform vorlag), während die anderen – in den späteren Monaten des Jahres 2010 befragt – die allgemeinbildende Schulzeit bereits beendet hatten und somit ihren Schulabschluss bereits mitteilen konnten.

Eine Information zum gegenwärtig höchsten Schulabschluss war deshalb im Rahmen des Mikrozensus nicht für alle Probanden zu gewinnen. Es konnte aus den Daten des Mikrozensus lediglich eine „kombinierte Schulform-/Schulabschlussvariable“ gebildet werden, deren Ausprägungen zum Teil nur die aktuell besuchte Schulform repräsentierten.

Aus den methodischen Abweichungen zwischen dem Mikrozensus und der BIBB-Übergangsstudie resultierten naturgemäß (und unabhängig von möglichen Stichprobenverzerrungen im Zusammenhang mit der BIBB-Übergangsstudie) deutliche Unterschiede zwischen der Verteilung der Schulabschlüsse: In der BIBB-Übergangsstudie mussten die Anteile der Probanden mit *höheren* Schulabschlüssen zwangsläufig wesentlich größer ausfallen. Denn einerseits war seit der letzten Mikrozensus-Erhebung bereits ein Jahr vergangen, und andererseits war die BIBB-Übergangsstudie erst zu Beginn der zweiten Jahreshälfte gestartet worden (und somit zu einem Zeitpunkt, zu dem die meisten Jugendlichen, die in diesem Jahr einen höheren Schulabschluss wie z.B. das Abitur erwerben wollten, diesen Schulabschluss auch bereits erreicht hatten).<sup>4</sup>

Die Probleme mit dem Schulabschluss als Gewichtungmerkmal ließen sich letztlich nur dadurch zufriedenstellend lösen, dass im Rahmen der BIBB-Übergangsstudie 2011 jene Befragungsverhältnisse im Jahre 2010 simulieren wurden, wie sie für den Mikrozensus galten. Dies war grundsätzlich möglich, da infolge des retrospektiven Längsschnittcharakters der BIBB-Übergangsstudie theoretisch für jeden Monat seit Eintritt in die Grundschule der jeweilige schulische Status der Probanden (Art des bis dato erreichten höchsten Schulabschlusses bzw. ggfs. auch der gerade besuchten Schulform) recherchiert werden konnte. Somit ließ sich auch das Problem entschärfen, das sich aus der unterjährig kontinuierlichen Datenerhebung des Mikrozensus ergeben hatte.

Zur Konstruktion dieser neuen Schulform-/Schulabschlussvariable für die Probanden der BIBB-Übergangsstudie (mit Simulation des Mikrozensus-Befragungsmodus im Jahr 2010) wurde wie folgt vorgegangen. Zunächst wurden die Probanden der BIBB-Übergangsstudie nach dem Zufallsprinzip auf 12 imaginäre „Befragungs“-Monate des Jahres 2010 verteilt. Wurde nun ein Proband z.B. dem Monat „Januar 2010“ zugeordnet, wurde für diesen Monat ermittelt, ob er sich zu diesem Zeitpunkt

---

<sup>4</sup> Deshalb konnte es auch keine zufriedenstellende Lösung sein, alternativ die Mikrozensus-Ergebnisse 2010 zu den Geburtsjahrgängen 1986-1991 heranzuziehen (die 2010 zwischen 18 und 24 Jahre alt waren) und diese Verteilung zur Abschätzung Schulabschlüsse der 18- bis 24-jährigen im Jahr 2011 zu verwenden. Zusätzliche Probleme hätten sich bei einem solchen Lösungsversuch ohnehin daraus ergeben, dass 2011 in Bayern und Niedersachsen doppelte Abiturientenjahrgänge die Schulen verlassen hatten und dort erstmalig ein Geburtsjahrgang zur Hochschulreife geführt worden war, der bereits nach 12 Schuljahren bzw. mit im Schnitt 18 Lebensjahren das Abitur erworben hatte.

in einer allgemeinbildenden Schule befand. War dies der Fall, wurde die entsprechende Schulform bestimmt, und diese Information ging – analog zum Mikrozensus – in die neue Schulform-/Schulabschlussvariable ein. Hatte der Proband die allgemeinbildende Schule zu diesem Zeitpunkt bereits verlassen, wurde sein bis dato höchster Schulabschluss rekonstruiert, und diese Information wurde verwendet, um die passende Ausprägung innerhalb der neuen Schulform-/Schulabschlussvariable zu finden.

Mit diesem Verfahren war ein unmittelbarer und weitgehend valider Vergleich der Verteilungen der Schulabschlüsse in der BIBB-Übergangsstudie und im Mikrozensus mit dem Bezugsjahr 2010 möglich (vgl. Tabelle 11)

Tabelle 11: Nicht angepasste Verteilung der Schulabschlüsse im Mikrozensus und in der BIBB-Übergangsstudie 2011 nach Geburtsjahrgängen

Kombinierte Schulform-/Schulabschlussvariable	Geburtsjahrgänge						Insgesamt		
	1987-1989			1990-1993			BIBB	MZ	Diff.
	BIBB	MZ	Diff.	BIBB	MZ	Diff.			
Haupt-/Volksschulabschluss	12,4	19,3	-7,0	16,1	16,5	-0,3	14,4	17,8	-3,4
Realschulabschluss	32,0	33,1	-1,1	31,7	27,9	3,8	31,8	30,2	1,6
Fachhochschulreife, Hochschulreife	49,0	42,4	6,7	16,2	11,3	4,9	31,7	25,4	6,2
Gegenwärtiger Besuch einer allgemeinbildenden Schule	1,4	1,7	-0,3	31,2	38,6	-7,5	17,1	21,8	-4,7
darunter:									
▪ Hauptschule	0,1	0,0	0,0	1,4	2,1	-0,8	0,8	1,2	-0,4
▪ Realschule	0,0	0,1	0,0	3,4	5,6	-2,2	1,8	3,1	-1,3
▪ Schulartunabhängige Orientierungsstufe, Schularten mit mehreren Bildungsgängen (z.B. Sekundarschule, Mittelschule)	0,0	0,0	0,0	0,0	0,8	-0,8	0,0	0,4	-0,4
▪ Gesamtschule, Waldorfschule	0,5	0,3	0,3	4,8	3,1	1,7	2,8	1,8	1,0
▪ Gymnasium	0,8	1,4	-0,6	21,5	26,3	-4,8	11,7	14,9	-3,2
▪ Sonderschule (Förderschule)	0,0	0,0	0,0	0,2	0,8	-0,6	0,1	0,4	-0,3
Kein Abschluss, Sonstige	5,1	3,4	1,7	4,9	5,7	-0,9	5,0	4,7	0,3
Insgesamt	100,0	100,0	0,0	100,0	100,0	0,0	100,0	100,0	0,0
Fallzahl (Mikrozensus: in Tsd.)	2.635	2.972	.	2.944	3.551	.	5.579	6.523	.

Abkürzungen: BIBB = BIBB-Übergangsstudie, MZ = Mikrozensus 2010, Diff. = Prozentpunktdifferenz

Quellen: Statistisches Bundesamt, BIBB-Übergangsstudie 2011

Wie die Übersicht zeigt, ließen sich größere Abweichungen zwischen der BIBB-Übergangsstudie und dem Mikrozensus vor allem in Hinblick auf die Anteile der Probanden mit Fachhochschul-/Hochschulreife feststellen (die Quote war im Mikrozensus deutlich niedriger). Dies lag jedoch zum Teil auch daran, dass der Anteil derjenigen, die zum „Befragungszeitpunkt“ ein Gymnasium be-

suchten (und somit das Abiturzeugnis noch nicht erworben hatten), im Mikrozensus ungeachtet der oben beschriebenen Anpassungen höher als in der BIBB-Übergangsstudie ausfiel.

Vereinigt man aber die beiden Kategorieformen („Schulabschluss bereits erworben“ und „aktueller Besuch einer allgemeinbildenden Schule, die zum entsprechenden Schulabschluss führt“) jeweils zu *einer* Variablenausprägung, so fallen die Differenzen zwischen Mikrozensus und BIBB-Übergangsstudie nur noch relativ gering aus (vgl. Tabelle 12)

Tabelle 12: Angepasste Verteilung der Schulabschlüsse im Mikrozensus und in der BIBB-Übergangsstudie 2011 nach Geburtsjahrgängen

	Geburtsjahrgänge						Insgesamt		
	1987-1989			1990-1993			BIBB	MZ	Diff.
	BIBB	MZ	Diff.	BIBB	MZ	Diff.			
▪ Haupt-/Volksschulabschluss (inklusive aktueller Besuch entsprechender Schulformen im allgemeinbildenden Sektor)	12,4	19,4	-7,0	17,5	18,6	-1,1	15,1	19,0	-3,9
▪ Realschulabschluss (inklusive aktueller Besuch entsprechender Schulformen im allgemeinbildenden Sektor)	32,6	33,4	-0,8	39,8	37,3	+2,5	36,4	35,6	+0,8
▪ Studienberechtigung (inklusive aktueller Besuch entsprechender Schulformen im allgemeinbildenden Sektor)	49,8	43,7	+6,1	37,7	37,6	+0,1	43,4	40,4	+3,0
▪ ohne Abschluss, sonstige	5,2	3,4	+1,8	5,1	6,5	-1,4	5,1	5,1	0,0
Insgesamt	100,0	100,0	0,0	100,0	100,0		100,0	100,0	0,0
Fallzahl (Mikrozensus: in Tsd.)	2.635	2.972	.	2.944	3.551	.	5.579	6.523	.

Abkürzungen: BIBB = BIBB-Übergangsstudie, MZ = Mikrozensus 2010, Diff. = Prozentpunktedifferenz

Quellen: Statistisches Bundesamt, BIBB-Übergangsstudie 2011

Bezogen auf die Geburtsjahrgänge 1990 bis 1993 sind die Unterschiede nun sogar nahezu vernachlässigenswert. Lediglich bei den Älteren, den Geburtsjahrgängen 1987 bis 1989 zeigt sich, dass im Rahmen der BIBB-Übergangsstudie Personen mit maximal Hauptschulabschluss „zu selten“ befragt worden waren (12,4% vs. 19,4% im Mikrozensus) und Studienberechtigte „zu häufig“ an der Untersuchung teilgenommen hatten (49,8% vs. 43,7% im Mikrozensus). Die Anteile der Befragten mit mittlerem Abschluss sind dagegen fast identisch.

Die auf diese Weise durchgeführten Berechnungen lassen somit erkennen, dass die Stichprobenziehung im Rahmen der BIBB-Übergangsstudie 2011 insgesamt gesehen nur zu relativ leichten Verzerrungen zugunsten höherer Schulabschlüsse führte. Und auch die übrigen Gewichtsmerkmale wichen in ihren Verteilungen nur wenig von den Verteilungen innerhalb des Mikrozensus ab (vgl. Tabelle 11).

Tabelle 12: Grundlage für die Anpassungsgewichtung

	<b>Stichprobe in %</b>	<b>Grundgesamtheit (Mikrozensus 2010) in %</b>	<b>Differenz %-Punkte</b>
<b>Geschlecht</b>			
Männlich	51,8	51,4	0,4
Weiblich	48,2	48,6	-0,4
<b>Jahrgang</b>			
1987	16,7	15,3	1,4
1988	15,8	15,4	0,4
1989	14,6	14,9	-0,3
1990	15,8	14,9	0,9
1991	12,2	13,7	-1,5
1992	13,1	13,2	-0,1
1993	11,7	12,6	-0,9
<b>Region*</b>			
Ost	20,1	18,3	1,8
Nordrhein-Westfalen	20,2	22,3	-2,1
Nord	15,5	16,2	-0,7
Mitte	14,6	13,4	1,2
Baden-Württemberg	12,4	14,1	-1,7
Bayern	17,1	15,7	1,4
Sonstiges	0,1	-	
<b>Schulabschluss</b>			
Hauptschule	15,1	19,0	-3,9
Mittlerer Abschluss	36,4	35,6	0,8
Studienberechtigung	43,4	40,4	3,0
ohne Abschluss, sonstige	5,1	5,1	0,0

\*Ost = Ostdeutschland (neue Länder, inkl. Berlin)

Nord = Schleswig-Holstein, Hamburg, Bremen, Niedersachsen

Mitte = Rheinland-Pfalz, Saarland, Hessen

Der Vergleich der demografischen Verteilung von Stichprobe und Grundgesamtheit zeigt, dass im Großen und Ganzen die Differenzen nicht sehr groß sind, was ein gute Qualität der vorliegenden Stichprobe belegt.

Hinsichtlich des Alters der Befragten ist in der Stichprobe ein leicht höherer Anteil der älteren Jahrgänge festzustellen, während die jüngeren Altersgruppen etwas unterrepräsentiert sind. In der regionalen Differenzierung zeigt sich ein etwas höherer Anteil in Ostdeutschland (einschl. Berlin), in

der Mitte sowie in Bayern, während vor allem Nordrhein-Westfalen und Baden-Württemberg leicht unterdurchschnittliche Anteile aufweisen.

Beim Schulabschluss ist ein Bias zu erkennen, der in ähnlichen Umfragen generell zu Tage tritt. Während der Anteil der Personen mit mittlerem Abschluss in etwa den Angaben aus der Grundgesamtheit entspricht, sind die Personen mit Hauptschulabschluss in der Stichprobe leicht unterrepräsentiert und Personen mit Studienberechtigung entsprechend überdurchschnittlich häufig vertreten.

## **7.1 Designgewichtung**

Der Anpassungsgewichtung vorgeschaltet ist standardmäßig die Designgewichtung. Grundlage für die Stichprobe waren Mobilfunkrufnummern, die zu Personen gehören, die aus der ADM Auswahlgrundlage Mobilfunk gezogen wurden. Da aus anderen Untersuchungen bekannt ist, dass Zielpersonen unter mehreren Mobilfunknummern zu erreichen sind, haben sie eine höhere Auswahlwahrscheinlichkeit als Personen mit nur einer Mobilfunknummer. So hat eine Person mit drei gültigen Mobilfunknummern die dreifache Wahrscheinlichkeit, in die Stichprobe zu gelangen.

Im Fragebogen wurde deshalb erfasst, unter wie vielen Rufnummern (mobil) die Zielperson erreichbar ist. Das Designgewicht errechnet sich nach folgender Formel:

$$\text{Designgewicht} = 1/\text{Anzahl der Mobilfunknummern}$$

## 7.2 Anpassungsgewichtung

Nach Berücksichtigung der unterschiedlichen Auswahlwahrscheinlichkeiten im Rahmen der Designgewichtung wurde die Anpassungsgewichtung vorgenommen<sup>5</sup>.

Ziel der Gewichtung ist es, die jeweilige Struktur der relevanten Grundgesamtheit in der Stichprobe zu reproduzieren. Um dies zu erreichen, werden den Fällen individuelle Gewichte so zugeordnet, dass die gewichtete Stichprobe die Vorgaben erfüllt. Das heißt, dass die Summe der Gewichte eines Segmentes (z.B. alle Fälle in einer Altersklasse oder Region) der gewünschten Segmentgröße innerhalb der Stichprobe entsprechen muss. In der Regel müssen mehrere solcher Vorgaben gleichzeitig erfüllt sein.

Mit der Gewichtung werden systematische Verzerrungen aus der Stichprobe entfernt. Während der systematische Fehler verringert wird, vergrößert die Gewichtung in der Regel allerdings den Zufallsfehler. Die Zunahme des Zufallsfehlers kann auch als Abnahme der sogenannten effektiven Stichprobengröße (Fallzahl) beschrieben werden<sup>6</sup>. Die effektive Stichprobengröße ist die gebräuchlichste Form der Quantifizierung der Gewichtungskosten und gilt auch bei der vorliegenden Untersuchung als Maßstab.

Für die vorliegende Studie wurde die Lineare Gewichtung verwendet.

Die meisten Gewichtungsalgorithmen sind iterativ. Beispiele hierfür sind Iterative Proportional Fitting IPF, Propensity Weighting und CALMAR. Diese Methoden verwenden unterschiedlichste Verfahren, haben aber alle gemeinsam, dass sie im Allgemeinen weder die beste Gewichtungslösung finden noch die Gewichtungsvorgaben perfekt reproduzieren.

Im Gegensatz dazu erzeugt die Lineare Gewichtung eine numerisch perfekte Übereinstimmung mit den Gewichtungsvorgaben und findet die rechnerisch beste Lösung. Bei dem eingesetzten Verfahren handelt es sich um einen eigenentwickelten Algorithmus der GfK, der die optimale Gewichtungslösung analytisch findet bzw. die effektive Fallzahl maximiert. Da es sich um das analytische Auffinden der besten Gewichtungslösung handelt, kann es unseres Ermessens kein anderes Verfahren geben, um eine bessere Lösung zu finden.

---

<sup>5</sup> Die folgenden Ausführungen wurden im Wesentlichen folgender Publikation entnommen: Volker Bosch, Die lineare Gewichtung der GfK, 18. Januar 2008, volker.bosch@gfk.com

<sup>6</sup> Dies wurde von Kish wegweisend definiert: Kish, L. Questions and Answers. The Survey Statistician, 1987, 17, 13-17.



Folgende Vorgaben aus dem Mikrozensus 2010 wurden für die Anpassungsgewichtung herangezogen:

Tabelle 13: Mikrozensusangaben zu Region x Geschlecht x Alter

	Geburtsjahr							<b>Gesamt</b>
	1987	1988	1989	1990	1991	1992	1993	
Ost, M	1,8%	1,7%	1,7%	1,6%	1,0%	0,9%	0,8%	9,6%
Ost, W	1,6%	1,6%	1,5%	1,5%	1,0%	0,8%	0,7%	8,7%
NRW, M	1,7%	1,7%	1,5%	1,7%	1,6%	1,6%	1,5%	11,3%
NRW, W	1,6%	1,6%	1,5%	1,6%	1,6%	1,6%	1,5%	11,0%
Nord, M	1,2%	1,2%	1,2%	1,2%	1,2%	1,2%	1,2%	8,4%
Nord, W	1,1%	1,2%	1,1%	1,0%	1,1%	1,2%	1,1%	7,8%
Mitte, M	1,0%	1,0%	0,9%	1,0%	1,0%	0,9%	1,0%	6,8%
Mitte, W	1,0%	0,9%	1,0%	1,0%	1,0%	0,9%	0,8%	6,5%
BW, M	1,0%	1,1%	1,0%	1,0%	1,0%	1,1%	0,9%	7,3%
BW, W	0,9%	1,1%	1,1%	1,0%	1,0%	0,9%	0,9%	6,9%
Bayern, M	1,2%	1,2%	1,2%	1,1%	1,1%	1,1%	1,1%	8,0%
Bayern, W	1,2%	1,2%	1,1%	1,1%	1,0%	1,1%	1,0%	7,7%
Sonstiges, M	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Sonstiges, W	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
<b>Gesamt</b>	15,3%	15,4%	14,9%	14,9%	13,7%	13,2%	12,6%	100,0%

M=Männlich

W=Weiblich

Tabelle 14: Mikrozensusangaben zu Schulabschluss x Alter

	Geburtsjahr		<b>Gesamt</b>
	1987-1989	1990-1993	
Hauptschule	8,8%	10,1%	19,0%
mittlerer Abschluss	15,2%	20,3%	35,6%
Studienberechtigung	19,9%	20,5%	40,4%
ohne Abschluss, sonstige	1,6%	3,5%	5,1%
<b>Gesamt</b>	45,6%	54,4%	100,0%

Es wurden zwei Varianten gerechnet:

**Variante 1: Berücksichtigung der Fälle mit Angabe „Sonstiges“ bei Bundesland**

Hierzu wurde der ungewichtete Anteil an Nennungen „sonstige“ beim Bundesland aus der Stichprobe übernommen und die Vorgaben des Mikrozensus entsprechend umprozentuiert. Es kam dabei zu sehr geringfügigen Abweichungen von den Vorgaben, da durch die Berücksichtigung der „Sonstigen“ bei der Bundeslandangabe die Verteilung Geschlecht x Altersgruppensplit nicht mehr exakt erfüllt werden konnte. Der Grund lag darin, dass im Datensatz in der Zelle „weiblich & sonstige Angabe bei Bundesland & Alterssplit 1990-1993“ keine Fälle realisiert wurden. Bei der ebenfalls durchgeführten Hochrechnung auf 6.523.000 Personen wurden die Fälle ohne Bundeslandangabe nicht berücksichtigt.

Tabelle 15 führt die Gewichtungs- bzw. Hochrechnungsvariablen auf, die für den Datensatz gebildet wurden.

Tabelle 15: Berechnete Gewichte unter Berücksichtigung der sonstigen-Kategorie bei Bundesland

<b>Variable</b>	<b>Beschreibung</b>
W0_mobile	Designgewicht Handy (Normierung auf Stichprobengröße)
Weight_w_mobile	Gewicht unter Berücksichtigung des Designgewichtes
Weight_wo_mobile	Gewicht ohne Berücksichtigung des Designgewichtes
pweight_w_mobile	Hochrechnungsgewicht unter Berücksichtigung Designgewicht
pweight_wo_mobile	Hochrechnungsgewicht ohne Berücksichtigung Designgewicht

Rim= 0,94, d.h. die effektive Fallzahl beträgt  $n_{eff}=5.244$  ( $5.579 \times 0,94$ ).

Die Güte der Gewichtung wird anhand der Rim Effizienz (rim) angegeben. Sie zeigt an, wie ausgewogen die Stichprobe ist. Sie gibt an, wie groß der Anteil der erhobenen Information (effektive Fallzahl) ist, der bei der Gewichtung erhalten geblieben ist. Ein Wert von 0,94 besagt, dass 94% der 5.579 Fällen als sogenannte effektive Fallzahl infolge der Gewichtung erhalten geblieben sind (Effektive Fallzahl somit  $n_{eff}=5.244$ ).

Als grober Richtwert gilt, dass bei proportionalen Stichproben, wie im vorliegenden Fall, die Rim Effizienz nicht unter 0,7 fallen sollte<sup>7</sup>.

---

<sup>7</sup> Vgl. hierzu: Bosch, Volker, Effektive Fallzahl und Rim-Effizienz bei varianzhomogenen Stichproben, 2003

**Variante 2: Keine Berücksichtigung der Fälle mit Angabe „Sonstiges“ bei Bundesland**

Hierbei mussten keine Umprozentuierungen vorgenommen werden.

Tabelle 16 führt die Gewichtung- bzw. Hochrechnungsvariablen auf, die für den Datensatz gebildet wurden.

Tabelle 16: Berechnete Gewichte ohne Berücksichtigung der sonstigen-Kategorie bei Bundesland

<b>Variable</b>	<b>Beschreibung</b>
w0_mobile_wo_others	Designgewicht Handy (Normierung auf Stichprobengröße)
weight_w_mobile_wo_others	Gewicht unter Berücksichtigung des Designgewichtes
weight_wo_mobile_wo_others	Gewicht ohne Berücksichtigung des Designgewichtes
pweight_w_mobile_wo_others	Hochrechnungsgewicht unter Berücksichtigung Designgewicht
pweight_wo_mobile_wo_others	Hochrechnungsgewicht ohne Berücksichtigung Designgewicht

Rim= 0,94, d.h. die effektive Fallzahl beträgt  $n_{eff}=5.240$ . Konkret heißt das, dass durch die Gewichtung 94% der Information erhalten geblieben ist und nur etwa 6% verloren gingen.

Das vorliegende Ergebnis von 0,94 für beide Varianten liegt deutlich über den als groben Richtwert geforderten 0,7 und gilt daher als exzellenter Wert, was ein weiterer Beleg für eine ausgesprochen hohe Qualität der Stichprobe ist.

## 8. Gesamtbewertung

Die Umstellung des Stichprobendesigns von einer Festnetz-Stichprobe zu einer Mobilfunk-Stichprobe in dieser Erhebung ist angesichts der Mobilität der Jugendlichen und jungen Erwachsenen notwendig und richtig gewesen. In verschiedenen aktuellen Untersuchungen lässt sich beobachten, dass gerade junge Menschen, die die allgemeinbildende Schule verlassen haben, schwer über das Festnetz zu erreichen sind, wenn sie denn überhaupt noch über einen Festnetzanschluss zum Telefonieren verfügen.

Aufgrund der sehr geringen Inzidenz der definierten Zielgruppe aber auch der schwierigen Erreichbarkeit von Zielpersonen in einem Mobilfunksample, z.B. aufgrund von automatischen Ansagen bzw. Begrüßungstexten von Mailboxen, ist der Aufwand von reinen Mobilfunkstichproben vergleichsweise hoch. Dennoch muss der Methodenwechsel bei der Übergangsstudie von einer Festnetz-Stichprobe auf ein Mobilfunk-Sample insgesamt positiv bewertet werden. Dies zeigen sowohl die diversen Qualitätsprüfungen des Datensatzes, die vorgenommen worden sind, als auch die sehr hohen Gütemaße im Rahmen der Gewichtung (Rim Effizienz). Die Ergebnisse können damit als überaus valide und zuverlässig angesehen werden.

GfK SE

Social & Economic Research

Nürnberg, 21. Dezember 2011



Methodenbericht BIBB-Übergangsstudie 2011

**Anhang:**

**Finale Fragebogenversion**