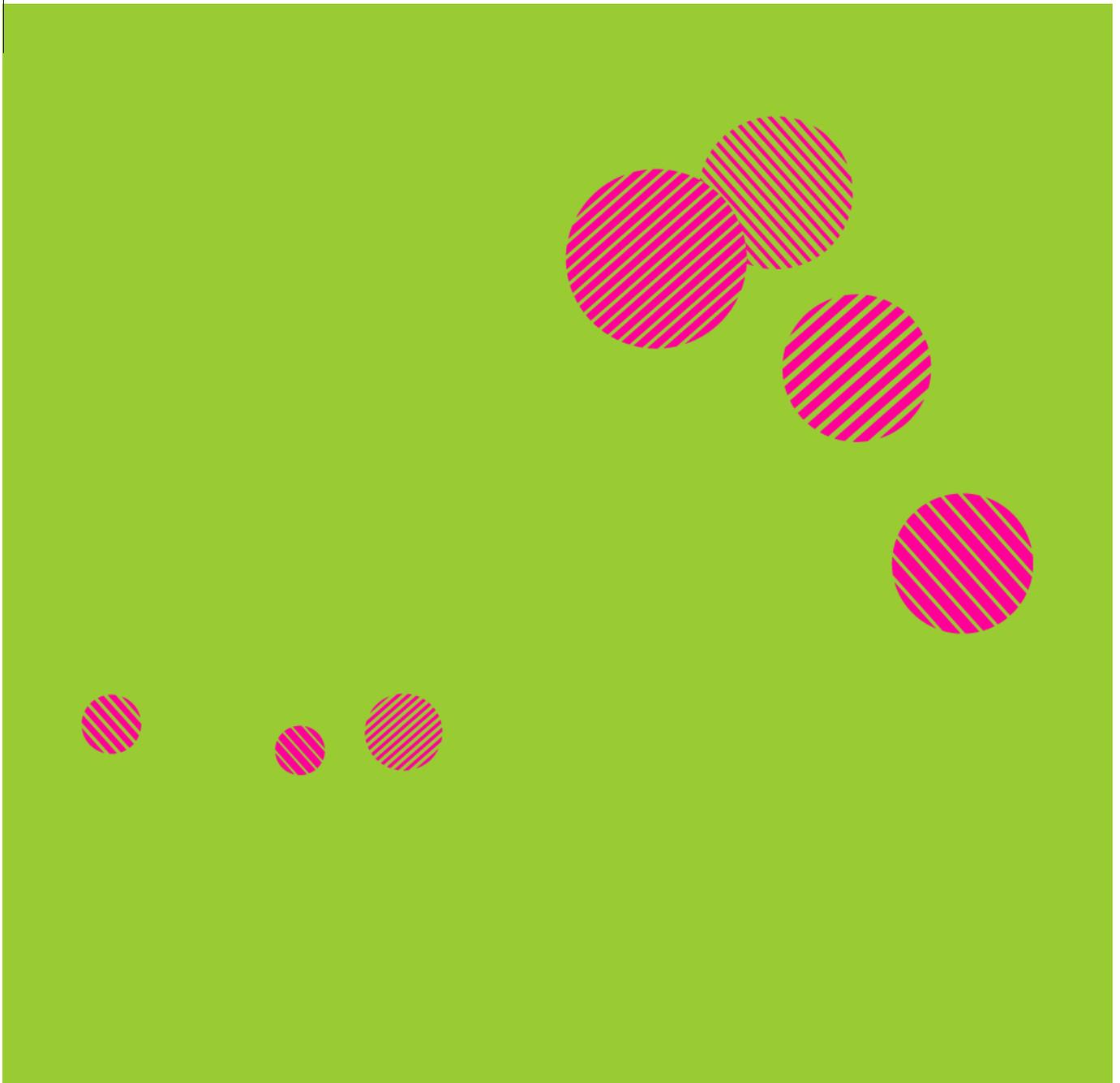


Methodenbericht

Bevölkerungsbefragung zum Thema
„Sicherheit und Kriminalität in Deutschland 2020“

Autoren: Stefan Schiel, Susann Südhof, Jonathan Ruiz-Marcos, Jan Niesmann, Moritz Fahrenholz



Bericht an

Bundeskriminalamt
Äppelallee 45
65173 Wiesbaden

Vorgelegt von

infas Institut für angewandte
Sozialwissenschaft GmbH
Friedrich-Wilhelm-Straße 18
53113 Bonn

Kontakt

Stefan Schiel
Senior-Projektleiter Sozialforschung

Tel. +49 (0)228/38 22-424
Fax +49 (0)228/310071
E-Mail s.schiel@infas.de

Autoren

Stefan Schiel, Susann Südhof, Jonathan Ruiz-Marcos,
Jan Niesmann, Moritz Fahrenholz

© infas Institut für angewandte Sozialwissenschaft GmbH
Der Inhalt dieses Berichts darf ganz oder teilweise
nur mit unserer schriftlichen Genehmigung veröffentlicht,
vervielfältigt, gedruckt oder in Informations- und
Dokumentationssystemen (information storage and
retrieval systems) gespeichert, verarbeitet oder ausgegeben
werden.

Projekt

7204
Bonn, September 2021
SI, Ds, Rj

infas ist zertifiziert
nach ISO 20252 für die Markt-,
Meinungs- und Sozialforschung

ISO 20252



infas ist Mitglied im
Arbeitskreis Deutscher Markt- und
Sozialforschungsinstitute e.V.
(ADM) und ESOMAR

ADM

ESOMAR
member

Vorbemerkung

Die infas Institut für angewandte Sozialwissenschaft GmbH wurde vom Bundeskriminalamt (BKA) beauftragt, im Jahr 2020 eine bundesweit repräsentative Bevölkerungsbefragung zum Thema „Sicherheit und Kriminalität in Deutschland (SKiD)“ durchzuführen. Bevölkerungsbefragungen zu Erfahrungen mit Kriminalität und zur Wahrnehmung von Sicherheit haben sich als wichtiges Mittel bewährt, die Verbreitung und Folgen von Kriminalität abzuschätzen und besser zu verstehen. In Deutschland fanden solche Befragungen bisher vor allem auf der kommunalen oder der Landesebene statt. In den Jahren 2012 und 2017 hat das BKA mit dem Deutschen Viktimisierungssurvey (DVS) erstmals seit längerer Zeit große bundesweite Befragungen dieser Art durchgeführt. Die Datenerhebung der beiden großen Telefonbefragungen übernahm infas.

Periodisch wiederholte Opferbefragungen ergänzen sinnvoll die Polizeiliche Kriminalstatistik (PKS), welche ausschließlich das Hellfeld, also der Polizei bekannt gewordenen Straftaten abbildet. Mit SKiD wird nun erstmals eine periodisch angelegte, bundesweit repräsentative Befragung in Deutschland etabliert.

Die Studie war als kombinierte Erhebung mit Papierfragebogen und Online-Fragebogen konzipiert. Im Hauptfeld wurde eine bundesweite Stichprobe von 122.667 in Deutschland lebenden Personen ab 16 Jahren eingesetzt. Zusätzlich war ein Methodenexperiment mit einer Stichprobengröße der Experimentalgruppe von 998 Fällen derselben Population vorgesehen.

Der Fragebogen wurde vom BKA und weiteren beteiligten Landeskriminalämtern entwickelt. infas beteiligte sich aktiv an der Endredaktion des Fragebogens und übernahm die Umsetzung der Fragen in ein professionelles Design. Die Feldfähigkeit des Fragebogens wurde einem quantitativen und mehreren qualitativen Pretests unterzogen. Der Fragebogen wurde in Deutsch (Papier und online) und in den Sprachen Türkisch (Papier und online) sowie Russisch und Arabisch (jeweils online) eingesetzt. Nach Prüfung aller eingehenden Fragebögen verblieben 45.351 auswertbare Fälle. 37.860 Fragebögen der Papierfassung (83,5 Prozent) und 7.491 Online-Fragebögen (16,5 Prozent) sind auswertbar.

Der vorliegende Bericht dokumentiert das Erhebungsdesign, die Durchführung und das Ergebnis der Haupterhebung.

infas Institut für angewandte Sozialwissenschaft GmbH

Inhaltsverzeichnis

1 Erhebungsdesign	7
2 Stichprobenziehung	9
2.1 Grundgesamtheit	9
2.2 Aufgabenteilung bei der Stichprobenziehung	9
2.3 Auswahlstufe 1: Ziehung der Gemeinden/Pointstichprobe	10
2.4 Auswahlstufe 2: Ziehung der Einsatzstichprobe aus den Registern der Einwohnermeldeämter	11
2.5 Umgang mit nicht teilnahmebereiten Gemeinden	12
2.6 Zusammenführung der Bundeslandstichproben	14
2.7 Onomastische Klassifizierung der Bruttostichprobe	15
2.8 Stichprobenumfang Methodenexperiment	16
3 Erhebungsunterlagen	17
3.1 Der Papierfragebogen als Leitinstrument	17
3.2 Der Online-Fragebogen (CAWI)	18
3.3 Weitere Erhebungsunterlagen	20
3.4 Fremdsprachige Erhebungsunterlagen	22
3.4.1 Übersetzungen	23
3.4.2 Kognitiver Pretest: Prüfung fremdsprachiger Erhebungsinstrumente	23
4 Durchführung der Hauptstudie	25
4.1 Feldzeit	25
4.2 Kommunikationskonzept	25
4.3 Feldsteuerung	27
4.4 Feldverlauf	30
4.5 Verarbeitung von Rückmeldungen aus dem Feld	32
4.6 Stichprobenausschöpfung	33
4.6.1 Basisstichprobe	33
4.6.2 Methodenexperiment	36
4.7 Interviewdauer CAWI-Fragebogen	38
4.8 Abbrüche im Fragebogen	38
5 Datenerfassung, -prüfung und -bereinigung	43
5.1 Eingangserfassung Papierfragebögen	43
5.2 Inhaltliche Datenerfassung	44
5.3 Datenprüfung und -bereinigung	44
5.3.1 Papierfragebogen	44
5.3.2 Online-Fragebogen	45
5.3.3 Umgang mit doppelt ausgefüllten Papier- und Online-Fragebögen	45
5.4 Ausschluss nicht verwertbarer Fälle	46
5.5 Fragebogendokumentation	47
6 Gewichtung	48
6.1 Designgewichtung	48
6.2 Ausfallgewichtung (Nonresponse-Adjustierung)	50
6.3 Kalibrierung (Randanpassung)	52
6.4 Kennwerte der Gewichte	56
Literaturverzeichnis	58

Abbildungsverzeichnis

Abbildung 1	Feldverlauf Papier- und Online-Fragebögen kumuliert	30
Abbildung 2	Fragebogeneingang pro Feldwoche	31

Tabellenverzeichnis

Tabelle 1	Überblick über das Studiendesign	7
Tabelle 2	Stichprobendesign und -beschaffung nach Bundesländern	10
Tabelle 3	Kontaktierte Gemeinden nach Bundesland	12
Tabelle 4	Gemeindeverteilung laut Stichprobenplan und in Einsatzstichprobe	13
Tabelle 5	Umfang Bruttostichprobe	14
Tabelle 6	Ergebnis der onomastischen Qualifizierung der Bruttostichprobe	15
Tabelle 7	Verteilung der Stichproben für Methodenexperiment und Kontrollgruppe nach Bundesländern	16
Tabelle 8	Zuordnung Fremdsprache und Erhebungsunterlagen	22
Tabelle 9	Fallzahlen pro Versandaktion	28
Tabelle 10	Versandtermine	29
Tabelle 11	Zugehörigkeit der Bundesländern zu Versandgruppe	29
Tabelle 12	Stichprobenausschöpfung	34
Tabelle 13	Vergleich Bruttostichprobe und auswertbare Fälle nach Bundesland	35
Tabelle 14	Auswertbare Fälle getrennt nach PAPI/CAWI	36
Tabelle 15	Auswertbare Fragebögen nach Sprache und Methode	36
Tabelle 16	Stichprobenausschöpfung	37
Tabelle 17	PAPI versus CAWI-Rücklauf getrennt nach Stichproben	37
Tabelle 18	Ausfülldauer vollständig bearbeitete CAWI-Fragebogen	38
Tabelle 19	Abbruchstellen im Papierfragebogen	39
Tabelle 20	Abbruchstellen im CAWI-Fragebogen	41
Tabelle 21	Ergebnis Ausfallmodell	51
Tabelle 22	Verteilungsvergleich gewichtete Stichprobe und Grundgesamtheit	54
Tabelle 23	Kennwerte der Gewichte	57

1 Erhebungsdesign

Viktimisierungsbefragungen sind ein bewährtes Mittel zur Beurteilung der Kriminalitätslage und des Sicherheitsempfindens der Bevölkerung. Hierzu wurden zuletzt in den Jahren 2012 und 2017 mit dem Deutschen Viktimisierungssurvey bundesweite Erhebungen in Deutschland durchgeführt. Zur langfristigen Beurteilung der Entwicklung von Kriminalität und Sicherheit sind jedoch regelmäßig wiederholte Befragungen nötig. Mit der Studie „Sicherheit und Kriminalität in Deutschland (SKiD) 2020“ soll solch eine periodisch angelegte, bundesweit repräsentative Erhebung etabliert werden.

Schwerpunkte der Befragung waren Einschätzungen zur Polizeiarbeit, Kriminalitätsfurcht, Opfererfahrungen und das Anzeigeverhalten der Bevölkerung. Für die SKiD-Befragung wurde als Erhebungsdesign ein Mixed-Mode-Ansatz aus schriftlich-postalischer und Online-Erhebung gewählt. Die folgende Tabelle gibt eine Übersicht über das Erhebungsdesign.¹

Tabelle 1 Überblick über das Studiendesign

Studientitel	Sicherheit und Kriminalität in Deutschland 2020 (SKiD)
Grundgesamtheit	Bevölkerung in Privathaushalten (Hauptwohnsitz) ab 16 Jahren in Deutschland
Stichproben	n= 123.665, davon: Basiserhebung: n=122.667 Methodenexperiment (Experimentalgruppe): n=998
Auswahlbasis	Register von Einwohnermeldeämtern
Erhebungsmethode	Mixed Mode: schriftlich-postalisch (PAPI) und online (CAWI)
Erhebungsinstrumente	– 24-seitiger Papierfragebogen, sog. „Kernfragebogen“ (PAPI), zzgl. 4 bundeslandspezifische Seiten in NRW und Hamburg – Online-Fragebogen (CAWI), in NRW und Hamburg inklusive bundeslandspezifischen Fragen
Erhebungssprachen PAPI	Deutsch, Türkisch
Erhebungssprachen CAWI	Deutsch, Türkisch, Russisch, Arabisch

¹ Das Design weicht in einigen Parametern von den ursprünglich ausgeschriebenen Größenordnungen ab. Eine Anpassung des Leistungskatalogs wurde einerseits notwendig, da sich aufgrund einer Rechtsänderung die Portokosten erheblich erhöhten. Zur Finanzierung der höheren Kosten für den Massenversand wurde die Bruttostichprobe um 6 Prozent reduziert, auf eine zusätzliche Onomastikstichprobe wurde verzichtet. Andererseits wurden Leistungen erweitert. So wurde die Seitenanzahl des Kernfragebogens von ursprünglich 20 auf 24 Seiten erhöht.

Weitere Erhebungsmaterialien	<ul style="list-style-type: none"> - Ankündigungsschreiben Deutsch (16 Varianten) - Ankündigungsschreiben Türkisch (16 Varianten) - Ankündigungsschreiben Methodenexperiment Deutsch (16 Varianten) - Begleitschreiben zum Fragebogen Deutsch (16 Varianten) - Begleitschreiben zum Fragebogen Türkisch (16 Varianten) - Anschreiben (inklusive Beiblatt mit Online-Zugangsdaten) Methodenexperiment (16 Varianten) - Datenschutzerklärung Deutsch, Türkisch, Russisch, Arabisch - 1. Erinnerungsschreiben Deutsch (16 Varianten) - 1. Erinnerungsschreiben Türkisch (16 Varianten) - 1. Erinnerungsschreiben Methodenexperiment Deutsch (16 Varianten) - 2. Erinnerungsschreiben Deutsch (16 Varianten) - 2. Erinnerungsschreiben Türkisch (16 Varianten) - 2. Erinnerungsschreiben Methodenexperiment Deutsch (16 Varianten)
Pretests	<ul style="list-style-type: none"> - Feldpretest (schriftlich-postalisch, online): Testung des deutschsprachigen Kernfragebogens (n=1.000 Bruttoadressen) - Kognitiver Pretest: Durchführung Leitfadenterviews in Deutsch (n=20) - Kognitiver Pretest der fremdsprachigen Fragebögen auf Türkisch, Russisch, Arabisch (n=14)
Feldzeit Hauptfeld	23.10.2020 bis 29.01.2021
Fragebogenrücklauf	<p>Rücklauf insgesamt (Basisstichprobe u. Methodenexperiment): 47.565 Fragebögen, davon:</p> <ul style="list-style-type: none"> - Papierfragebögen: 39.459 (83,0 Prozent) - Online-Fragebögen: 8.106 (17,0 Prozent) <p>Aufschlüsselung Fallzahlen: 47.565 Fragebögen, davon</p> <ul style="list-style-type: none"> - 295 doppelte PAPI-Bögen - 67 Fragebögen mit nicht lesbarem QR-Code - 126 Doppelrealisierungen (PAPI und CAWI)
Auswertbare Fälle	45.666 auswertbare Fragebögen insgesamt; davon entfallen 45.351 Fragebögen auf die Basisstichprobe und 315 auf das Methodenexperiment.
Auswertbare Fälle nach Erhebungsmethode	45.351 auswertbare Fragebögen Basisstichprobe, davon: <ul style="list-style-type: none"> - Papierfragebogen: 37.860 (83,5 Prozent) - Online-Fragebogen: 7.491 (16,5 Prozent)
Interviewdauer CAWI	Durchschnittlich 27,2 Minuten
Rückversicherungsmöglichkeit	<ul style="list-style-type: none"> - Telefonische Ansprechpartner beim BKA (Hotline) - Studienspezifische E-Mail-Adresse des BKA - Studienspezifische Webseite des BKA - Studienspezifische Webseite von infas - Informationen auf infas-Homepage - Hotline infas
Datensätze und Dokumentation	<ul style="list-style-type: none"> - Datensatz mit Informationen zu den Gemeinden, aus deren Melderegistern die Adressen gezogen wurden (Gemeindedatensatz) - Befragungsdaten inklusive gesondert gekennzeichnete bereinigter Variablen (Datensatz ohne Angabe der Postleitzahlen) - Datensatz mit vergrößerten Merkmalen und inklusive Angaben zur Postleitzahl - Ausgewählte Bildarchive der Papierfragebögen - Methodendatensatz (Bruttodatensatz) - GewichtungsvARIABLEN - Schriftgut, das dem Fragebogen beigelegt war - Dokumentationen des Papier- und Online-Fragebogens

2 Stichprobenziehung

2.1 Grundgesamtheit

Die Grundgesamtheit der Erhebung umfasst alle in Privathaushalten lebenden Personen in Deutschland, die vor dem 01.07.2004 geboren wurden.

2.2 Aufgabenteilung bei der Stichprobenziehung

Die Stichprobenziehung erfolgte in Einwohnermeldeämtern ausgewählter Gemeinden. Bei einer Einwohnermeldeamtsstichprobe handelt es sich zunächst im Grundsatz um eine zweistufige geklumpnte Stichprobe. In der ersten Stufe (Primary Sample Units = PSU) wird eine Stichprobe von Gemeinden bzw. Sample Points ausgewählt, wobei eine Gemeinde mehrere Sample Points umfassen kann. Die Auswahl der zu befragenden Zielpersonen (Secondary Sample Units = SSU) erfolgt in einem zweiten Schritt direkt aus den Registern der Einwohnermeldeämter.

Die Zweistufigkeit der Stichprobenziehung und mithin die Klumpung ist dem Umstand geschuldet, dass kein zentrales Einwohnermelderegister für die Bundesrepublik existiert, aus dem eine uneingeschränkte Zufallsauswahl (srs = simple random sampling) möglich wäre. Allerdings war es zumindest in drei Bundesländern (Berlin, Schleswig-Holstein und Thüringen) möglich, einen Auszug aus einem zentralen bundeslandweiten Einwohnermelderegister zu erhalten, so dass in diesen Bundesländern eine uneingeschränkte Zufallsauswahl (im Fall von Schleswig-Holstein nach Kreisen geschichtet) auf Bundeslandebene erfolgen konnte. In den übrigen Bundesländern war eine zweistufige Ziehung erforderlich. Dafür wurden zunächst bundesweit Gemeinden ausgewählt (siehe Abschnitt 2.3) und dann auf Gemeindeebene eine uneingeschränkte Zufallsauswahl von Personen vorgenommen. Außerdem haben in Niedersachsen und Bremen die LKÄ die Kontaktierung der Einwohnermeldeämter für die Pointstichprobe übernommen, so dass sich die Stichprobenziehung in den Bundesländern insgesamt wie folgt darstellte (Tabelle 2).

Tabelle 2 Stichprobendesign und -beschaffung nach Bundesländern

Bundesland	Stichprobenansatz	Beschaffung der Stichprobe durch
Schleswig-Holstein	Nach Kreisen geschichtete Zufallsauswahl	LKA
Hamburg	Zweistufige Zufallsauswahl	infas
Niedersachsen	Zweistufige Zufallsauswahl	LKA
Bremen	Zweistufige Zufallsauswahl	LKA
Nordrhein-Westfalen	Zweistufige Zufallsauswahl	infas
Hessen	Zweistufige Zufallsauswahl	infas
Rheinland-Pfalz	Zweistufige Zufallsauswahl	infas
Baden-Württemberg	Zweistufige Zufallsauswahl	infas
Bayern	Zweistufige Zufallsauswahl	infas
Saarland	Zweistufige Zufallsauswahl	infas
Berlin	Uneingeschränkte Zufallsauswahl	LKA
Brandenburg	Zweistufige Zufallsauswahl	infas
Mecklenburg-Vorpommern	Zweistufige Zufallsauswahl	infas
Sachsen	Zweistufige Zufallsauswahl	infas
Sachsen-Anhalt	Zweistufige Zufallsauswahl	infas
Thüringen	Uneingeschränkte Zufallsauswahl	Fachhochschule für öffentliche Verwaltung (FHöV)

2.3 Auswahlstufe 1: Ziehung der Gemeinden/Pointstichprobe

Für die Auswahl der Sample Points wurden die Gemeinden in Deutschland nach Kombination von Kreisen und BIK-Größenklasse (Stand Gemeindeverzeichnis 2018) geschichtet. Innerhalb der Schichtungszellen erfolgte die Auswahl entsprechend dem Bedeutungsgewicht der Gemeinden, das sich aus der Zahl der Bewohner in der Zielgruppe je Gemeinde ergibt. Die Leistungsbeschreibung sah die Ziehung von 450 Sample Points vor. Da Niedersachsen entgegen der Leistungsbeschreibung nicht über ein zentrales Melderegister ziehen konnte, wurde in Absprache mit dem BKA die Anzahl der Sample Points auf insgesamt 505 erhöht.²

Die Anzahl der innerhalb jeder expliziten Schicht per Stratifizierungszelle auszuwählenden Sample Points wird jeweils so auf das Allokationstableau aus Kreis x BIK-Größenklasse verteilt, dass ein bevölkerungsrepräsentatives Abbild entsteht, das heißt proportional zur Zahl der Einwohner in der Zielgruppe. Das Allokationstableau enthält für jede existierende Kombination aus Kreis und BIK-Größenklasse den jeweiligen Bevölkerungsanteil. Die proportionale Verteilung der Sample Points anhand dieser Anteile resultiert in zu erwartenden Werten für die Zellenanteile, die normalerweise Nachkommastellen aufweisen. Nachkommastellen in den Zellenanteilen werden mithilfe spezieller Allokationsverfahren

² Die Pointanzahl bezieht sich auf Bundesländer ohne Zentralregister.

zufällig gerundet (vgl. Cox 1987: 520-524), so dass sich ganzzahlige Werte in den Zellen aus der Kombination von Kreisen und Gemeindegrößenklassen für die Auswahl ergeben.

In vier Bundesländern (Schleswig-Holstein, Hamburg, Nordrhein-Westfalen, Thüringen) wurde die Stichprobe aufgestockt, so dass diese Bundesländer in der finalen Bruttostichprobe überproportional vertreten waren. Da keine zusätzliche Gemeindestichprobe für die Aufstockungen vorgesehen und angesichts der sehr hohen Zahl an Sample Points auch nicht notwendig war, erfolgte die Ziehung der Aufstockungstichproben in den ausgewählten Gemeinden der Basisstichprobe. Dies bedeutet, dass in diesen Bundesländern eine höhere Anzahl an Personenadressen gezogen werden musste. In den Bundesländern mit zentralem Melderegister wurde entsprechend eine größere einstufige Stichprobe bestellt.

2.4 Auswahlstufe 2: Ziehung der Einsatzstichprobe aus den Registern der Einwohnermeldeämter

Nach der Auswahl der Gemeinden erfolgte auf der zweiten Stufe die Ziehung der Personenadressen (SSU) aus den Registern der Meldeämter. Es handelt sich um eine systematische Zufallsauswahl (Intervallziehung). Ausgehend von einer zufälligen Startadresse werden über eine feste Schrittweite die übrigen Adressen systematisch ausgewählt. Dabei wurde für jeden Point die gleiche Anzahl an Personenadressen ($n = 55$) angefordert. In den Bundesländern mit Aufstockung wurde die Anzahl entsprechend erhöht.

Zum Zweck der Adressziehung erhielt jede ausgewählte Gemeinde die folgenden Unterlagen:

- Ein Anschreiben von infas inklusive eines Begleitschreibens des BKA-Präsidenten mit Informationen über die Studie, den Auftraggeber, eine Terminvorgabe sowie einer Erklärung zum Datenschutz.
- Die Ziehungsanweisung für die Einwohnermeldeämter bzw. für die zuständigen Rechenzentren mit Angabe der Auswahlkriterien sowie Vorgaben zu Ziehungsintervall und Startzahl.
- Ein Statistikformblatt für Informationen zur Grundgesamtheit in der Gemeinde sowie zur Kontrolle des Auswahlalgorithmus.

In der Ziehungsanweisung wurde der Ziehungsalgorithmus detailliert erläutert. In einem ersten Schritt sollte die Auswahl der Personen aus der Zielgruppe erfolgen, im zweiten Schritt sollte die Startadresse ermittelt und im letzten Schritt die Schrittweite bestimmt werden und anschließend die Auswahl der Personen/Adressen erfolgen.

2.5 Umgang mit nicht teilnahmebereiten Gemeinden

Nicht alle Gemeinden zeigten sich kooperativ und lieferten die gewünschten Stichproben. Für Gemeinden, die die Adresslieferung verweigerten, wurden in Abstimmung mit dem BKA strukturgleiche Ersatzgemeinden ausgewählt.

Diese wurden ebenfalls von infas kontaktiert und um eine Stichprobe gebeten. Es wurden insgesamt 24 Gemeinden zusätzlich kontaktiert, also 24 mehr als die ursprüngliche Gemeindestichprobe umfasste.

In Summe wurden 396 Gemeinden³ kontaktiert und um eine Stichprobenziehung aus dem Melderegister ihrer Gemeinde gebeten. Diese verteilen sich wie folgt auf die Bundesländer.

Tabelle 3 Kontaktierte Gemeinden nach Bundesland

Bundesland	Gemeindeanzahl laut Stichprobenplan	Kontaktierte Ersatzgemeinde	Kontaktierte Gemeinden gesamt
Hamburg	1	0	1
Nordrhein-Westfalen	94	2	96
Hessen	37	5	42
Rheinland-Pfalz	27	0	27
Baden-Württemberg	67	4	71
Bayern	76	7	83
Saarland	7	0	7
Brandenburg	18	0	18
Mecklenburg-Vorpommern	11	1	12
Sachsen	21	3	24
Sachsen-Anhalt	13	2	15
Gesamt	372	24	396

Die intensiven Bemühungen um die Stichprobenlieferung haben die Ausfälle durch nicht teilnahmebereite Gemeinden erfolgreich beschränken können. Der Austausch von Gemeinden stieß allerdings an Grenzen. Im Ergebnis wurden für 25 Gemeinden keine Stichprobeninformationen geliefert. Die Gründe dafür waren:

- Gemeinden, für die keine strukturgleiche Ersatzgemeinde⁴ gezogen wurde oder bei denen auch die Ersatzgemeinde eine Stichprobenziehung verweigerte.

³ In diesem Abschnitt werden lediglich die Bemühungen um eine Stichprobenlieferung dokumentiert, für die infas verantwortlich war, s. Verteilung der Zuständigkeiten bei der Stichprobenbeschaffung in Tabelle 2. Bei der Kontaktilierung der Gemeinden wurde infas vom BKA unterstützt.

⁴ In einem Fall konnte die ausgefallene Gemeinde nur strukturähnlich ersetzt werden, weil nur eine Option mit gleicher BIK-Größenklasse im Bundesland noch nicht in der ursprünglichen Gemeindestichprobe enthalten war. Strukturähnlich heißt in diesem Fall, dass die BIK-Größenklasse konstant gehalten wurde, jedoch die Ersatzgemeinde nicht im selben Kreis liegt.

- Zusagen der Gemeinden lagen vor, waren aber bis zum Ablauf der Lieferfrist noch in Bearbeitung bei den Gemeinden.
- Bis zuletzt keine Rückmeldung aus den Gemeinden.
- Wunsch nach Neuversand der Unterlagen durch Gemeinden, aber ohne anschließende Lieferung der Stichprobendaten.
- Explizite Verweigerung der Stichprobenlieferung, zeitlich keine Möglichkeit für Ersatz durch eine alternative Gemeinde.

Die Bemühungen wurden im September 2020 beendet, damit die Einsatzstichprobe für die onomastische Qualifizierung und den anschließenden Feldstart aufbereitet werden konnte. Statt in – laut Stichprobenplan ursprünglich vorgesehenen – 372 Gemeinden konnte letztlich die Erhebung in 347 Gemeinden durchgeführt werden. Die Reduktion um 25 Gemeinden verteilt sich auf unterschiedliche Bundesländer (Tabelle 4). Die prozentuale Verteilung der Gemeinden auf die Bundesländer verändert sich dadurch allerdings nur geringfügig. Da der Umfang der Sample Points von Beginn an ausreichend groß gewählt war, ließ sich das vereinbarte Gesamtbrutto trotz ausgefallener Gemeinden über eine leicht erhöhte Zahl von Adressen pro Sample Point realisieren.

Tabelle 4 Gemeindeverteilung laut Stichprobenplan und in Einsatzstichprobe

Bundesland	Umfang Gemeinden nach Stichprobenplan		Einsatzstichprobe inklusive Ersatzgemeinden		Differenz in Prozentpunkten
	abs.	%	abs.	%	
Hamburg	1	0,3	1	0,3	0,0
Nordrhein-Westfalen	94	25,3	89	25,6	0,3
Hessen	37	9,9	32	9,2	0,7
Rheinland-Pfalz	27	7,3	27	7,8	0,5
Baden-Württemberg	67	18,0	62	17,9	0,1
Bayern	76	20,4	71	20,5	0,1
Saarland	7	1,9	6	1,7	0,2
Brandenburg	18	4,8	18	5,2	0,4
Mecklenburg-Vorpommern	11	3,0	11	3,2	0,2
Sachsen	21	5,6	17	4,9	0,7
Sachsen-Anhalt	13	3,5	13	3,7	0,2
Gesamt	372	100,0	347	100,0	

2.6 Zusammenführung der Bundeslandstichproben

Die Gesamtstichprobe ergibt sich aus den Lieferungen aus Bundesländern mit zentralem Melderegister und den Lieferungen aus den Einwohnermeldeämtern der Gemeinden, die in 11 Fällen von infas beschafft wurden und in 5 Fällen von den LKÄ. Insgesamt wurden Adressen aus 2.010 Gemeinden eingesetzt.

Die Lieferungen aus zentralen Registern wurden proportional zur Bevölkerung des Bundeslands bzw. entsprechend der vereinbarten Aufstockung eingesetzt. Für die Einsatzstichprobe waren jeweils 53 Adressen pro Point in der Haupterhebung angestrebt, bzw. entsprechend mehr in den aufgestockten Bundesländern (Hamburg: 1.917, Nordrhein-Westfalen: 324). Einzelne Gemeinden hatten etwas weniger als die angeforderten Adressen geliefert. Um die Varianz von Auswahlwahrscheinlichkeiten (und damit der Designgewichte) möglichst klein zu halten, wurde darauf verzichtet, die Unterbesetzung von Gemeinden durch Überbesetzung anderer Gemeinden auszugleichen.

Tabelle 5 Umfang Bruttostichprobe

Spalten%	Bruttostichprobe	
	abs.	%
Schleswig-Holstein	23.535	19,0
Hamburg	23.026	18,6
Niedersachsen	3.011	2,4
Bremen	273	0,2
Nordrhein-Westfalen	37.770 ⁵	30,5
Hessen	2.026	1,6
Rheinland-Pfalz	1.480	1,2
Baden-Württemberg	3.842	3,1
Bayern	4.552	3,7
Saarland	330	0,3
Berlin	1.275	1,0
Brandenburg	979	0,8
Mecklenburg-Vorpommer	603	0,5
Sachsen	1.320	1,1
Sachsen-Anhalt	817	0,7
Thüringen	18.826	15,2
Gesamt	123.665	100,0

⁵ Eine Gemeinde in Nordrhein-Westfalen hatte Adressen mit falsch zugeordneten Vornamen geliefert. Die zweite (fehlerfreie) Stichprobe dieser Gemeinde wurde vor ihrem Feldeinsatz mit der ersten Lieferung abgeglichen. Die zweite Stichprobe wurde um Fälle bereinigt, die bereits in der ersten, fehlerhaften Stichprobe enthalten waren. Im Ergebnis reduzierte sich die korrigierte Einsatzstichprobe um 5 Fälle von 37.775 (ursprünglich geplant) auf 37.770 (tatsächlich eingesetzt). 2 dieser 5 Fälle entfallen auf die Experimentalgruppe des Methodenexperiments, die restlichen 3 auf die Basisstichprobe.

2.7 Onomastische Klassifizierung der Bruttostichprobe

Für Zielpersonen, die nicht ausreichend Deutsch lesen und verstehen können, aber die türkische Sprache beherrschen, sollten neben den deutschsprachigen Erhebungsunterlagen sämtliche Unterlagen in Türkisch beigelegt werden. Allerdings lagen keine Informationen über die Sprachfähigkeiten der Zielpersonen vor. Da der Migrationshintergrund kein Merkmal ist, das bei den Einwohnermeldeämtern bekannt ist, wurde die Bruttostichprobe über ein onomastisches Verfahren⁶ im Anschluss an die Ziehung qualifiziert. Mit diesem Verfahren sollten möglichst viele Personen, die nicht ausreichend Deutsch, aber gut Türkisch verstehen, erreicht werden. Dabei wurde anhand der Vor- und Nachnamen ein vermuteter türkischer Migrationshintergrund ermittelt. Von 123.665 Personennamen wurde für 5.024 (4,1 Prozent) ein mutmaßlich türkischer Hintergrund identifiziert.

Tabelle 6 Ergebnis der onomastischen Qualifizierung der Bruttostichprobe

Spalten%	Bruttostichprobe insgesamt		Personen mit vermutlich türkischem Migrationshintergrund	
	abs.	%	abs.	%
Schleswig-Holstein	23.535	19,0	561	11,2
Hamburg	23.026	18,6	1.365	27,2
Niedersachsen	3.011	2,4	74	1,5
Bremen	273	0,2	16	0,3
Nordrhein-Westfalen	37.770	30,5	2.407	47,9
Hessen	2.026	1,6	92	1,8
Rheinland-Pfalz	1.480	1,2	34	0,7
Baden-Württemberg	3.842	3,1	196	3,9
Bayern	4.552	3,7	142	2,8
Saarland	330	0,3	4	0,1
Berlin	1.275	1,0	75	1,5
Brandenburg	979	0,8	8	0,2
Mecklenburg-Vorpommern	603	0,5	3	0,1
Sachsen	1.320	1,1	2	0,0
Sachsen-Anhalt	817	0,7	3	0,1
Thüringen	18.826	15,2	42	0,8
Gesamt	123.665	100,0	5.024	100,0

Quelle: infas, eigene Berechnung

⁶ Onomastik = Namenforschung, Namenkunde

2.8 Stichprobenumfang Methodenexperiment

Parallel zur Haupterhebung wurde ein Methodenexperiment durchgeführt. Für das Experiment wurde eine zusätzliche, bundeslandproportionale Stichprobe gezogen und im Hauptfeld eingesetzt.⁷ Die Stichprobe der Experimentalgruppe des Methodenexperiments umfasst 998 zufällig ausgewählte Personen⁸. Sie erhielten abweichend von der Kontrollgruppe den Papierfragebogen erst mit dem zweiten Erinnerungsschreiben (s. Abschnitt 4.6.2).

In der Einsatzstichprobe der Haupterhebung wurde eine Kontrollgruppe im gleichen Umfang, das heißt von ebenfalls 998 Fällen auf Gemeindeebene strukturgleich zur Experimentalgruppe zufällig bestimmt und mit einem Stichprobenkennner versehen. Die Kontrollgruppe blieb stets Teil der regulären Einsatzstichprobe (Basisstichprobe) und wurde wie die restlichen Fälle der Basisstichprobe behandelt.

Tabelle 7 Verteilung der Stichproben für Methodenexperiment und Kontrollgruppe nach Bundesländern

Spalten%	Methodenexperiment (Experimentalgruppe)		Kontrollgruppe (Teil der Basisstichprobe)	
	abs.	%	abs.	%
Schleswig-Holstein	35	3,5	35	3,5
Hamburg	22	2,2	22	2,2
Niedersachsen	96	9,6	96	9,6
Bremen	8	0,8	8	0,8
Nordrhein-Westfalen	213	21,3	213	21,3
Hessen	75	7,5	75	7,5
Rheinland-Pfalz	49	4,9	49	4,9
Baden-Württemberg	134	13,4	134	13,4
Bayern	158	15,8	158	15,8
Saarland	12	1,2	12	1,2
Berlin	44	4,4	44	4,4
Brandenburg	30	3,0	30	3,0
Mecklenburg-Vorpommern	20	2,0	20	2,0
Sachsen	49	4,9	49	4,9
Sachsen-Anhalt	27	2,7	27	2,7
Thüringen	26	2,6	26	2,6
Gesamt	998	100,0	998	100,0

Quelle: infas, eigene Berechnung

⁷ Länderaufstockungen wurden dabei nicht berücksichtigt, d.h. die Stichprobe des Methodenexperiments ist im Gegensatz zur Haupterhebung auf Bundeslandebene bevölkerungsproportional.

⁸ Ursprünglich geplant war ein Stichprobenumfang von 1.000 Zielpersonen. Nach dem Abgleich einer fehlerhaften Adresslieferung aus einem Einwohnermeldeamt mit der nachgelieferten korrigierten Stichprobe mussten zwei Fälle für den Feldeinsatz ausgeschlossen werden.

3 Erhebungsunterlagen

3.1 Der Papierfragebogen als Leitinstrument

Der Fragebogen wurde vom Bundeskriminalamt (BKA) und weiteren Landeskriminalämtern (LKÄ) entwickelt. infas beteiligte sich aktiv an der Endredaktion des Fragebogens und überführte zudem den Fragenkatalog in ein professionelles Layout. Die farbliche Gestaltung des Fragebogens verdeutlicht, dass es sich um eine bundesweite Befragung handelt. Die Farben Schwarz, Rot und Gold aus dem SKiD-Logo finden sich mehrfach im Fragebogen wieder, sei es zum Beispiel im Farbschema der Texte, bei der farblich hinterlegten Fragennummer oder in den Säulendarstellungen neben den Seitenzahlen am unteren Rand jeder Fragebogenseite. Das gewählte Layout verleiht dem Fragebogen einen individuellen und unverwechselbaren Charakter. Der mehrfarbige Druck erhöht die Attraktivität des Fragebogens und motiviert zusätzlich zur Teilnahme. Zudem erschwert der Mehrfarbendruck die Möglichkeit, den Fragebogen zu fälschen, um beispielsweise unauffällig mehrfach teilzunehmen.

Der finale Kernfragebogen⁹ umfasste 24 Seiten. Der Aufbau des Fragebogens gliederte sich in eine Titelseite, auf der zentrale Hinweise zur Erhebung sowie die Zugangsdaten für den Online-Fragebogen notiert waren. Ebenfalls auf der Titelseite waren das Bundesministerium des Innern als Auftraggeber der Studie, infas als durchführendes Erhebungsinstitut sowie der Hinweis auf die Förderung durch den Fonds für die Innere Sicherheit der Europäischen Union genannt. Für Rückfragen wurden zudem Kontaktdaten (Telefonnummer und E-Mail-Adresse) des BKA angegeben.

Auf der ersten Innenseite (Rückseite der Titelseite) waren auf einer halben Seite Ausfüllhinweise eingetragen. Die weiteren 22,5 Seiten waren mit inhaltlichen Fragen gefüllt. Im Zentrum des Befragungsinteresses standen die Themen Bewertung der Polizeiarbeit, Opfererfahrungen, Kriminalitätsfurcht, Anzeigeverhalten sowie die Nutzung von Hilfeangeboten. Im letzten Abschnitt wurden Fragen zur Soziodemografie gestellt. Der Kernfragebogen gliedert sich wie folgt:

- Fragen zur Wohngegend,
- Fragen zur Bewertung der Polizei,
- Fragen zum Sicherheitsgefühl,
- Fragen zur Erfahrung mit Kriminalität (seit 2015 und in den letzten 12 Monaten) und Nachfragen zu berichteten Erfahrungen,
- Fragen zur Soziodemografie.

⁹ Mit „Kernfragebogen“ werden die ersten 24 Seiten bezeichnet, die bundesweit einheitlich eingesetzt wurden. In den Bundesländern Nordrhein-Westfalen und Hamburg kamen zusätzlich jeweils 4 landesspezifische Seiten hinzu. Sie werden im Folgenden als „Ländermodule“ bezeichnet.

Die Abfrage nach Art und Anzahl der Erfahrungen mit Kriminalität nahmen Raum über mehrere Seiten ein. Die Beantwortung dieser langen Fragebatterien erforderte von den Zielpersonen ein maßgebliches Verständnis der zum Teil komplexen Frageformulierungen. Zudem erforderten die Nachfragen, die sich auf zuvor gemachte Angaben bezogen, einen hohen Grad an Konzentration der Zielpersonen. Durch eine unterstützende grafische Gestaltung und ergänzende Ausfüllhinweise sollte die Fehleranfälligkeit beim Ausfüllen des sehr anspruchsvollen Fragekomplexes reduziert werden.

In den Bundesländern Nordrhein-Westfalen (NRW) und Hamburg wurden neben dem Fragenprogramm des Kernfragebogens zusätzlich landesspezifische Fragen im Umfang von 4 weiteren Fragebogenseiten in den Fragenkatalog aufgenommen. Die zusätzlichen Fragen wurden am Ende des Kernfragebogens eingefügt. Lediglich die abschließende offene Frage nach weiteren Themen, die die Zielpersonen hier notieren konnten, wurde auch in den beiden genannten Bundesländern als abschließende Frage im Bogen beibehalten. Der Umfang der Fragebögen in NRW und Hamburg umfasst demzufolge 28 Seiten.

Der Papierfragebogen wurde nach der Fertigstellung der deutschen Fassung ins Türkische übersetzt. Neben dem Kernfragebogen wurden auch die zusätzlichen Fragen der Ländermodule Hamburg und NRW ins Türkische übersetzt.

3.2 Der Online-Fragebogen (CAWI)

Als Alternative zum Papierfragebogen bestand die Möglichkeit, den Fragebogen online auszufüllen (CAWI). Über die verschiedenen Teilnahmemöglichkeiten wurden die Zielpersonen im Begleitschreiben zum Fragebogen aufgeklärt.

Auf der Titelseite des Papierfragebogens waren der Weblink zur Online-Befragung, ein individueller Zugangscode sowie ein individueller QR-Code mit Direktlink zum Online-Fragebogen aufgedruckt. Der Zugang zum Fragebogen war somit entweder über den Aufruf der Internetadresse und Eingabe des Zugangscodes oder direkt durch Einscannen des QR-Codes mit einem geeigneten Endgerät (z.B. Smartphone, Tablet) möglich. Über die aufgedruckten individuellen Zugangscodes war die Zuordnung des Fragebogens zur jeweiligen Zielperson für die Eingangserfassung und Steuerung der Erinnerungsaktionen gewährleistet. Anhand des Zugangscodes wurde die für die jeweilige Zielperson vorgesehene Variante (ausschließlich Kernfragebogen, Kernfragebogen mit Hamburger Modul oder Kernfragebogen mit NRW-Modul) des Online-Fragebogens angesteuert.

Technische Rahmenbedingungen

Der Login in den Fragebogen erfolgte entweder durch Aufruf der Internetseite und Eingabe des Zugangscodes oder per Direktaufruf mittels QR-Code. Mit dem persönlichen Zugangscode bzw. dem Direktlink per QR-Code konnte der Fragebogen nur einmal vollständig ausgefüllt werden. Sobald die Zielpersonen den Fragebogen vollständig ausgefüllt und dies auf der letzten Bildschirmseite bestätigt hatten, war kein erneuter Zugriff auf den Fragebogen möglich. Eine Unterbrechung der Befragung und ein erneuter Login zu einem späteren Zeitpunkt waren technisch möglich. Die Befragung konnte in diesem Fall an der Stelle fort-

gesetzt werden, an der sie unterbrochen worden war. Ein automatisches Timeout erfolgte nach 5 Minuten Verweilzeit auf einer Befragungsseite, in diesem Fall bekamen die Zielpersonen einen Hinweis zur automatischen Abmeldung und konnten die Befragung nach erneuter Anmeldung an derselben Stelle fortsetzen. Ebenso ermöglicht wurde die Veränderung von Antworten durch Zurückklicken innerhalb der Befragung.

Im Online-Fragebogen war bei jeder Frage eine Prüfung auf Vollständigkeit der Eingaben vorgesehen. Bei fehlenden oder unvollständigen Angaben wurde eine Erinnerung an das (vollständige) Beantworten der Frage eingeblendet sowie eine zusätzliche Antwortkategorie „kann/möchte ich nicht beantworten“ angeboten. So konnte einerseits sichergestellt werden, dass die Zielpersonen nicht unbeabsichtigt eine Antwort auslassen und andererseits bestand analog zum Papierfragebogen jederzeit die Möglichkeit, eine Frage bewusst unbeantwortet zu lassen.¹⁰

Layout in responsivem Design

Das Layout eines Online-Fragebogens ist nicht nur eine Geschmacksfrage, sondern hat unmittelbare Auswirkungen auf die Qualität der erhobenen Daten. Dies wird besonders deutlich bei der weit verbreiteten Verwendung von mobilen Endgeräten. Der Anteil von Zielpersonen, der Online-Befragungen auch über Smartphones oder Tablets aufruft, steigt beständig. Der Online-Fragebogen wurde daher spezifisch auch für die Nutzung auf mobilen Endgeräten optimiert, so dass sich die einzelnen Fragebogenelemente bis zu einem gewissen Grad an die jeweilige Bildschirmgröße bzw. Bildschirmauflösung des genutzten Endgeräts anpassen. Dadurch wird vermieden, dass die Nutzer vertikal und horizontal scrollen müssen und dabei gegebenenfalls nicht alle notwendigen Fragetexte und -elemente sehen. Dies verbessert nicht nur die Nutzbarkeit für die Zielpersonen, sondern es wird auch eine Fehlerquelle bei der Beantwortung des Fragebogens vermieden. Zu Beginn des Fragebogens wurde den Zielpersonen ein Hinweis eingeblendet, bei der Nutzung von mobilen Endgeräten (Smartphone, Tablet) diese Geräte für eine übersichtlichere Darstellung im Querformat zu bedienen.

Um das Fragebogenlayout optimal an die Gerätegröße anzupassen, wurde ein responsives Fragebogendesign verwendet. Dies stellt spezifische Anforderungen an den Fragebogen, die in enger Abstimmung mit dem BKA umgesetzt wurden. So wurden Itembatterien bzw. Matrixfragen aus dem Papierfragebogen in Einzelitems aufgelöst und jeweils ein Item pro Bildschirmseite eingeblendet. Um die Nutzung des Fragebogens zu erleichtern, erfolgt bei (geschlossenen) Einfachnennungen nach der Beantwortung mit zeitlicher Verzögerung eine automatische Weiterleitung zur nächsten Frage. Dies vereinfacht insbesondere bei der Nutzung von Endgeräten mit Touch-Display die Bedienung, da eine zusätzliche Betätigung des „Weiter“-Buttons entfällt. Bei Mehrfachnennungen, Fragen mit

¹⁰ Eine Dokumentation des Online-Fragebogens inklusive der Ausfüllhinweise und automatisierten Prüfungen von Eingaben liegt dem BKA gesondert vor.

offenen Antwortkategorien sowie einer ausgewählten Einfachnennung¹¹ war eine aktive Bestätigung der (vollständigen) Eingabe durch die Nutzer mittels des "Weiter"-Buttons unterhalb der Frage notwendig.

Qualitätssicherung

Die Umsetzung der Papierfragebögen in die Online-Fassung erfolgte in enger Abstimmung mit dem BKA und folgte dem Ziel, den CAWI-Fragebogen möglichst nahe am Layout des Papierfragebogens umzusetzen und die technischen Möglichkeiten für eine gute Datenqualität zu nutzen. Die programmierten Fragebögen wurden vor Feldeinsatz in einem mehrstufigen Verfahren geprüft. Zunächst wurden die Fragebögen bei infas unter Leitung des Fragebogenprogrammierers durch mehrere Fragebagentester intensiv getestet. Um eine fehlerfreie Datenerhebung zu gewährleisten, wurden im Anschluss in der Abteilung Datenaufbereitung die Fragebogensteuerung, also der Filterverlauf im Fragebogen, sowie die korrekte Datenablage im Kernfragebogen und in den beiden Zusatzmodulen geprüft. Die Datenablageprüfung umfasste die Einhaltung der vorgesehenen Wertebereiche, also die laut Codebook vorgesehenen Codes bei jeder Frage bzw. die gültigen Werte bei offenen Zahlenangaben, sowie die korrekte Ablage aller offenen Angaben.

Nach Abschluss der Prüfungen aller beteiligten Abteilungen wurden die programmierten Fragebögen abschließend durch die infas-Projektleitung und das BKA getestet und für den Feldeinsatz freigegeben.

3.3 Weitere Erhebungsunterlagen

Die Zielpersonen erhielten begleitend zum schriftlichen Fragebogen weitere ergänzende Erhebungsunterlagen, die ihnen zu unterschiedlichen Zeitpunkten während des Erhebungszeitraums zugeschickt wurden. Die Unterlagen variierten in Abhängigkeit davon, ob es sich rein um die Standardversion, den zusätzlichen Versand türkischsprachiger Unterlagen oder das Methodenexperiment handelte.

Briefpapier

Für die verschiedenen Anschreiben wurde für jedes Bundesland ein spezifisches Briefpapier entworfen. In Anlehnung an die Titelseite des Fragebogens wurde neben dem SKiD-Logo unter anderem auch – mit Ausnahme von Niedersachsen – das jeweilige Logo der Landespolizei auf dem Briefbogen berücksichtigt. Für Rückfragen zur Studie waren zusätzlich Hotline-Nummer und die E-Mail-Adresse des BKA-Projektteams aufgeführt. Der Druck auf dem Briefpapier war mehrfarbig angelegt.

¹¹ Aufgrund der besonderen Textlänge und der grafischen Darstellung der Antwortskala musste die Eingabe in Frage 46 („Leiterfrage“) aktiv bestätigt werden.

Ankündigungsschreiben

Zum Start der Erhebung wurde zunächst ein einseitiges Ankündigungsschreiben versendet. In diesem Anschreiben wurden die Zielpersonen über die bevorstehende Befragung informiert und die Zusendung des Fragebogens in etwa zwei Wochen angekündigt. Die Anschreiben enthielten landesspezifische Formulierungen und wurden sowohl vom Präsidenten des BKA als auch einem Leiter des jeweiligen Landeskriminalamts oder einer ähnlichen Stelle sowie dem Geschäftsführer von infas unterschrieben.

Begleitschreiben zum Fragebogen

In einem zweiten Schritt erhielten die Zielpersonen den Papierfragebogen, welchem ein Begleitschreiben sowie ein an infas adressierter und frankierter Rückumschlag beigelegt waren. In dem Begleitschreiben wurden die Zielpersonen über das Ziel, zentrale Fragestellungen der Studie und die Stichprobenquelle aufgeklärt. Sie wurden zusätzlich informiert, dass sie zwischen zwei Möglichkeiten der Teilnahme wählen konnten, und zwar zwischen dem Ausfüllen des Papierfragebogens oder des Online-Fragebogens. Für Rückfragen wurden die Hotlinenummer und die E-Mail-Adresse des BKA-Projektteams aufgeführt. Unterschrieben war das Begleitschreiben von denselben drei Personen wie das des Ankündigungsschreibens. Es wurden 16 landesspezifische Begleitschreiben eingesetzt.

Datenschutzerklärung

Zusätzlich zum Fragebogen und Begleitschreiben erhielten die Zielpersonen in der zweiten Versandrunde eine Datenschutzerklärung. Es gab lediglich eine Fassung der Datenschutzerklärung, die in allen Bundesländern eingesetzt wurde. Sie wurde immer dann mit versendet, wenn einem Anschreiben ein Fragebogen beigelegt war.

Erstes Erinnerungsanschreiben

In einer dritten Versandwelle wurde ein erstes Erinnerungsschreiben versendet. Es war als kombiniertes Dank- und Erinnerungsschreiben konzipiert, da von Überschneidungen beim Rückversand ausgefüllter Fragebögen und dem Versand des Erinnerungsschreibens auszugehen war. Im Anschreiben wurde noch einmal für die Teilnahme geworben. Auch dieses Anschreiben wurde wieder von denselben drei Personen unterzeichnet, die auch die vorangegangenen Anschreiben unterschrieben hatten.

Zweites Erinnerungsanschreiben

In der letzten Versandrunde erfolgte erneut der Versand eines kombinierten Dank- und Erinnerungsschreibens. Ausführlicher als im ersten Erinnerungsschreiben wurde noch einmal auf die Zielrichtung und die Wichtigkeit der Studie eingegangen. Es wurde erstmalig mit dem 31.12.2020 auch ein Datum genannt, bis wann der Fragebogen ausgefüllt an infas zurückgeschickt werden sollte. Mit dem zweiten Erinnerungsschreiben erhielten die Zielpersonen erneut den Frage-

bogen, die Datenschutzerklärung sowie einen an infas adressierten und frankierten Rückumschlag.

Erhebungsunterlagen Methodenexperiment

Für die Stichprobe des Methodenexperiments wurde ein etwas anderes Vorgehen gewählt und die oben genannten Unterlagen entsprechend angepasst.

Im Ankündigungsschreiben wurde nur die Online-Befragung im Internet angekündigt. Dem anschließend versendeten Begleitschreiben lag ein DIN-A4-Blatt mit den Zugangsdaten für den Online-Fragebogen bei. Erst mit dem zweiten Erinnerungsschreiben erhielten die Zielpersonen, die sich bisher nicht an der Befragung beteiligt hatten, zusätzlich zu den Online-Zugangsdaten den Papierfragebogen zugeschickt.

Den Druck sowie den Versand des Erhebungsmaterials übernahm infas.

3.4 Fremdsprachige Erhebungsunterlagen

Die Teilnahme an der Befragung war unter bestimmten Voraussetzungen wahlweise auf Deutsch oder in den Fremdsprachen Türkisch, Russisch und Arabisch möglich. Für die Befragung der Personen mit vermutetem türkischem Migrationshintergrund wurden der Papier- und Online-Fragebogen sowie sämtliche Erhebungsunterlagen in enger Abstimmung mit dem BKA und infas ins Türkische übersetzt. Diese Personen hatten jederzeit die Wahlmöglichkeit, den Online- bzw. Papierfragebogen in der deutschen oder türkischen Fassung auszufüllen.

Allen Teilnehmern an der Online-Befragung stand neben der deutschen und der türkischen auch eine russische und arabische Fassung des Fragebogens zur Auswahl. Die Zielpersonen wurden über diese Wahlmöglichkeit im Begleitschreiben zum Fragebogen informiert. Im Online-Fragebogen konnte über ein Drop-down-Menü jederzeit die Spracheinstellung geändert werden. Die im Online-Fragebogen hinterlegte Datenschutzerklärung stand ebenfalls in allen vier Sprachen zur Verfügung. Die in die einzelnen Sprachen übersetzten Dokumente sind in der nachfolgenden Tabelle aufgeführt.

Tabelle 8 Zuordnung Fremdsprache und Erhebungsunterlagen

Sprache	Erhebungsunterlagen
Türkisch	<ul style="list-style-type: none">- Fragebögen PAPI und CAWI- Ankündigungsschreiben in 16 Varianten- Begleitschreiben in 16 Varianten- Datenschutzerklärung- Erstes Erinnerungsschreiben in 16 Varianten- Zweites Erinnerungsschreiben in 16 Varianten
Russisch, Arabisch	<ul style="list-style-type: none">- Fragebogen CAWI- Datenschutzerklärung

3.4.1 Übersetzungen

Die Übersetzungen erfolgten durch muttersprachliche Übersetzer und wurden durch ein Übersetzungsbüro vorgenommen, mit dem infas schon seit vielen Jahren zusammenarbeitet. Die Übersetzung der Erhebungsmaterialien orientiert sich an den Standards des European Social Survey (ESS) (vgl. u.a. Dorer/Krajceva 2018, Zavala-Rojas 2014) und wurde als inhaltliche 1:1-Übersetzung angelegt. Dies meint nicht eine rein wörtliche Übersetzung, sondern dass sie den Kriterien der funktionalen Äquivalenz genügen muss.

Nach der Erstübersetzung wurde die Richtigkeit der Übersetzungen jeweils von einem zweiten muttersprachlichen Übersetzer geprüft. Diese Zweitübersetzer protokollierten dann bei Bedarf Änderungs- und Verbesserungsvorschläge zur Erstübersetzung. Sämtliche Anmerkungen der Zweitübersetzer wurden an die Erstübersetzer zurückgespielt. Diese nahmen zu den Anmerkungen schriftlich Stellung und die beiden Übersetzer verständigten sich auf eine gültige Fassung. Die auf diese Weise geprüften und diskutierten Übersetzungen der fremdsprachigen Texte wurden von infas in die Erhebungsunterlagen überführt und die korrekte Übernahme der Übersetzungen geprüft.

3.4.2 Kognitiver Pretest: Prüfung fremdsprachiger Erhebungsinstrumente

Die im beschriebenen Verfahren geprüften Übersetzungen der Erhebungsinstrumente wurden zusätzlich einem kognitiven Pretest unterzogen. Dabei kamen jeweils die für die Erhebung vorgesehenen Haupteinsatzvarianten zum Einsatz (Türkisch: PAPI; Russisch und Arabisch: CAWI).

Die fremdsprachigen Pretestinterviews wurden im Zeitraum vom 21.09.2020 bis zum 23.09.2020 (Türkisch) und vom 24.09.2020 bis zum 30.09.2020 (Russisch und Arabisch) durchgeführt. Es wurden jeweils 5 Interviews auf Russisch und Türkisch sowie 4 Interviews auf Arabisch geführt.

Die Pretestteilnehmer wurden aus dem persönlichen Netzwerk von infas-Mitarbeitern rekrutiert. Es wurde darauf geachtet, dass die Pretestinterviews mit Personen durchgeführt wurden, die selbst keine Interviewer sind. Aufgrund der Kontaktbeschränkungen während der Corona-Pandemie wurden die Pretestinterviews telefonisch durchgeführt.¹²

Die Pretestteilnehmer erhielten im Vorfeld des Interviewtermins den Papierfragebogen (Türkisch) bzw. Zugangsdaten zum Online-Fragebogen (Russisch/Arabisch) zugeschickt und sollten den Fragebogen möglichst erst 1 bis 2 Stunden vor dem Gesprächstermin bearbeiten. Die Teilnehmer wurden gebeten, während des Ausfüllens Stellen zu markieren (Türkisch) bzw. zu notieren (Russisch/Arabisch), die ihnen unklar oder unverständlich waren. Zu Beginn des Interviews wurden diese Unklarheiten besprochen. Die Interviewer notierten bei

¹² Das Konzept für die telefonische Durchführung der Pretestinterviews wurde mit dem BKA abgestimmt. Das Konzeptpapier liegt dem BKA vor.

Bedarf eine Alternativformulierung, die aus Sicht der Pretest-teilnehmer der deutschen Formulierung besser entsprach.

Den Teilnehmern wurden in einem weiteren Schritt zu ausgewählten Fragen oder Items sowohl die Erstübersetzung als auch die deutsche Formulierung vorgelesen und gefragt, ob in der Übersetzung die Bedeutung der deutschen Formulierung verstanden wurde oder nicht. Im Anschluss wurde eine alternative Übersetzung (Übersetzung des Zweitübersetzers) für die Frage vorgelesen und gefragt, ob diese Zweitübersetzung für die Zielperson das gleiche bedeutet wie die erste Formulierung oder eher etwas anderes. Die Rückmeldungen der Teilnehmer wurden durch die Interviewer für die spätere Auswertung protokolliert und kommentiert.

Insgesamt gab es nur wenige Anmerkungen der Pretestteilnehmer zu den Übersetzungen. Nach Auswertung der Rückmeldungen aus dem Pretest wurden diese anschließend von der infas-Projektleitung gemeinsam mit zweisprachigen infas-Mitarbeitern abschließend geprüft und bei Bedarf Anpassungen an den Textstellen vorgenommen. Der Umfang der notwendigen Anpassungen war recht gering.¹³

¹³ Insgesamt wurden 13 Textstellen korrigiert (türkisch 3x, arabisch 4x, russisch 6x).

4 Durchführung der Hauptstudie

4.1 Feldzeit

Die Feldzeit startete mit dem Versand der ersten Erhebungsunterlagen in Form des Ankündigungsschreibens in KW 43/2020. Zusammen mit dem Start des Fragebogenversandes ab KW 45 wurden die Online-Versionen des Fragebogens freigeschaltet. Das erste Erinnerungsschreiben wurde ab KW 48 versendet und ab KW 51 erfolgte abschließend der Versand des zweiten Erinnerungsschreibens.

Im Anschreiben der zweiten Erinnerung wurden die Zielpersonen gebeten, den Fragebogen möglichst bis zum 31.12.2020 online auszufüllen bzw. den ausgefüllten Papierfragebogen an infas zu senden. Im Januar 2021 erreichten infas weitere ausgefüllte Fragebögen. In Abstimmung mit dem BKA wurde der 29.01.2021 als Feldende festgelegt. Die Online-Fassungen der Fragebögen wurden am 29.01.2021 auf inaktiv gesetzt, so dass ab diesem Zeitpunkt keine Fragebögen mehr online ausgefüllt werden konnten.¹⁴

4.2 Kommunikationskonzept

Die Kommunikation mit den Zielpersonen ist neben der Gestaltung der Fragebögen ein zentraler Bestandteil für eine hohe Stichprobenausschöpfung. Das Erhebungsdesign sah vor, dass den zufällig ausgewählten Personen der Bruttostichprobe die Erhebung in einem ersten Schritt per Anschreiben angekündigt und ihnen rund 14 Tage später der Fragebogen zugesandt wurde. Zudem waren bis zu zwei Erinnerungsaktionen geplant, die auch beide durchgeführt wurden.

Bei der Konzeption der vier Anschreiben (Ankündigungsschreiben, Begleitschreiben zum Fragebogen und die beiden Erinnerungsschreiben) wurde stets darauf geachtet, dass die Ansprachen der Zielpersonen aufeinander aufbauten, widerspruchsfrei waren, ausreichend über das Forschungsvorhaben und den Datenschutz aufklärten und für eine Teilnahme motivierten. Zur vollständigen Aufklärung über den Datenschutz lag dem Fragebogen jeweils eine ausführliche Datenschutzerklärung bei. Auf der Titelseite des Fragebogens wurden zusätzliche Hinweise zur Wichtigkeit der Studie und zum Datenschutz aufgeführt.

Mit dem gewählten Kommunikationskonzept sollten zum einen Vertrauen in die Erhebung und zum anderen die Motivation zu einer Teilnahme erreicht werden. Vertrauensbildende Maßnahmen waren nicht zuletzt deshalb zentral, da mit dem Fragebogen sensible und persönliche Fragestellungen verbunden waren.

¹⁴ Insgesamt trafen nach dem Feldende am 29.01.2021 noch 154 Papierfragebögen bei infas ein. Ihre Inhalte wurden nicht mehr erfasst.

Die Inhalte aller Anschreiben wurden aufeinander abgestimmt. Die Textlängen unterschieden sich erkennbar und auch die inhaltliche Argumentation zwischen den Anschreiben variierte mehrfach. Die unterschiedlichen Anschreibenvarianten sollten möglichst alle Zielpersonen ansprechen und sie für eine Teilnahme an der Erhebung werben. Um die Zielpersonen in den 16 Bundesländern möglichst konkret anzusprechen, wurden auch landesspezifische Textbausteine in die Anschreiben integriert. Zudem hatte jeweils ein leitender Vertreter der Landespolizei bzw. des Landeskriminalamts oder des Landesinnenministeriums neben dem Präsidenten des BKA und dem Geschäftsführer von infas die Anschreiben unterzeichnet.

Bereits im Ankündigungsschreiben wurden die wissenschaftlichen Ziele der Erhebung vorgestellt und betont, dass die Zielpersonen durch Ihre Teilnahme wichtige Erkenntnisse zur aktuellen Sicherheitslage beitragen konnten. Das Begleitschreiben, welches in der zweiten Versandwelle gemeinsam mit dem Fragebogen versendet wurde, enthielt vertiefende Erläuterungen zu den zentralen Themeninhalten und Fragestellungen der Befragung und die erneute Bitte, die Erhebung durch die eigene Teilnahme zu unterstützen. Ergänzend erhielt das Begleitschreiben Kontaktdaten von Hilfsangeboten für Betroffene von Straftaten. Zudem wurden die Zielpersonen aufgeklärt, dass die Teilnahme an der Befragung nicht geeignet ist, eine Straftat anzuzeigen. Es wurde darum gebeten, sich für eine Strafanzeige an die örtliche Polizeidienststelle zu wenden.

Die beiden Erinnerungsschreiben wurden als kombinierte Dank- und Erinnerungsschreiben konzipiert. Das erste Erinnerungsschreiben wurde bewusst kurz gehalten. Der Text passte auf eine Seite. Das zweite Erinnerungsschreiben umfasste eine ausführlichere Argumentation, so dass das Anschreiben auf zwei Seiten gedruckt wurde. Aufgrund der zeitlichen Nähe der aufeinander folgenden Versandaktionen musste davon ausgegangen werden, dass eine nennenswerte Anzahl Fragebögen parallel zum Versand der Erinnerungsschreiben von den Zielpersonen im Internet ausgefüllt bzw. ausgefüllte Papierfragebögen an infas gesendet werden. Um Irritationen bei den Zielpersonen vorzubeugen, wurde das Erinnerungsschreiben um eine Komponente erweitert, die den Dank an alle richtete, die bereits teilgenommen hatten. Damit möglichst wenig Personen ein Erinnerungsanschreiben erhielten, obwohl sie bereits teilgenommen hatten, wurde vor dem Versand der Erinnerungsschreiben die Stichprobe um Personen bereinigt, für die bereits ein Fragebogen vorlag. Zusätzlich wurden auch die Personen nicht erneut angeschrieben, die bereits explizit ihre Teilnahme verweigert hatten oder für die bekannt war, dass sie ins Ausland verzogen oder verstorben waren bzw. aufgrund der bisherigen Versandaktionen bekannt war, dass die vorliegende Adresse nicht mehr aktuell war.

Dem zweiten Erinnerungsschreiben lagen erneut ein Fragebogen und ein an infas adressierter und frankierter Rückumschlag bei. Sowohl beim Erstversand des Fragebogens als auch beim Zweitversand des Bogens mit der zweiten Erinnerung erhielten die Zielpersonen eine Datenschutzerklärung. Mit diesem ausführlichen Dokument wurden die Zielpersonen unter anderem über den Zweck und Rechtsgrundlage der Verarbeitung ihrer Daten, die Sicherheit der Datenverarbeitung, ihre Betroffenenrechte sowie ihr Beschwerderecht aufgeklärt. Ebenso wurden die Zielpersonen darüber unterrichtet, dass das BKA die datenschutzrechtli-

che Verantwortung für die Bevölkerungsbefragung trägt und infas die Rolle als Auftragsverarbeiter zukommt.

Grundsätzlich wurden auch die Erhebungsunterlagen für das Methodenexperiment anhand des Kommunikationskonzeptes konzipiert. Sie wurden auf der Grundlage der Anschreiben für die Basisstichprobe erstellt und entsprechend der Vorgehensweise im Methodenexperiment – soweit es notwendig war – angepasst. Es wurde somit sichergestellt, dass mögliche Effekte bei der Teilnahme auf die Unterschiede im Design zurückzuführen sind.

Neben den genannten schriftlichen Unterlagen bestand ebenfalls die Möglichkeit, dass sich die Zielpersonen im Internet über die Befragung informieren. Das BKA setzte zu diesem Zweck ausführliche Informationen auf seine Webseite. infas informierte ebenfalls auf einer studienspezifischen Webseite über das Ziel der Erhebung. Zusätzlich verfasste infas einen Kurztext für die Rubrik „Aktuelle Befragungen“ auf der Internetseite des Instituts. Die Internetseiten des BKA und infas waren miteinander verlinkt, um die Zusammenarbeit deutlich zu machen und die Seriosität der Erhebung zu unterstreichen.

Den Zielpersonen standen während der gesamten Feldzeit zusätzliche Möglichkeiten zur Verfügung, direkt mit dem BKA als Träger der datenschutzrechtlichen Verantwortung in Kontakt zu treten. Dafür gab es zum einen eine Festnetzrufnummer sowie die studienspezifische E-Mail-Adresse des BKA.¹⁵ Die Rufnummer und die E-Mail-Adresse waren auf sämtlichen Erhebungsunterlagen aufgeführt. infas unterstützte das BKA bei der Erstellung der Schulungsunterlagen für die Hotline-Mitarbeiter beim BKA. Bei Bedarf unterstützte infas das BKA bei der Beantwortung der Rückfragen von Zielpersonen oder Angehörigen.

Wenn sich in Ausnahmefällen Zielpersonen direkt an infas wendeten, antworteten die geschulten Mitarbeiterinnen und Mitarbeiter der infas-Hotline auf die Fragen der Anrufer und informierten im Bedarfsfall das BKA.

4.3 Feldsteuerung

Das Studiendesign sah insgesamt vier Versandaktionen vor.¹⁶ Für einen zielgerichteten Versand der Erhebungsunterlagen war es erforderlich, dass alle relevanten Informationen zur Teilnahme, zur Aktualität der Kontaktdaten, zur Teilnahmeverweigerung und anderen Ausfällen zeitnah im infas Sample-Management-System vorlagen. Die Rückmeldungen vom BKA zu Teilnahmeverweigerungen, Adresskorrekturen etc. wurden bei infas zeitnah im System

¹⁵ Auf die zusätzliche Nennung der Kontaktdaten von infas wurde nach eingehender Beratung verzichtet.

¹⁶ Die zweite Erinnerung war als optionale Leistung ausgeschrieben. Es war ursprünglich vorgesehen, über ihren Einsatz in Abhängigkeit vom Stand der Fragebogenrückläufe nach der ersten Erinnerungsaktion zu entscheiden. Eine Entscheidung für die Durchführung der zweiten Erinnerung erst aufgrund der Rücklaufentwicklung nach der ersten Erinnerung, hätte in der Praxis eine deutliche Verzögerung des Versands des zweiten Erinnerungsschreibens bedeutet. Um die zweite Erinnerung ohne Verzögerung durchführen zu können, beauftragte das BKA bereits vor der ersten Erinnerung die Durchführung der zweiten Erinnerungsaktion. Somit wurde letztlich in allen Bundesländern eine zweite Erinnerungsaktion durchgeführt. Für das Methodenexperiment waren von Beginn an zwei Erinnerungen vorgesehen.

erfasst. Die zügige Bearbeitung der Rückmeldungen garantierte einen aktuellen Stand der Stichprobe für jede weitere Versandaktion.

Die hohe Anzahl Versendungen benötigte einen gewissen zeitlichen Vorlauf und entsprechend Zeit für die Konfektionierung und Postauslieferung der Versandtaschen bzw. Briefumschläge. Die ersten beiden Versandaktionen waren zeitlich sehr eng getaktet, so dass bis auf vier keine weiteren Rückmeldungen aus dem Feld für den Versand des Fragebogens berücksichtigt werden konnten.¹⁷ Ankündigungsschreiben und Fragebogenversand richteten sich an den (nahezu) gleichen Stichprobenumfang (d.h. gleiche Anzahl an Zielpersonen).

Tabelle 9 Fallzahlen pro Versandaktion

Versandaktion	Basisstichprobe		Methodenexperiment		Gesamt	
	abs.	%	abs.	%	abs.	%
Ankündigungsschreiben	122.667	100,0	998	100,0	123.665	100,0
Basisstichprobe: Anschreiben, Fragebogen mit Zugangsdaten für Online- Fragebogen und Datenschutz- erklärung	122.663	100,0	998	100,0	123.661	100,0
Experimentalgruppe: Anschreiben und Beiblatt mit Zugangsdaten für Online-Frage- bogen und Datenschutzerklärung						
1. Erinnerungsschreiben	111.786	91,1	822	82,4	112.608	91,1
2. Erinnerungsschreiben mit Frage- bogen und Datenschutzerklärung	85.593	69,8	755	75,7	86.348	69,8

Quelle: infas, eigene Berechnungen

Wie weiter oben ausgeführt, wurden neben 16 bundeslandspezifischen Anschreiben auch 16 bundeslandspezifische Anschreiben für das Methodenexperiment eingesetzt. In den Bundesländern NRW und Hamburg kam ein um vier landesspezifische Seiten ergänzter Kernfragebogen zum Einsatz. Durch den Einsatz der türkischsprachigen Erhebungsunterlagen ergaben sich pro Versandaktion in Summe 56 spezifisch zu konfektionierende Versandpakete.¹⁸ Der Versandprozess erfolgte sukzessive über mehrere Tage. Dafür wurden die Bundesländer in „Versandgruppen“ eingeteilt, die immer in einer festgelegten Reihenfolge im Rahmen des Versands abgearbeitet wurden. Das Vorgehen garantierte, dass über alle Versandaktionen hinweg die Abstände zwischen den Versandaktionen in jedem Bundesland identisch waren. Daraus folgte, dass alle Zielpersonen unabhängig von der Bundeslandzugehörigkeit die Erhebungsunterlagen immer mit demselben zeitlichen Abstand erhielten.

¹⁷ Bei den vier Fällen handelt es sich um Teilnahmeverweigerungen.

¹⁸ Die 56 Varianten verteilen sich wie folgt: 16 Varianten Basisstichprobe, 16 Varianten Basisstichprobe türkisch, 16 Varianten Methodenexperiment, 8 Varianten Methodenexperiment türkisch (in der vergleichsweise kleinen Stichprobe für das Methodenexperiment wurden nur in 8 von 16 Bundesländern Personen mit vermutlich türkischem Migrationshintergrund identifiziert).

Tabelle 10 Versandtermine

Versandaktion	Versandzeitraum
Ankündigungsschreiben:	KW 43/KW 44
Basisstichprobe: Anschreiben, Fragebogen mit Zugangsdaten für Online-Fragebogen und Datenschutzerklärung	KW 45/KW 46
Experimentalgruppe: Anschreiben und Beiblatt mit Zugangsdaten für Online-Fragebogen und Datenschutzerklärung	
1. Erinnerungsschreiben	KW 48
2. Erinnerungsschreiben mit Fragebogen und Datenschutzblatt	KW 51

Tabelle 11 Zugehörigkeit der Bundesländern zu Versandgruppe

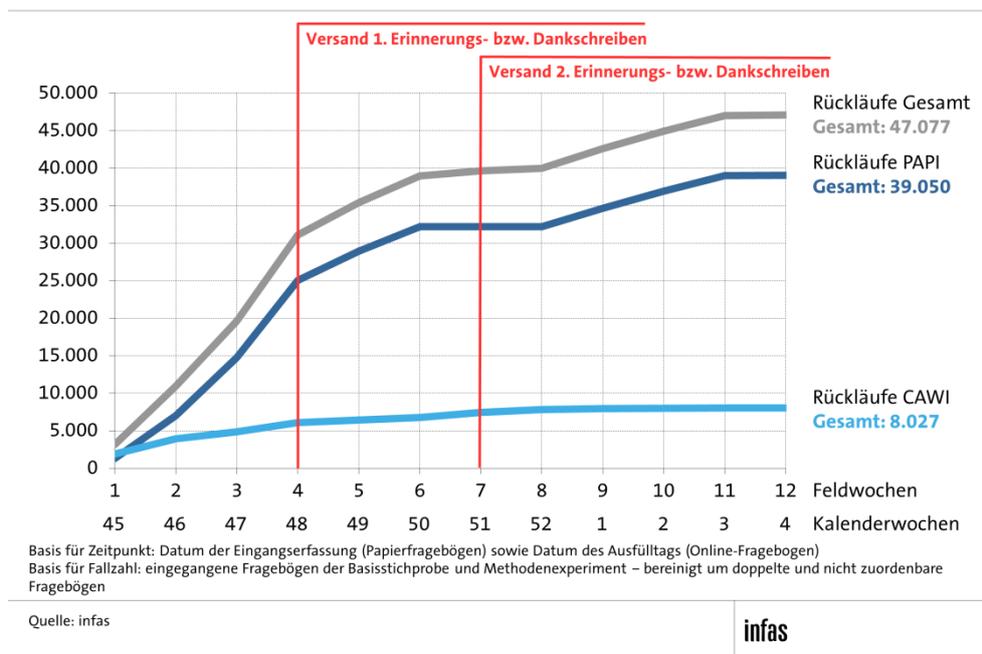
Versandgruppennummer	Bundesland
1	Bremen Nordrhein-Westfalen Saarland Mecklenburg-Vorpommern
2	Schleswig-Holstein Berlin Brandenburg Sachsen-Anhalt
3	Hamburg Hessen Rheinland-Pfalz Sachsen
4	Baden-Württemberg Bayern Niedersachsen Thüringen
5	Methodenexperiment

4.4 Feldverlauf

Bereits in den Folgetagen nach dem Versand der ersten Fragebögen lagen in der ersten Feldwoche (KW 45/2020) ausgefüllte Fragebögen vor. Abgeschlossene Online-Fragebögen lagen direkt digital vor. Die Eingangserfassung von ausgefüllten Papierfragebögen kann aufgrund des Postlaufs erst zeitverzögert erfolgen. Bereits nach wenigen Tagen überstieg jedoch die Anzahl eingegangener Papierfragebögen die Anzahl ausgefüllter Online-Fragebögen sehr deutlich (Abbildung 1¹⁹). Insgesamt entfallen rund 17 Prozent auf einen Online-Fragebogen und 83 Prozent auf die Papierfassungen.

Der Online-Rücklauf steigt in der ersten Feldwoche am stärksten an. Im weiteren Verlauf zeigt sich nach jeder Erinnerungsaktion erneut ein kleiner Anstieg, der dann wieder abflacht. Ähnlich lässt sich bei den Rückläufen der Papierfragebögen feststellen, dass die Erinnerungsaktionen – mit leichter zeitlicher Verzögerung – zu einem temporären Anstieg des Rücklaufs beigetragen haben. Der Gesamtverlauf des Felds ist stark vom Rückversand der Papierfragebögen geprägt, da deren Anteil mit 83 Prozent den Rücklauf dominiert. Der Graph der Papierfragebögen flacht im Vergleich zum Online-Instrument nach jeder Versandaktion zudem deutlich langsamer ab. Grundsätzlich ist der Feldverlauf typisch für eine schriftlich-postalische Erhebung. Erfreulich und nicht selbstverständlich ist die hohe Fallzahl der Teilnahme insgesamt.

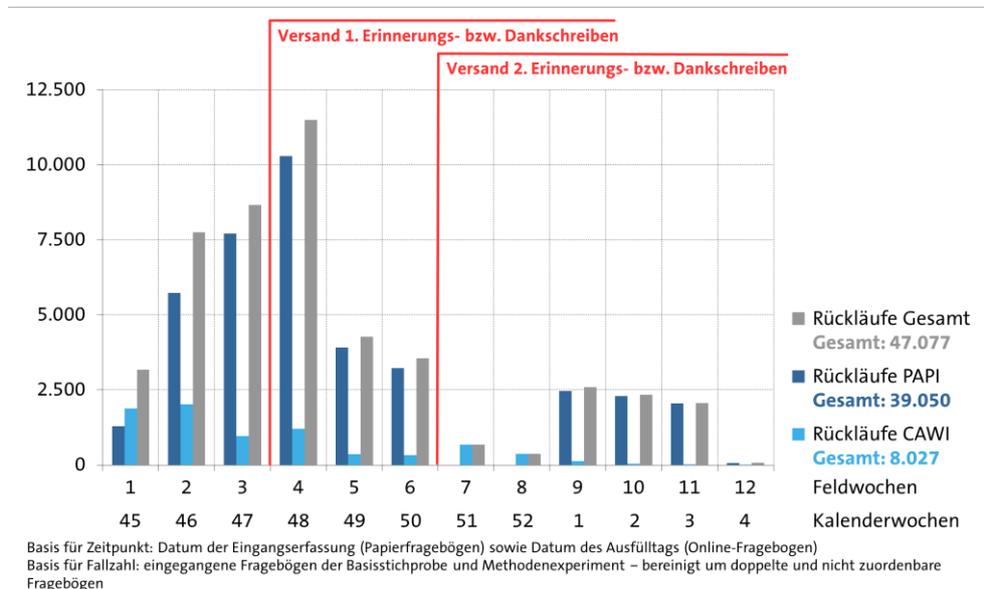
Abbildung 1 Feldverlauf Papier- und Online-Fragebögen kumuliert



¹⁹ Der Feldverlauf in Abbildung 1 weist den Zeitpunkt der Eingangserfassung von 39.050 Papierfragebögen und 8.027 Online-Fragebögen aus. Die 47.077 Fragebögen stellen die Summe dar aus den Rückläufen der Basisstichprobe und des Methodenexperiments, abzüglich der doppelt eingegangenen Fälle und jener Papierfragebögen, die aufgrund eines entfernten QR-Codes keiner konkreten Person aus der Stichprobe zugeordnet werden konnten.

Die Darstellung des Fragebogeneingangs pro Kalenderwoche in Abbildung 2 unterstreicht noch einmal die berichteten Befunde. Die erste Erinnerungsaktion wurde bereits frühzeitig gestartet, als der Rücklauf online zwar schon abflachte, beim Papierfragebogen allerdings noch zunahm. Der Effekt der zweiten Erinnerungsaktion zeigt sich beim Online-Fragebogen sofort, aber nur mit einem vergleichsweise kleinen Ausschlag nach oben. Die verzögerte Wirkung der zweiten Erinnerung beim Papierfragebogen beruht auf den infas-Betriebsferien: Der Versand der zweiten Erinnerung erfolgte kurz vor den Weihnachtsfeiertagen. Die Eingangserfassung wurde direkt nach den Betriebsferien in der ersten Kalenderwoche 2021 wieder aufgenommen. Erfreulicherweise wurde bis KW 3 noch eine nennenswerte Anzahl Papierfragebögen zurückgesendet. Online wurden im Januar 2021 hingegen fast keine Fragebögen mehr ausgefüllt. Insgesamt erstreckte sich die Feldzeit über 12 Wochen.²⁰

Abbildung 2 Fragebogeneingang pro Feldwoche



Quelle: infas

infas

²⁰ Inklusive Versand des Ankündigungsschreibens ergeben sich 14 Wochen.

4.5 Verarbeitung von Rückmeldungen aus dem Feld

Das BKA trug für die Erhebung die datenschutzrechtliche Verantwortung. Fragen zu den Betroffenenrechten sollten direkt vom BKA beantwortet werden. Für Rückfragen der Zielpersonen standen deshalb während der gesamten Feldzeit eine Festnetzrufnummer sowie eine studienspezifische E-Mail-Adresse des BKA zur Verfügung. Beide Kontaktdaten waren auf allen Erhebungsunterlagen genannt.

Nahezu sämtliche Anfragen der Zielpersonen gingen im BKA ein und wurden von dort bearbeitet. infas erhielt zur Steuerung der einzelnen Versandaktionen regelmäßig ausgewählte Informationen vom BKA. Es handelt sich dabei um Informationen, die bei infas

- zur Löschung von Adressdaten und/oder Befragungsdaten,
- zur Sperrung der Adresse für weitere Versendungen aus anderen Gründen
- oder zur Korrektur von Kontaktdaten bzw. Verwendung aktueller Kontaktdaten für zukünftige Versendungen benötigt wurden.

Darüber hinaus unterstützte infas das BKA bei Auskunftersuchen von Zielpersonen zu gespeicherten personenbezogenen Daten. Es wurde ein regelmäßiger Informationsaustausch etabliert, damit beidseitig zeitnah relevante Informationen vorlagen. Äußerten Zielpersonen den Wunsch, das Anschreiben und/oder den Fragebogen erneut zugesendet zu bekommen, wurden sie vom BKA informiert, dass ein erneuter Versand des Fragebogens für den Dezember 2020 geplant war.

Nachdem das BKA Rückmeldungen von Zielpersonen an infas übermittelte, wurden diejenigen Informationen, die zur Feldsteuerung benötigt wurden, in das infas-Sample-Management-System eingepflegt. Auf diese Weise wurde sichergestellt, dass alle stichprobenrelevanten Informationen zugänglich vorlagen und für die nächste Versandaktion berücksichtigt werden konnten.

Für den Umgang mit Löschanforderungen wurde ein ergänzender Prozessschritt vereinbart. Dieser sah vor, dass sobald die Adressen der betroffenen Zielperson bei infas gelöscht wurden, die Übermittlung der Löschbestätigungen ans BKA erfolgte. Nach der erfolgten Übermittlung der Übergabedateien wurden auch diese Dateien abschließend bei infas gelöscht.

In den Fällen, in denen die Zielpersonen Unterstützung beim Online-Fragebogen benötigten, nahmen die geschulten infas-Mitarbeiter Kontakt zu den Zielpersonen auf, um eine entsprechende Hilfestellung zu geben. Eine direkte Weiterleitung an die infas Hotline war für die Fälle vorgesehen, wenn die Zielpersonen beim Kontakt mit dem BKA akute technische Probleme mit dem Online-Fragebogen hatten.

Auch wenn das BKA als einziger Ansprechpartner in den Erhebungsunterlagen genannt war, kam es vor, dass sich einzelne Zielpersonen mit einem Anliegen direkt an infas wandten. Unabhängig davon, ob diese Anliegen per E-Mail oder per Telefon bei infas eintrafen, wurden sie von dafür geschulten Mitarbeitern

aufgenommen und weiterbearbeitet. Sofern es sich dabei um Betroffenenanfragen nach Datenschutz-Grundverordnung (DSGVO) handelte, wie z.B. die Bitte um Adresslöschung, wurden diese Anfragen ans BKA weitergeleitet, um auch für diese Fälle eine Bearbeitung beim BKA wie oben beschrieben zu gewährleisten.

Sofern es sich um Anliegen oder Rückfragen zum Ausfüllen des Papier- oder Online-Fragebogens handelte, wurde den Zielpersonen ad hoc weitergeholfen.

In einigen Fällen legten die Zielpersonen dem zurückgesendeten Fragebogen zusätzliches Schriftgut bei. Das BKA erhielt physische Kopien dieser Schriftstücke, die vor der Übergabe ans BKA anonymisiert wurden. Die Originale wurden bei infas nach Absprache mit dem BKA vernichtet.

4.6 Stichprobenausschöpfung

Die Analyse des Feldendstandes wird getrennt für die Basisstichprobe und das Methodenexperiment vorgenommen. Die Stichprobenausschöpfung im Methodenexperiment wird dabei mit der Ausschöpfung in der Kontrollgruppe verglichen. Die Kontrollgruppe ist als Bestandteil der Basisstichprobe auch in der Analyse der Basisstichprobe enthalten.

4.6.1 Basisstichprobe

Von den 122.667 ausgewählten Zielpersonen liegt für 46.813 Personen ein vollständig oder teilweise ausgefüllter Fragebogen²¹ vor. Dies entspricht einem Anteil von 38,2 Prozent der Bruttostichprobe. Abzüglich der nicht auswertbaren Fälle²² (n=1.462) verbleiben 45.351 Fragebögen (37,0 Prozent der Bruttostichprobe) für die inhaltliche Auswertung.²³ Bereinigt um die stichprobenneutralen Ausfälle (7.617 Zielperson sind verzogen oder verstorben) steigt die Ausschöpfungsquote insgesamt auf 40,7 Prozent. Bezogen auf die auswertbaren Fälle wurde eine Stichprobenausschöpfung von 39,4 Prozent der bereinigten Bruttostichprobe erreicht. Die Stichprobenausschöpfung ist erfreulich hoch und kann als sehr gut bewertet werden.

Wie üblich bei postalischen Befragungen liegt mit 66.496 Fällen für mehr als die Hälfte der Bruttostichprobe keine Rückmeldung vor. Lediglich 1,4 Prozent der Bruttostichprobe (entspricht 1.740 Personen) hat die Teilnahme explizit verweigert.²⁴

²¹ In dieser Summe sind auch die 67 PAPI-Fragebögen enthalten, deren QR-Code nicht lesbar war.

²² Die nicht auswertbaren 1.462 Fälle enthalten 67 Papierfragebögen ohne lesbaren QR-Code sowie 2 bereits vor der Datenprüfung bei infas vom BKA ausgeschlossene Fragebögen.

²³ Die Ermittlung der nicht auswertbaren Fälle ist in Kapitel 5 beschrieben.

²⁴ Die Rückmeldungen der Zielpersonen erreichten in den allermeisten Fällen das BKA (s. Kapitel 4.5). Die in Tabelle 12 ausgewiesenen Rückmeldungen basieren auf den vom BKA übermittelten Informationen zu Rückmeldungen von Ziel-

Tabelle 12 Stichprobenausschöpfung

	abs.	%	abs.	%
Bruttostichprobe	122.667	100,0		
Nicht Zielgruppe	7.617	6,2		
Zielperson verstorben	328	0,3		
Anschreiben unzustellbar, Zielperson ins Ausland verzogen	7.289	5,9		
Bereinigte Bruttostichprobe	115.050	93,8	115.050	100,0
Nonresponse – verweigert	1.740	1,4	1.740	1,5
Zielperson verweigert die Teilnahme grund- sätzlich/bittet um Adresslöschung	57	0,0	57	0,0
Zielperson verweigert Start des Interviews/ Fragebogen leer zurück	360	0,3	360	0,3
Sonstige Teilnahmeverweigerung	1.323	1,1	1.323	1,1
Nonresponse – Sonstige	66.497	54,2	66.497	57,8
Keine Rückmeldung	66.496	54,2	6.6496	57,8
Zielperson wünscht erneute Zusendung des Fragebogens/des Online-Zugangscodes	1	0,0	1	0,0
Rückgesendete Fragebögen (mindestens eine Frage beantwortet, um doppelt ausgefüllte Fragebögen bereinigt)	46.813	38,2	46.813	40,7
Nicht auswertbare Fälle	1.462	1,2	1.462	1,3
Auswertbare Fragebögen	45.351	37,0	45.351	39,4

Quelle: Rücklaufdatenbank, eigene Berechnungen

personen oder anderen Dritten an infas. Der Fragebogenrücklauf erfolgte direkt an infas und wurde wie in Kapitel 5.1 beschrieben erfasst.

Der Verteilungsvergleich zwischen der Bruttostichprobe und den auswertbaren Fällen zeigt, dass sich die auswertbaren Fälle nahezu in gleicher Weise über die Bundesländer verteilen wie in der Bruttostichprobe. Es konnten sehr ähnliche bis nahezu identische Ausschöpfungen in allen 16 Bundesländern erzielt werden. Lediglich Schleswig-Holstein sticht mit einer erkennbar höheren Bruttoausschöpfung heraus. Die Differenz zwischen dem Anteil der Bruttostichprobe und den auswertbaren Fällen jeweils an der Summe über alle Bundesländer beträgt 1,2 Prozentpunkte. In Nordrhein-Westfalen ist mit minus 0,8 Prozentpunkten die größte Abweichung nach unten erkennbar.

Tabelle 13 Vergleich Bruttostichprobe und auswertbare Fälle nach Bundesland

Bundesland	Bruttostichprobe		Auswertbare Fragebögen		Differenz %-Punkte
	abs.	%	abs.	%	
<i>Spalten%</i>					
Schleswig-Holstein	23.500	19,2	9.239	20,4	1,2
Hamburg	23.004	18,8	8.551	18,9	0,1
Niedersachsen	2.915	2,4	1.046	2,3	-0,1
Bremen	265	0,2	103	0,2	0,0
Nordrhein-Westfalen	37.557	30,6	13.505	29,8	-0,8
Hessen	1.951	1,6	671	1,5	-0,1
Rheinland-Pfalz	1.431	1,2	493	1,1	-0,1
Baden-Württemberg	3.708	3,0	1.486	3,3	0,3
Bayern	4.394	3,6	1.695	3,7	0,1
Saarland	318	0,3	103	0,2	-0,1
Berlin	1.231	1,0	453	1,0	0,0
Brandenburg	949	0,8	331	0,7	-0,1
Mecklenburg-Vorpommern	583	0,5	209	0,5	0,0
Sachsen	1.271	1,0	458	1,0	0,0
Sachsen-Anhalt	790	0,6	248	0,5	-0,1
Thüringen	18.800	15,3	6.760	14,9	-0,4
Gesamt	122.667	100,0	45.351	100,0	

Quelle: Rücklaufdatenbank, eigene Berechnungen

Die auswertbaren Fälle verteilen sich zu 16,5 Prozent auf die CAWI-Fragebögen und zu 83,5 Prozent auf die Papierfragebögen.

Tabelle 14 Auswertbare Fälle getrennt nach PAPI/CAWI

Realisierungsmethode	abs.	%
Online-Fragebogen	7.491	16,5
Papierfragebogen	37.860	83,5
Gesamt	45.351	100,0

Quelle: Rücklaufdatenbank, eigene Berechnungen

Von den 45.351 auswertbaren Fragebögen sind 338 in einer der drei Fremdsprachen (Türkisch, Russisch, Arabisch) beantwortet worden. Von diesen 338 Fragebögen entfallen 195 auf die türkische Fassung des Papierfragebogens. Die restlichen 143 Fragebögen sind online ausgefüllt worden; 26 auf Türkisch, 52 auf Russisch und 65 auf Arabisch. Zielpersonen, die sich für das Ausfüllen des türkischsprachigen Fragebogens entschieden hatten, wählten mehrheitlich die Papierfassung.

Tabelle 15 Auswertbare Fragebögen nach Sprache und Methode

Spalten%	Papierfragebogen		Online-Fragebogen		Gesamt	
	abs.	%	abs.	%	abs.	%
Deutsch	37.665	99,5	7.348	98,1	45.013	99,3
Türkisch	195	0,5	26	0,3	221	0,5
Russisch	-	-	52	0,7	52	0,1
Arabisch	-	-	65	0,9	65	0,1
Gesamt	37.860	100,0	7.491	100,0	45.351	100,0

Quelle: Rücklaufdatenbank, eigene Berechnungen

4.6.2 Methodenexperiment

Das Methodenexperiment sah vor, dass die ausgewählten Zielpersonen erst mit der zweiten Erinnerungsaktion neben den Online-Zugangsdaten auch den Papierfragebogen erhalten. Als Kontrollgruppe war im gleichen Umfang eine Stichprobe aus der Basisstichprobe ausgewählt worden (s. Kapitel 2.8). Die Kontrollgruppe erhielt bereits mit der zweiten Versandaktion neben den Zugangsdaten für den Online-Fragebogen auch eine Papierfassung des Erhebungsinstruments.

Die Bruttoausschöpfung liegt in der Kontrollgruppe um 6 Prozentpunkte höher. Nach der Datenprüfung verbleiben für die Kontrollgruppe 380 auswertbare Fragebögen (38 Prozent), für die Experimentalgruppe sind es mit 315 (rund 32 Prozent) deutlich weniger.

Tabelle 16 Stichprobenausschöpfung

Spalten%	Experimentalgruppe		Kontrollgruppe		Gesamt	
	abs.	%	abs.	%	abs.	%
Bruttostichprobe	998	100,0	998	100,0	1.996	100,0
Nicht in Zielgruppe	48	4,8	60	6,0	108	5,4
Zielperson verstorben	2	0,2	5	0,5	7	0,4
Anschreiben unzustellbar, Zielperson ins Ausland verzogen	46	4,6	55	5,5	101	5,1
Nonresponse – verweigert	9	0,9	13	1,3	22	1,1
Zielperson verweigert Start des Interviews/Fragebogen leer zurück	2	0,2	1	0,1	3	0,2
Sonstige Teilnahmeverweigerung	7	0,7	12	1,2	19	1,0
Nonresponse – Sonstige	610	61,1	533	53,4	1.143	57,3
Keine Rückmeldung	609	61,0	533	53,4	1.142	57,2
Zielperson wünscht erneute Zusendung des Fragebogens/Online-Zugangscode	1	0,1	-	-	1	0,1
Rückgesendete Fragebögen (mindestens eine Frage beantwortet, um doppelt ausgefüllte Fragebögen bereinigt)	331	33,2	392	39,3	723	36,2
Nicht auswertbare Fälle	16	1,6	12	1,2	28	1,4
Auswertbare Fragebögen	315	31,6	380	38,1	695	34,8

Quelle: Rücklaufdatenbank, eigene Berechnungen

Erwartungsgemäß verteilen sich die Rückläufe unterschiedlich auf die Methoden CAWI und Papierfragebogen. Während in der Kontrollgruppe der Anteil der Papierfragebogen unter den auswertbaren Fällen mit rund 83 Prozent auf dem Niveau der Basisstichprobe liegt, haben in der Experimentalgruppe des Methodenexperiments rund 76 Prozent den Fragebogen online ausgefüllt. Immerhin liegt für ein knappes Viertel der Antwortenden der Experimentalgruppe ein auswertbarer Papierfragebogen vor, obwohl der Papierfragebogen in dieser Teilstichprobe nur einmal, und zwar erst mit der 2. Erinnerung, verschickt wurde.

Tabelle 17 PAPI versus CAWI-Rücklauf getrennt nach Stichproben

Realisierungsmethode	Experimentalgruppe		Kontrollgruppe	
	abs.	%	abs.	%
Online-Fragebogen (CAWI)	239	75,9	66	17,4
Papierfragebogen	76	24,1	314	82,6
Gesamt	315	100,0	380	100,0

Quelle: Rücklaufdatenbank, eigene Berechnungen

4.7 Interviewdauer CAWI-Fragebogen

Für das Ausfüllen des Online-Fragebogen benötigten die 7.540 Zielpersonen im Durchschnitt rund 27 Minuten²⁵. Die Ausfülldauer variiert zwischen knapp 6 Minuten und rund 5 Stunden. Die extrem lange gemessene Zeit ist ein deutlicher Ausreißer. Die nächst niedrigere Ausfülldauer liegt bei 159 Minuten. Abgesehen von dem Maximalwert liegen lediglich acht weitere Zeitmessungen über 2 Stunden.

Tabelle 18 Ausfülldauer vollständig bearbeitete CAWI-Fragebogen

Anzahl der Interviews	7.540
Mittelwert	27,2
Standardabweichung	14,5
Minimum	5,8
Maximum	312,2

Quelle: infas, eigene Berechnung

4.8 Abbrüche im Fragebogen

In 480 Fällen ist der zurückgesandte Papierfragebogen nicht bis zur letzten Frage ausgefüllt worden, das heißt das Ausfüllen wurde vorzeitig abgebrochen. Die Abbrüche verteilen sich über den gesamten Papierfragebogen.

Im Kernfragebogen gibt es lediglich einen erhöhten Anteil Abbrüche nach den Fragen 15 (Matrixabfrage zu Erfahrungen mit Kriminalität ohne Gewaltanwendung), 29 (Nachfrage zu Bedrohungen bzw. körperlicher oder sexueller Angriffen durch den (Ex-)Partner/die (Ex-)Partnerin), 46 (subjektiver sozialer Status), 51 und 52 (Fragen zur eigenen Staatsangehörigkeit) sowie Frage 53 (Geburtsland der Eltern). Der höchste Anteil Abbrüche ist dabei nach Frage 53 zu verzeichnen.

Im Hamburger Fragenmodul, das sich an den Kernfragebogen anschließt, gibt es ebenfalls Fragen, nach welchen ein vergleichsweise hoher Anteil an Abbrüchen erfolgte. Dazu zählen insbesondere die Fragen 60 bis 62, 66 und, in etwas geringerem Umfang, die Fragen 63 und 65. Bei diesen Fragen wird nach dem Interesse der Zielperson an und ihre Wahrnehmung von polizeilichen Informationen sowie zu Zeugenerlebnissen in den letzten 12 Monaten gefragt. Die meisten Abbrüche sind dabei nach Frage 66 zu verzeichnen. Mit dieser Frage sollte ermittelt werden, weshalb die Zielperson die Polizei über beobachtete Straftaten informiert bzw. nicht informiert hat.

²⁵ Die Dauerberechnung basiert auf allen abgeschlossenen CAWI-Fragebögen. D.h. nicht berücksichtigt sind CAWI-Fragebögen, die nicht bis zum Ende bearbeitet, also vorzeitig abgebrochen wurden.

Im nordrhein-westfälischen Modul, das ebenfalls an den Kernfragebogen anschließt, steigt die Abbruchquote nach den Fragen 62 bis 64 an und ist nach Frage 65 am höchsten. Die Frage 62 bezieht sich auf die direkte Nachbarschaft, Fragen 63 bis 65 beschäftigen sich mit dem Wohnort der Zielperson.

Tabelle 19 Abbruchstellen im Papierfragebogen

Fragennummer	<i>abs.</i>	%
F02	3	0,63
F03	2	0,42
F04	3	0,63
F05	2	0,42
F06	3	0,63
F09	4	0,83
F10b	1	0,21
F11	4	0,83
F12	2	0,42
F13	7	1,46
F14	1	0,21
F15	15	3,13
F16	2	0,42
F18	4	0,83
F23	1	0,21
F24	7	1,46
F25	1	0,21
F27	5	1,04
F28	5	1,04
F29	10	2,08
F31	3	0,63
F35	1	0,21
F36	1	0,21
F37	1	0,21
F38	1	0,21
F41	2	0,42
F42	2	0,42
F44	3	0,63
F45	5	1,04
F46	14	2,92
F47	1	0,21
F48	1	0,21

Fragennummer	<i>abs.</i>	%
F49	7	1,46
F50	3	0,63
F51	10	2,08
F52	10	2,08
F53	39	8,13
F54	2	0,42
<i>Hamburger Modul</i>		
HH_F56	1	0,21
HH_F57	4	0,83
HH_F58	3	0,63
HH_F59	7	1,46
HH_F60	24	5,00
HH_F61	42	8,75
HH_F62	33	6,88
HH_F63	10	2,08
HH_F64	6	1,25
HH_F65	13	2,71
HH_F66	73	15,21
<i>Nordrhein-westfälisches Modul</i>		
NRW_F56	2	0,42
NRW_F57	4	0,83
NRW_F59	4	0,83
NRW_F61	5	1,04
NRW_F62	11	2,29
NRW_F63	11	2,29
NRW_F64	8	1,67
NRW_F65	31	6,46
Gesamt	480	100

Quelle: eigene Auswertungen

In 249 Fällen wurde der CAWI-Fragebogen nicht bis zum Ende ausgefüllt. Ähnlich wie im Papierfragebogen verteilen sich die Abbrüche über den gesamten Fragebogen. Anders als im Papierfragebogen gibt es im CAWI-Fragebogen eine deutliche Häufung zu Beginn des Bogens. Der mit Abstand größte Anteil der Abbrüche ereignete sich nach der Frage 3 (Beginn der Fragen zur Bewertung der Polizeiarbeit.). Auch bei einigen weiteren Fragen liegt die Abbruchquote nach Beantwortung dieser Fragen etwas erhöht. Dies betrifft die Fragen 2 (Anzeichen physischer und sozialer Unordnung in der Wohngegend), 12 (Wahrscheinlichkeit, dass man von ausgewählten Straftaten betroffen sein wird), 13 (Maßnah-

men, um sich vor Kriminalität zu schützen), 15 (Matrixabfrage zu Erfahrungen mit Kriminalität ohne Gewaltanwendung), 24 (Matrixabfrage zu Erfahrungen mit Gewaltkriminalität) und 54 (Staatsangehörigkeit der Eltern).

Im Hamburger Modul gibt es mit 6 Abbrüchen insgesamt keine Auffälligkeiten. Sie ereignen sich im Anschluss an die Beantwortung der Fragen 59, 61, 64 und 66.

Im nordrhein-westfälischen Modul verteilen sich die wenigen Abbrüche ebenfalls über eine Reihe von Fragen zwischen den Fragen 56 und 66. Absolut gab es mit 7 die meisten Abbrüche nach der Frage 66. Bei der Frage 66 handelte sich lediglich um eine offene Nachfrage. Dort konnten die Zielpersonen optional einen Text eintragen, wenn sie noch ein Thema benennen mochten, dass aus ihrer Sicht für die Sicherheit wichtig ist. Es handelt sich somit weniger um einen Abbruch im Fragebogen, der fehlende inhaltliche Angaben zur Konsequenz hätte.

Tabelle 20 Abbruchstellen im CAWI-Fragebogen

Frage	abs.	%
F01	5	2,01
F02	14	5,62
F03	46	18,47
F04	8	3,21
F05	8	3,21
F06	2	0,80
F08	1	0,40
F09	3	1,20
F10a	1	0,40
F10b	6	2,41
F11	5	2,01
F12	10	4,02
F13	15	6,02
F14	7	2,81
F15	16	6,43
F18	3	1,20
F23	2	0,80
F24	14	5,62
F32	2	0,80
F34	2	0,80
F35	2	0,80
F38	2	0,80
F39	1	0,40
F40	1	0,40

Frage	<i>abs.</i>	%
F41	1	0,40
F42	1	0,40
F43	1	0,40
F44	2	0,80
F45	4	1,61
F46	2	0,80
F47	1	0,40
F48	3	1,20
F49	1	0,40
F50	1	0,40
F51	1	0,40
F53	7	2,81
F54	11	4,42
<i>Hamburger Modul</i>		
HH_F59	2	0,80
HH_F61	1	0,40
HH_F64	2	0,80
HH_F66	1	0,40
<i>Nordrhein-westfälisches Modul</i>		
NRW_F56	1	0,40
NRW_F57	2	0,80
NRW_F58	1	0,40
NRW_F59	2	0,80
NRW_F60	1	0,40
NRW_F61	2	0,80
NRW_F62	7	2,81
NRW_F63	2	0,80
NRW_F64	4	1,61
NRW_F65	2	0,80
NRW_F66	7	2,81
Gesamt	249	100

Quelle: eigene Auswertungen

5 Datenerfassung, -prüfung und -bereinigung

5.1 Eingangserfassung Papierfragebögen

Insgesamt wurden 39.459 Papierfragebögen innerhalb der Feldzeit an infas zurückgeschickt. Jeder dieser Papierfragebögen wurde in einem ersten Schritt einer formalen Prüfung unterzogen und für die Weiterverarbeitung sortiert. Die Eingangskontrolle umfasste die Prüfung, ob mindestens eine Frage im Bogen beantwortet war und die Qualität der Kreuze im gesamten Fragebogen. Für die maschinelle Erfassung musste gut erkennbar sein, an welcher Stelle das Antwortkreuz gesetzt ist. Wenn die Kreuze nicht stark genug waren, wurden sie vor der Erfassung nachgezeichnet, damit keine Informationen verloren gingen. Nach der Sichtung und Prüfung der Ausfüllqualität und Vollständigkeit wurden die Fragebögen nach unterschiedlichen Merkmalen sortiert.

Vollständig leere Fragebögen und Bögen, auf denen lediglich die Teilnahmeverweigerung notiert war, wurden aussortiert. Zur Datenerfassung wurden nur Fragebögen weitergeleitet, wenn mindestens eine Frage beantwortet war. Anschließend wurden die ausgefüllten Fragebögen zur Vorbereitung der maschinellen Erfassung nach der jeweiligen Version sortiert:

- Kernfragebogen deutsch
- Fragebogen mit Hamburger Modul deutsch
- Fragebogen mit NRW-Modul deutsch
- Kernfragebogen türkisch
- Fragebogen mit Hamburger Modul türkisch
- Fragebogen mit NRW-Modul türkisch

Weitere Sortiermerkmale für die Dokumentation der eingegangenen Fragebögen waren doppelt ausgefüllte Fragebögen und Fragebögen mit nicht mehr lesbarem QR-Code.

Doppelte Papierfragebögen²⁶ und Fragebögen mit nicht mehr lesbarem QR-Code²⁷ wurden nach der Eingangsprüfung aussortiert und nicht in der Rücklaufdatenbank erfasst. Für alle anderen Fragebögen wurde im Anschluss an die Prüfung und Sortierung der Eingang der Fragebögen anhand der Fallnummer elektronisch erfasst.

²⁶ Nicht mehr erfasst wurde jeweils der später eintreffende Papierfragebogen. Der Abgleich wurde bei der elektronischen Eingangserfassung vorgenommen. War bereits der Eingang eines Papierfragebogens unter einer Fallnummer erfasst, wurde der zweite Bogen nicht mehr berücksichtigt. Dies betraf insgesamt 294 Fälle.

²⁷ Eine Erfassung in der Datenbank war nicht möglich, da diese 67 Papierfragebögen keiner Fallnummer zugeordnet werden konnten.

5.2 Inhaltliche Datenerfassung

Die Antworten aus den Papierfragebögen wurden maschinell erfasst. Die Erfassungsregeln waren zuvor mit dem BKA abgestimmt worden. Der Regelkatalog beinhaltete unter anderem Regeln für den Umgang mit Filterverstößen, die Markierung der Stellen, wo Angaben fehlten sowie die Dokumentation von zusätzlichen Angaben, die vom jeweils vorgegebenen Antwortformat abwichen. Darunter fielen beispielsweise Textangaben anstelle der geforderten Zahlenangaben.

Die elektronisch erfassten Befragungsdaten aus den Papierfragebögen wurden in einen Befragungsdatensatz (Rohdatenebene) überführt.

Für den Online-Fragebogen (CAWI) erfolgte die elektronische Datenerfassung bereits direkt bei der Eingabe der jeweiligen Antworten im Fragebogen durch die Zielperson. Für die anschließende inhaltliche Prüfung der Befragungsdaten wurden die Befragungsdaten aus dem Papier- und dem CAWI-Fragebogen in einen Datensatz zusammengeführt.

5.3 Datenprüfung und -bereinigung

5.3.1 Papierfragebogen

Nach der inhaltlichen Prüfung des Papierfragebogens schloss sich eine Reihe von Bereinigungsschritten an. Die Bereinigung von Ausfüllfehlern wurde anhand eines ausführlichen Regelkatalogs vorgenommen, der zuvor mit dem BKA abgestimmt und schriftlich fixiert wurde. Die Regeln wurden anhand eines Zwischendatensatzes entwickelt und mit dem Enddatensatz vervollständigt. Die vorgenommenen Bereinigungen bezogen sich in der Regel auf folgenden Fehlerarten:

1. Unzulässige Mehrfachnennung
2. Widersprüchliche/inkonsistente Angaben
3. Filterverstöße
4. Offene Angabe von Fallzahlen, Jahresangaben etc.

Neben den einzelnen Bereinigungsregeln wurde auch festgehalten, in welcher Fragen-Reihenfolge die Bereinigungen vorgenommen werden sollen.

Nach Abschluss der Datenbereinigung erhielt das BKA den Befragungsdatensatz mit allen ursprünglichen Angaben (Rohdaten). Sämtliche Variablen, die im Bereinigungsprozess verändert worden waren, wurden in den Datensatz integriert und mit „_e“ für „editiert“ gekennzeichnet.

Für die abschließende inhaltliche Prüfung beim BKA erhielt das BKA eine Auswahl an Bildarchiven der Papierfragebögen. Dabei handelte es sich, um:

- Bildarchive mit handschriftlichen Randnotizen oder Bemerkungen der Zielpersonen und
- Fragebögen des Hamburger Moduls mit unzulässiger Mehrfachangabe bei Frage 65.

Vor der Übergabe an das BKA wurden die Bildarchive durchgesehen und anonymisiert.

5.3.2 Online-Fragebogen

Ein Vorteil des elektronischen Fragebogens besteht in der Möglichkeit, direkt während der Beantwortung einer Frage die Eingabe prüfen zu können. Inkonsistente Angaben und Eingaben außerhalb des gültigen Wertebereichs können somit vermieden werden. Das gleiche gilt für unzulässige Doppelnennungen oder das unabsichtliche Auslassen einer Frage.

Die Datenprüfung vollzog sich bereits im programmierten Instrument. Die anschließende Datenprüfung beschränkte sich nach dem Feldende auf die Prüfung von Abbruchstellen im Fragebogen und die Ausfülldauern.

5.3.3 Umgang mit doppelt ausgefüllten Papier- und Online-Fragebögen

Aufgrund der zwei Erhebungsmethoden war es prinzipiell möglich, den Fragebogen sowohl auf Papier als auch online auszufüllen. Die Rückläufe wurden entsprechend auf mögliche Mehrfachteilnahme hin geprüft. Zum Feldende lagen für 126 Zielpersonen zwei ausgefüllte Fragebögen vor; bei diesen insgesamt 252 doppelt ausgefüllten Fragebögen handelt es sich jeweils um einen Papier- und einen CAWI-Fragebogen.

Mit dem BKA wurde ein Regelwerk für die Auswahl des Fragebogens abgestimmt, der im Enddatensatz verbleiben sollte. Das Regelwerk umfasst eine Hierarchisierung von Kriterien.²⁸

Im ersten Schritt wurden die Angaben zu Geschlecht und Alter mit den Stichprobeninformationen abgeglichen. In einer abgestuften Form wurde geprüft, welcher der beiden Fragebögen bei den beiden Angaben den höchsten Übereinstimmungsgrad mit den Stichprobeninformationen aufwies. Dieser Fragebogen wurde für die Auswertung im Datensatz belassen.

²⁸ Das Konzeptpapier mit allen konkreten Regeln liegt dem BKA vor.

In Konstellationen, die nach Durchlauf dieses ersten Prüfschritts nicht geklärt werden konnten, wurde im zweiten Schritt der Missing-Anteil als Auswahlkriterium herangezogen. Es wurde derjenige Fragebogen im Auswertungsdatensatz behalten, welcher den geringeren Missing-Anteil enthielt.²⁹

Für Fälle, die aufgrund des Vergleichs des Missing-Anteils noch nicht entschieden werden konnten, wurde das Datum des Fragebogeneingangs bei infas herangezogen. Für den Verbleib im Befragungsdatensatz wurde derjenige Fragebogen ausgewählt, welcher als erstes bei infas eingetroffen war.

5.4 Ausschluss nicht verwertbarer Fälle

Bevor die Gewichtung der Befragungsdaten vorgenommen wurde, verständigten sich infas und BKA über den Ausschluss von Fällen aus dem Befragungsdatensatz sowie die auswertbaren Fälle. Das BKA entwickelte auf Basis des aufbereiteten Gesamtbefragungsdatensatzes die folgenden Kriterien für den Ausschluss von nicht auswertbaren Fällen:

1. Ausfüllgrad allgemein: weitgehend unvollständige Fragebögen oder vollständig unbeantwortete Fragebögen
2. Ausfüllgrad speziell bei Einstellungsfragen und Viktimisierungserfahrungen
3. Ausfüllgrad Gewichtungsfragen
4. Auffällige Antwortmuster
5. Unrealistische Angaben
6. Sehr kurze Interviewdauer (nur CAWI)
7. Personen außerhalb der Zielgruppe: zu junge Personen
8. Zu große Abweichungen bei Meldedaten zu Geburtsdatum und Geschlecht im Vergleich zu den Angaben im Fragebogen

Für jedes dieser Kriterien wiederum wurde definiert, welche Bedingungen die Fragenbogeninhalte erfüllen mussten, um den Fragebogen insgesamt als auswertbar in die weiteren Datenanalysen einbeziehen zu können.

Aufgrund eines zu geringen Ausfüllgrads bei den Fragen zu Einstellungen und Viktimisierungserfahrungen wurden 211 Fragebögen aussortiert. Weitere 214 Fragebögen wurden als nicht auswertbar definiert, weil sie bei den ausgewählten gewichtungsrelevanten Fragen keinerlei gültige Antworten aufwiesen. Der mit Abstand größte Anteil an Fragebögen wurde ausgeschlossen, weil hier große Abweichungen zwischen den im Fragebogen genannten Altersangaben und denjenigen in den Meldeamtsdaten vorlagen. Als Kriterium für den Ausschluss

²⁹ In einem Einzelfall sind diese Regeln auch für einen doppelt vorliegenden Papierfragebogen umgesetzt worden. Hier lag für einen der Rückläufer kein Eingangsdatum vor, so dass die Regel, dass bei doppelten PAPI-Bögen der zuerst eingegangene PAPI-Bogen ausgewählt wird, nicht angewendet werden konnte.

eines Fragebogens wurde eine Abweichung zwischen den Angaben zum Alter aus dem Fragebogen und den Meldeamtsdaten von mehr als 2 Jahren festgelegt. Im Ergebnis wurden nach diesem Prüfschritt insgesamt 965 Fragebögen ausgeschlossen. In Summe sind 1.393 Fragebögen der Basisstichprobe durch die beschriebene Prüfung des BKA von der Auswertung ausgeschlossen worden.

Des Weiteren hatte das BKA bereits vor Beginn der Datenerfassung 2 Fälle als nicht auswertbar definiert.

Von den Fragebögen aus dem Methodenexperiment wurden 16 Fälle durch die vorab beschriebenen Prüfkriterien des BKA als nicht auswertbar definiert.

In Summe wurden aufgrund der vom BKA aufgestellten PrüfregeIn 1.411 Fragebögen als nicht auswertbar deklariert. Zusätzlich mussten weitere 67 Papierfragebögen von der Auswertung ausgeschlossen werden, da der QR-Code entfernt oder unkenntlich gemacht worden war und diese Fragebögen keiner Fallnummer und damit keiner Person aus der Stichprobe zugeordnet werden konnten.

Für die gesamte Studie ergibt sich damit eine Anzahl von 1.478 nicht auswertbaren Fällen.

5.5 Fragebogendokumentation

Das BKA erhielt für beide Fragebogenvarianten – Papier- und Online-Fassung – eine getrennte Dokumentation der Fragebögen.

Als Basis für die Fragebogendokumentation des Papierfragebogens wurde die PDF-Druckvorlage genutzt. In das Dokument wurden die Variablennamen und die Codes für jede Antwortmöglichkeit eingetragen. Ergänzend wurde eine Seite erstellt, auf der die Regeln für die Vergabe von Missing-Codes notiert waren. Das BKA erhielt drei PDF-Dokumente, jeweils eines für den Kernfragebogen, für den Kernfragebogen einschließlich NRW-Modul und den Kernfragebogen einschließlich HH-Modul.

Der CAWI-Fragebogen wurde anhand von Screenshots dokumentiert. Zusätzlich wurden Filterführung, die einzelnen Ausfüllhinweise, Wertebereiche und allgemeine technische Hinweise in einem separaten EXCEL-Dokument dokumentiert und an das BKA übergeben. Zusätzlich wurden in diesem Dokument die Prüfungen hinsichtlich Vollständigkeit der gegebenen Antworten oder die Vorgehensweise beim Abgleich von zwei Angaben zum selben Sachverhalt festgehalten.

6 Gewichtung

Grundlage der Berechnung der Gewichtungs- und Hochrechnungsfaktoren ist die Untergliederung des Erhebungsprozesses in die Phase der Designauswahl (Auswahlwahrscheinlichkeit) und die Phase der an die Auswahl anschließenden Antwortgewährung (Adjustierung der Designgewichte). Für beide Phasen kann das Gewicht jeweils über die Berechnung der inversen Auswahlwahrscheinlichkeit (sog. Horvitz-Thompson-Schätzer, vgl. Horvitz, Thompson 1952: 663–685) erfolgen. Das Gesamtgewicht wird als Inverse des Produkts der beiden Wahrscheinlichkeiten bestimmt. Alle berechneten Gewichtungsfaktoren wurden auch als fallzahlnormierte Gewichte ausgeliefert.

Dem mehrstufigen Stichprobendesign entsprechend, werden für die Gewichtungs- und Hochrechnungsfaktoren die folgenden Schritte definiert.

1. Designgewichtung (Auswahl der Gemeinden und anschließende Auswahl der Personen)
2. Ausfallgewichtung (Nonresponse-Adjustierung)
3. Kalibrierung (Randanpassung) an Verteilungen der offiziellen Bevölkerungsstatistik (Mikrozensus, fortgeschriebener Bevölkerungsstand etc.)

6.1 Designgewichtung

Das Stichprobendesign unterscheidet sich zum Teil zwischen den Bundesländern. So wurden in einigen Bundesländern größere („aufgestockte“) Stichproben gezogen als ihnen proportional, d.h. bezogen auf ihren Anteil an der Gesamtbevölkerung, entsprächen (Schleswig-Holstein, Hamburg, Nordrhein-Westfalen, Thüringen). Außerdem haben einige Bundesländer über ein zentrales Register Adressen gezogen (Berlin, Schleswig-Holstein und Thüringen); dementsprechend liegt in diesen Bundesländern keine geklumpte Pointstichprobe vor.

Für den Großteil der Befragung wurde eine Einwohnermeldeamtsstichprobe gezogen, bei der es sich im Grundsatz um eine zweistufige, geklumpte Stichprobe handelt.

- 1. Stufe (Primary Sample Units = PSU): Auswahl von Gemeinden bzw. Sample Points, wobei eine Gemeinde mehrere Sample Points umfassen konnte
- 2. Stufe (Secondary Sample Units = SSU): Auswahl der zu befragenden Zielpersonen

Stufe 1 Primary Sampling Unit (PSU): Gemeinden

Zunächst wurden die Gemeinden gezogen, weshalb für diese die Auswahlwahrscheinlichkeit berechnet werden musste. Gegeben waren m Sampling Points (PSU), ausgewählt mittels geschichteter Zufallsziehung (siehe 2.3) aus M Gemeinden in Deutschland ohne die oben genannten Bundesländer, die zentral gezogen haben. Die Anzahl der Sampling Points in einer Gemeinde s kann dabei größer sein als 1. Weiterhin wird für die Berechnung die Anzahl an Einwohnerinnen und Einwohnern aus der Grundgesamtheit N_i in der Gemeinde i benötigt. Somit ergibt sich:

$$\sum_{i=1}^M N_i = N$$

Die Auswahlwahrscheinlichkeit in Stufe 1 für einen Sampling Point in Gemeinde i ergibt sich dann durch:

$$P_i^{Stufe\ 1} = \frac{m \cdot N_i}{N \cdot s}$$

Stufe 2 Secondary Sampling Unit (SSU): Personen in ausgewählten Gemeinden

In einem zweiten Schritt wurden die Personen in den ausgewählten Gemeinden gezogen. Die Auswahlwahrscheinlichkeit in dieser Stufe 2 wird beschrieben durch:

$$P_{j|i}^{Stufe\ 2} = \frac{n}{m \cdot N_i}, [n = \text{ausgewählte Personen in der Bruttostichprobe}]$$

Die Anzahl der gezogenen Adressen pro Gemeinde und Sampling Point kann theoretisch vom geplanten Wert abweichen, wenn beispielsweise nicht genügend Adressen von der Gemeinde geliefert wurden. Für die korrekte Berechnung der Auswahlwahrscheinlichkeit wird dann die variierende Anzahl gezogener Adressen pro Sample Point (n_i) in der jeweiligen Gemeinde i verwendet. Die korrekte Auswahlwahrscheinlichkeit in Stufe 2 ergibt sich dann für Gemeinden mit einem Sample Point durch:

$$P_{j|i}^{Stufe\ 2} = \frac{n_i}{N_i}$$

Da einige Gemeinden in mehrere Sample Points aufgeteilt waren, errechnet sich die Inklusionswahrscheinlichkeit für die Personen durch:

$$P_{j|i}^{Stufe\ 2} = \frac{n_i \cdot s}{N_i}$$

Für die Bildung des Designgewichts für diesen Teil der Stichprobe wurde das inverse Produkt aus den Wahrscheinlichkeiten der Stufen 1 und 2 wie folgt berechnet:

$$dw = \frac{1}{P_i^{Stufe1} \cdot P_{j|i}^{Stufe2}}$$

Wobei P_i^{Stufe1} die Auswahlwahrscheinlichkeit für einen Sampling Point in Gemeinde i und $P_{j|i}^{Stufe2}$ die Auswahlwahrscheinlichkeit der Person j in der jeweiligen Gemeinde bezeichnet.

Für die Bundesländer mit zentralem Register ergibt sich durch die einstufige Zufallsauswahl („(stratified) simple random sampling“) ein fallzahlnormierter Gewichtungsfaktor von 1 für alle Einheiten.

Für diese Bundesländer beschränkt sich die Designgewichtung auf den Ausgleich des nach Bundesländern disproportionalen Stichprobenansatzes:

$$dw = \frac{N_{BL}}{n_{BL}}$$

Wobei N_{BL} die Größe der Grundgesamtheit und n_{BL} die Größe der Bruttostichprobe im jeweiligen Bundesland bezeichnet.

6.2 Ausfallgewichtung (Nonresponse-Adjustierung)

Bei der Nonresponse-Adjustierung der Designgewichte mithilfe eines logistischen Regressionsmodells (Ausfallmodell) wird die Teilnahmewahrscheinlichkeit mit der abhängigen Variable „Teilnahme am Interview“ (Teilnahme: $y = 1$, Nichtteilnahme: $y = 0$) modelliert. Die erklärenden Variablen stammen aus den mitgelieferten Merkmalen der Einwohnermeldeämter bzw. zusätzlichen bekannten regionalen Merkmalen, also dem Alter, dem Geschlecht, der Staatsangehörigkeit, dem Bundesland und der Gemeindegrößenklasse.

Die aus dem Ausfallmodell abgeleitete Teilnahmewahrscheinlichkeit fließt in die Berechnung der adjustierten Designgewichte ein, indem die Gesamtauswahlwahrscheinlichkeit auf Basis der Stufen 1 und 2 mit der Teilnahmewahrscheinlichkeit ($P_{y=1}$) aus dem Ausfallmodell multipliziert wird.

$$dw_{adj} = \frac{1}{P_i^{Stufe1} \cdot P_{j|i}^{Stufe2} \cdot P_{y=1}}$$

Tabelle 21 Ergebnis Ausfallmodell

Abhängige Variable: Y = 1: Teilnahme, Y = 0: Keine Teilnahme	Referenzkategorie	Odds Ratio	P-Wert
Altersgruppe	16-24 Jahre		
20-24 Jahre		1,00	0,93
25-29 Jahre		1,06	0,18
30-34 Jahre		1,23	0,00
35-39 Jahre		1,29	0,00
40-44 Jahre		1,28	0,00
45-49 Jahre		1,46	0,00
50-54 Jahre		1,72	0,00
55-59 Jahre		1,95	0,00
60-64 Jahre		2,19	0,00
65-74 Jahre		2,49	0,00
75 Jahre u. älter		1,51	0,00
Geschlecht	Weiblich		
Männlich		1,06	0,00
Nationalität	Deutsch		
Nicht deutsch		0,34	0,00
Keine Angabe		0,92	0,28
Bundesland	Schleswig-Holstein		
Hamburg		0,98	0,44
Niedersachsen		0,86	0,00
Bremen		1,03	0,82
Nordrhein-Westfalen		0,92	0,00
Hessen		0,87	0,01
Rheinland-Pfalz		0,83	0,00
Baden-Württemberg		1,15	0,00
Bayern		1,05	0,15
Saarland		0,74	0,02
Berlin		1,04	0,50
Brandenburg		0,78	0,00
Mecklenburg-Vorpommern		0,81	0,02
Sachsen		0,84	0,00
Sachsen-Anhalt		0,68	0,00
Thüringen		0,85	0,00

Abhängige Variable: Y = 1: Teilnahme, Y = 0: Keine Teilnahme	Referenzkategorie	Odds Ratio	P-Wert
GKBIK10	Unter 2.000 Einw.		
2.000 bis unter 5.000 Einw.		0,92	0,16
5.000 bis unter 20.000 Einw.		0,98	0,71
20.000 bis unter 50.000 Einw.		1,01	0,84
50.000 bis unter 100.000 Einw. + SGTyp 2/3/4		1,06	0,26
50.000 bis unter 100.000 Einw. + SGTyp 1		0,98	0,70
100.000 bis unter 500.000 Einw. + SGTyp 2/3/4		1,12	0,03
100.000 bis unter 500.000 Einw. + SGTyp 1		1,08	0,10
500.000 Einw. und mehr + SGTyp 2/3/4.		1,18	0,00
500.000 Einw. und mehr + SGTyp 1		1,14	0,01
Pseudo R2 (McFadden)			0,04
Fallzahl			122.667

P-Werte basieren auf einfachen Standardfehlern.
Quelle: infas, eigene Berechnung

6.3 Kalibrierung (Randanpassung)

Die ausfalladjustierte Designgewichtung wird im Anschluss weiter adjustiert, um Undercoverage, Nonresponse und weitere systematische Ausfallprozesse auszugleichen. Da die Varianz der ausfalladjustierten Gewichte weiterhin zum größten Teil auf den disproportionalen Stichprobenansatz nach Bundesländern zurückgeht, wurde auf eine Trimmung der Gewichte vor der Kalibrierung verzichtet.

Zum Ausgleich von systematischen Ausfällen, die auch als unterschiedliche Teilnahme- bzw. als Realisierungswahrscheinlichkeiten beschrieben werden können, erfolgt eine Anpassung der Verteilungen interessierender Merkmale in der Stichprobe an bekannte Verteilungen in der Grundgesamtheit über den Generalized Regression Estimator (GREG). Bei dieser Kalibrierung werden die adjustierten Designgewichte je Untersuchungseinheit mit einem Faktor multipliziert, so dass die bekannten Verteilungen in der Grundgesamtheit durch die Stichprobe exakt abgebildet werden. Diese Faktoren wurden dabei auf den Bereich zwischen 0,05 und 8 begrenzt.

Die multiplikative Verknüpfung von adjustiertem Designgewicht und Anpassungsgewicht liefert die kalibrierten Endgewichte. Diese werden ebenso wie die Auswahlwahrscheinlichkeiten und die Designgewichte als reine Gewichtungsfaktoren auf die realisierten Fallzahlen normiert und als Hochrechnungsfaktoren auf die Grundgesamtheit dem Datensatz zugespielt.

Für die Randanpassung ist es notwendig, Randverteilungen der Population verfügbar zu haben. Hier wurde auf die aktuellsten verfügbaren Daten der amtlichen Statistik zurückgegriffen. Als zentrale Merkmale wurden verwendet:

- Geschlecht x Altersgruppen,
- höchster allgemeinbildender Schulabschluss,
- Erwerbsstatus,

- Haushaltsgröße,
- BIK-Gemeindegrößenklasse,
- Migrationshintergrund.

Aufgrund der Disproportionalität auf Bundeslandebene erfolgte die Anpassung jeweils auf Bundeslandebene. Dies hat weiterhin den Vorteil, dass sich Gewichte für jedes Bundesland ohne zusätzliche Anpassungsschritte einzeln erstellen lassen. Lediglich die Fallzahlnormierung muss gesondert durchgeführt werden. Für Schleswig-Holstein wurde Geschlecht x Altersgruppe zusätzlich auf Kreisebene angepasst.

Für die Kalibrierung mittels GREG ist es notwendig, dass jedes Merkmal für jeden Befragten gefüllt ist. Daher wurde bei Item-Nonresponse imputiert über die Stichprobe (Geschlecht, Alter, Bundesland, BIK-Gemeindegrößenklasse) bzw. den Median (Schulabschluss, Haushaltsgröße) oder den Modus (Erwerbsstatus). Wenn der Migrationshintergrund nicht bestimmt werden konnte, wurde nicht imputiert. Stattdessen wurde der Anteil der Missings bei diesem Merkmal in der Gewichtung konstant gehalten.

Da auf Bundeslandebene für die Kalibrierung die BIK-Größenklassen und Migrantengruppen teilweise stark zusammengefasst werden mussten, wurden auf Bundesebene zusätzlich die BIK-Größenklassen und der Migrationshintergrund ohne Zusammenfassen von Kategorien angepasst. So ist gewährleistet, dass die Verteilungen dieser Variablen auch innerhalb der zusammengefassten Kategorien auf Bundesebene passen.

In der folgenden Tabelle sind die Ergebnisse der Gewichtung zusammengefasst. Dargestellt sind die Verteilungen der aktiven Gewichtungsmerkmale gewichtet einerseits mit dem reinen Designgewicht, andererseits mit dem kalibrierten Gewichtungsfaktor. Dabei entsprechen die gewichteten Verteilungen dem fortgeschriebenen Bevölkerungsstand des Statistischen Bundesamts ab 16 Jahren (Stand 31.12.2018) und dem Mikrozensus (Stand 31.12.2019).

Tabelle 22 Verteilungsvergleich gewichtete Stichprobe und Grundgesamtheit

	Designgewicht	Randangepasst	Sollverteilung (Grundgesamtheit)	
	%	%	abs.	%
Gesamt	100,0	100,0	69.893.000	100,0
Bundesland				
Schleswig-Holstein	3,9	3,5	2.445.000	3,5
Hamburg	2,3	2,2	1.552.000	2,2
Niedersachsen	9,9	9,6	6.688.000	9,6
Bremen	1,0	0,8	577.000	0,8
Nordrhein-Westfalen	21,1	21,5	15.041.000	21,5
Hessen	6,5	7,6	5.274.000	7,5
Rheinland-Pfalz	4,7	4,9	3.439.000	4,9
Baden-Württemberg	14,2	13,3	9.326.000	13,3
Bayern	16,2	15,8	11.023.000	15,8
Saarland	1,0	1,2	847.000	1,2
Berlin	4,6	4,4	3.068.000	4,4
Brandenburg	3,2	3,0	2.123.000	3,0
Mecklenburg-Vorpommern	2,0	2,0	1.364.000	2,0
Sachsen	4,4	4,9	3.435.000	4,9
Sachsen-Anhalt	2,4	2,7	1.874.000	2,7
Thüringen	2,7	2,6	1.817.000	2,6
Geschlecht x Altersgruppen				
16-24 Jahre männlich	3,6	5,8	4.039.102	5,8
16-24 Jahre weiblich	4,6	5,3	3.686.197	5,3
25-34 Jahre männlich	5,5	7,7	5.394.894	7,7
25-34 Jahre weiblich	6,5	7,2	5.045.028	7,2
35-44 Jahre männlich	5,7	7,2	5.018.083	7,2
35-44 Jahre weiblich	6,2	7,0	4.914.484	7,0
45-54 Jahre männlich	8,1	8,9	6.190.846	8,9
45-54 Jahre weiblich	8,9	8,7	6.086.490	8,7
55-64 Jahre männlich	10,5	8,5	5.903.365	8,4
55-64 Jahre weiblich	10,9	8,6	6.007.830	8,6
65-74 Jahre männlich	8,6	5,6	3.913.077	5,6
65-74 Jahre weiblich	8,2	6,3	4.365.693	6,2
75 Jahre u. älter männlich	6,4	5,4	3.789.111	5,4
75 Jahre u. älter weiblich	6,4	7,9	5.538.800	7,9

	Designgewicht	Randangepasst	Sollverteilung (Grundgesamtheit)	
	%	%	abs.	%
Höchster allgemeinbildender Schulabschluss				
Ohne (allg. bild.) Schulabschluss	2,3	6,5	4.596.000	6,6
Haupt-/Sonder-/Förderschule	20,1	30,3	21.107.100	30,2
Mittlere Reife	34,1	29,4	20.533.500	29,4
Fachhochschulreife	12,5	8,2	5.749.500	8,2
Hochschulreife	31,1	25,6	17.906.900	25,6
Erwerbsstatus				
Voll Erwerbstätig	40,9	40,3	28.199.000	40,3
Erwerbslos	1,9	1,8	1.238.000	1,8
Andere	57,2	57,9	40.456.000	57,9
Haushaltsgröße				
Einpersonenhaushalt	21,5	24,4	17.065.800	24,4
Haushalte mit 2 Personen	45,4	38,2	26.686.300	38,2
Haushalte mit 3 Personen	15,1	17,0	11.869.600	17,0
Haushalte mit 4 Personen	12,7	14,2	9.927.600	14,2
Haushalte mit 5 oder mehr Personen	5,3	6,2	4.343.700	6,2
BIK-Gemeindegrößenklasse				
Unter 2.000 Einw.	1,6	1,7	1.183.560	1,7
2.000 bis unter 5.000 Einw.	2,4	2,5	1.710.657	2,4
5.000 bis unter 20.000 Einw.	8,5	8,7	6.084.481	8,7
20.000 bis unter 50.000 Einw.	10,2	10,7	7.444.577	10,7
50.000 bis unter 100.000 Einw. + SGTYP 2/3/4	8,0	7,8	5.425.076	7,8
50.000 bis unter 100.000 Einw. + SGTYP 1	1,7	2,2	1.507.770	2,2
100.000 bis unter 500.000 Einw. + SGTYP 2/3/4	16,1	14,9	10.434.012	14,9
100.000 bis unter 500.000 Einw. + SGTYP 1	13,9	15,0	10.497.554	15,0
500.000 Einw. und mehr + SGTYP 2/3/4.	9,9	9,7	6.773.771	9,7
500.000 Einw. und mehr + SGTYP 1	27,8	26,9	18.831.542	26,9
Migrationshintergrund				
Ohne Migrationshintergrund	77,2	72,2	50.469.114	72,2
Herkunftsland: Polen	2,0	2,6	1.783.458	2,6
Herkunftsland: Türkei	1,0	3,1	2.194.963	3,1
Herkunftsland: Russland, ehem. Sowjetunion	2,1	1,6	1.100.452	1,6
Herkunftsland: Krisenländer (Syrien, Afghanistan, Irak, Eritrea)	0,7	1,4	954.369	1,4
Herkunftsland: sonstige Herkunftsländer	7,5	13,3	9.292.858	13,3
Herkunftsland: unbestimmt	1,6	0,3	206.922	0,3
Herkunftsland: keine Angabe	2,5	0,3	178.445	0,3
Unbekannt ob MHG	5,4	5,3	3.712.419	5,3

Quellen: Statistisches Bundesamt (Destatis); Ergebnisse des Mikrozensus 2019.
Statistisches Bundesamt (Destatis); Fortschreibung des Bevölkerungsstandes 2018.
BIK ASCHPURWIS + BEHRENS GMBH; BIK Basisdaten/Gemeindedatei.
Eigene Berechnungen.

Die in der vorherigen Tabelle ausgewiesenen Verteilungen zeigen, dass die lediglich designgewichtete Stichprobe bereits bei fast allen Merkmalen recht nah an der realen Verteilung in der Grundgesamtheit liegt. Vergleichsweise hohe Abweichungen finden sich lediglich beim höchsten Bildungsabschluss. Wie in anderen sozialwissenschaftlichen Studien zeigt sich auch hier eine erkennbare Bildungsverzerrung durch systematischen Nonresponse, der durch die Ausfalladjustierung und Kalibrierung der Designgewichtung ausgeglichen wird.

6.4 Kennwerte der Gewichte

Um die Güte der Gewichtung und deren Einfluss auf die Konfidenzintervalle von Populationsschätzungen auf Basis der gewichteten bzw. hochgerechneten Daten beurteilen zu können, empfiehlt sich anschließend an die Gewichtung die Berechnung von Effektivitätsparametern (Designeffekt bzw. Effektivitätsmaß und effektive Fallzahl, vgl. Schnell, Noack 2015: 8-75; Gabler, Hoffmeyer-Zlotnik, Krebs 1994). Diese Effektivitätsparameter beziffern über die Varianz der Gewichte jeweils den Stichprobenfehler im Verhältnis zum Stichprobenfehler bei einer uneingeschränkten Zufallsauswahl (srs = simple random sampling) ohne Gewichtung. Darüber lässt sich angeben, welche Auswirkungen eine Fallzahlreduktion im ungünstigsten Fall (keine Korrelation eines Merkmals mit den Gewichten) auf das Vertrauensintervall der Populationsschätzungen hat; das heißt, wie sich die Schwankungsbreiten durch die Gewichtung verändern.

Zur Beurteilung der Gewichte wird im Folgenden neben dem Mittelwert und der Standardabweichung auch das Effektivitätsmaß (E) ausgewiesen. Dieses Maß basiert auf der Varianz des Gewichtungsfaktors. Es gibt in Prozent der realisierten Fallzahl an, wie groß bei Verwendung des Gewichts die effektive Fallzahl (n') bei einem passiven Merkmal ist, das mit den aktiven Merkmalen nicht korreliert (Worst-Case-Szenario). Die effektive Fallzahl entspricht der Anzahl Befragter, die bei einer uneingeschränkten Zufallsauswahl – gegeben die Varianz des Merkmals in der Stichprobe – den gleichen Stichprobenfehler produziert hätte.

Das Effektivitätsmaß drückt das Verhältnis von n zu n' als Prozentwert aus und wird berechnet über:

$$E = \frac{n'}{n} \text{ wobei: } n' = \frac{(\sum_i g_i)^2}{\sum_i g_i^2}$$

Tabelle 23 Kennwerte der Gewichte

Variable	Bezeichnung	Fallzahl	Mittelwert	Standardabweichung	Min.	Max.	Effektivitätsmaß (%)
dw_br	Fallzahlnormiertes Designgewicht Bruttostichprobe	122.667	1,0	1,48	0,1	5,0	31,4
dw_ne	Fallzahlnormiertes Designgewicht Nettostichprobe	45.351	1,0	1,49	0,1	5,0	31,1
hoch_kal	Kalibrierter Hochrechnungsfaktor	45.351	1.541,2	3.240,77	6,2	107.220,8	18,4
gew_kal	Kalibriertes fallzahlnormiertes Gewicht	45.351	1,0	2,10	0,0	69,6	
dw_ne_hh	Fallzahlnormiertes Designgewicht Nettostichprobe Hamburg	8.534	1,0	0,39	1,0	35,9	87,0
gew_kal_hh	Kalibriertes fallzahlnormiertes Gewicht Hamburg	8.534	1,0	1,05	0,0	68,5	47,6
hoch_kal_hh	Kalibrierter Hochrechnungsfaktor Hamburg	8.534	181,9	190,65	6,2	12.453,1	
dw_ne_nw	Fallzahlnormiertes Designgewicht Nettostichprobe Nordrhein-Westf.	13.475	1,0	0,10	0,2	6,1	99,0
gew_kal_nw	Kalibriertes fallzahlnormiertes Gewicht Nordrhein-Westfalen	13.475	1,0	0,79	0,0	10,6	61,5
hoch_kal_nw	Kalibrierter Hochrechnungsfaktor Nordrhein-Westfalen	13.475	1.116,2	883,46	37,5	11.845,1	
dw_ne_sh	Fallzahlnormiertes Designgewicht Nettostichprobe Schleswig-Holstein	9.229	1,0	0,32	0,6	22,	90,8
gew_kal_sh	Kalibriertes fallzahlnormiertes Gewicht Schleswig-Holstein	9.229	1,0	0,71	0,0	18,3	66,4
hoch_kal_sh	Kalibrierter Hochrechnungsfaktor Schleswig-Holstein	9.229	264,9	188,35	9,5	4.844,4	
dw_ne_th	Fallzahlnormiertes Designgewicht Nettostichprobe Thüringen	6.734	1,0	0,28	1,0	23,9	92,7
gew_kal_th	Kalibriertes fallzahlnormiertes Gewicht Thüringen	6.734	1,0	0,77	0,0	35,4	62,5
hoch_kal_th	Kalibrierter Hochrechnungsfaktor Thüringen	6.734	269,8	209,11	10,2	9.542,6	

Quelle: infas, eigene Berechnung

Die mehrfache Integration disproportionaler Stichproben erhöht zwar die Fallzahlen in einigen Bundesländern, führt aber auch zu einem Absinken der Effektivität und damit der effektiven Fallzahl in der Gesamtstichprobe – das heißt, einem Ansteigen des Stichprobenfehlers bei gewichteten Populationsschätzungen. Dies zeigt sich im Effektivitätsmaß des reinen Designgewichts von ca. 31 Prozent. Angesichts der insgesamt sehr hohen Fallzahl kann dies hingenommen werden. Das ausgewiesene Effektivitätsmaß von ca. 18 Prozent kann vor diesem Hintergrund als gut eingeschätzt werden. Die Effektivitätsmaße der kalibrierten Gewichte in den aufgestockten Bundesländern zwischen 48 und 66 Prozent zeigen weiterhin, dass die niedrige Effektivität in der Gesamtstichprobe im Wesentlichen auf Effekte des disproportionalen Stichprobendesigns zurückzuführen ist.

BKA – Sicherheit und Kriminalität 2020

Trimmung der Eingangsgewichte

Für die ursprüngliche Gewichtung wurde auf eine Trimmung der ausfalladjustierten Designgewichte vor der Kalibrierung verzichtet, da deren Varianz zum größten Teil auf den disproportionalen Stichprobenansatz nach Bundesländern zurückgeht und eine globale Trimmung somit den wichtigen Ausgleich des disproportionalen Stichprobendesigns verhindern würde. Bei der gesonderten Betrachtung einzelner Bundesländer kann es allerdings bedingt durch die nach Bundesländern disproportionaler Ziehung und Zielpersonen die nach der Ziehung in den Einwohnermeldeämtern aus einem aufgestockten Bundesland in ein nicht aufgestocktes Bundesland oder umgekehrt gezogen sind zu starken Ausreißern kommen. Diese Werte sind im Sinne der Verzerrungsreduktion inhaltlich korrekt, allerdings ist hier im Sinne der in der Statistik immer wieder anzutreffenden Abwägung zwischen Verzerrung und Varianz eine Trimmung der adjustierten Designgewichte auf Bundeslandebene vertretbar und im Sinne der Varianzreduktion für die Analyse einzelner Bundesländer wünschenswert. Daher wurden für die neu kalibrierten Gewichte (Dezember 2021) die ausfalladjustierten Designgewichte auf Bundeslandebene am 1%- und am 99%-Quantil getrimmt bevor sie kalibriert wurden.

Tabelle 1 zeigt die Kennwerte der Gewichte inklusive eines Effektivitätsvergleichs zwischen ursprünglichen und neuen Gewichten. Die Trimmung zeigt kaum einen Effekt auf die Effektivität der Gewichte für die Gesamtstichprobe. Auf Ebene einzelner Bundesländer zeigen sich dafür zum Teil deutliche Verbesserungen der Effektivität, vor allem in Hamburg und Thüringen.

Tabelle 1 Kennwerte der neu kalibrierten Gewichte

Variable	Bezeichnung	Fallzahl	Mittelwert	Standardabweichung	Min.	Max.	Effektivitätsmaß (%)	altes Effektivitätsmaß (%)
hoch_kal	Kalibrierter Hochrechnungsfaktor	45.351	1541,2	3223,89	6,2	107069,1	18,6	18,4
gew_kal	Kalibriertes fallzahlnormiertes Gewicht	45.351	1,0	2,09	0,0	69,5		
gew_kal_hh	Kalibriertes fallzahlnormiertes Gewicht Hamburg	8.534	1,0	0,75	0,0	8,1	63,9	47,6
hoch_kal_hh	Kalibrierter Hochrechnungsfaktor Hamburg	8.534	181,9	136,62	6,2	1466,3		
gew_kal_nw	Kalibriertes fallzahlnormiertes Gewicht Nordrhein-Westfalen	13.475	1,0	0,78	0,0	8,2	62,2	61,5
hoch_kal_nw	Kalibrierter Hochrechnungsfaktor Nordrhein-Westfalen	13.475	1116,2	870,83	38,1	9153,9		
gew_kal_sh	Kalibriertes fallzahlnormiertes Gewicht Schleswig-Holstein	9.229	1,0	0,66	0,0	9,9	69,5	66,4
hoch_kal_sh	Kalibrierter Hochrechnungsfaktor Schleswig-Holstein	9.229	264,9	175,60	9,8	2635,1		
gew_kal_th	Kalibriertes fallzahlnormiertes Gewicht Thüringen	6.734	1,0	0,65	0,0	22,7	70,1	62,5
hoch_kal_th	Kalibrierter Hochrechnungsfaktor Thüringen	6.734	269,8	176,22	10,2	6120,4		

Literaturverzeichnis

Cox, L. H. (1987): A Constructive Procedure for Unbiased Controlled Rounding. In: Journal of the American Statistical Association, Vol. 82, No. 398, S. 520–524

Dorer, B., Krajevca, E. (2018): ESS Round 9. Verification Instructions for National Coordinators. ESS

Gabler, S., Hoffmeyer-Zlotnik, J., Krebs, D. (1994): Gewichtung in der Umfragepraxis. Opladen: Westdeutscher Verlag

Horvitz, D. G., Thompson, D. J. (1952): A generalization of sampling without replacement from a finite universe. In: Journal of the American Statistical Association, 47, S. 663–685

Schnell, R., Noack, M. (2015): Stichproben, Nonresponse und Gewichtung für Viktimisierungsstudien. In: Guzy, N., Birkel, C., Mischkowitz, R. (Hg.): Viktimisierungsbefragungen in Deutschland, Band 2 Methodik und Methodologie, Wiesbaden: Bundeskriminalamt, S. 8–75

Zavala-Rojas, D. (2014): A procedure to prevent differences in translated survey items using SQP. RECSM Working Paper Number 38