

Methodendokumentation

Weizenbaum Report: Politische Partizipation in Deutschland 2019

Die [Studie](#) wurde von der Forschungsgruppe (FG) 13 „Digital Citizenship“ am Weizenbaum-Institut für die vernetzte Gesellschaft in Berlin durchgeführt. Finanziert wurde sie im Rahmen der Projektförderung des Weizenbaum-Instituts durch Mittel des Bundesministeriums für Bildung und Forschung. Sie ist Teil eines längerfristigen Forschungsprojekts und wird weitere Datenerhebungen in den kommenden Jahren umfassen.

Datenerhebungsmethode

Die Daten wurden mit Hilfe telefonischer Befragungen im CATI-Modus („Computer Administered Telephone Interviewing“) unter Einsatz des Systems VOXCO CATI erhoben. Die Erhebung wurde im Telefonstudio der Freien Universität Berlin, Campus Lankwitz, von insgesamt 45 Interviewer*innen durchgeführt, die auf Honorarbasis arbeiteten. Geleitet wurde die Datenerhebung von Mitarbeiter*innen der Freien Universität Berlin (Gesamtleitung: [Prof. Dr. Martin Emmer](#)) und der FG 13 des Weizenbaum-Instituts (Leitung: [Dr. Pablo Portencheé](#), [Christian Strippel](#), [Nadja Schaetz](#) und [Laura Leißner](#)).

Umfang und Zeitraum der Befragung

Die Befragung fand im Herbst 2019 vom 09.09. bis 30.11. statt. In die Auswertung gingen die Antworten von insgesamt 1.298 Personen ein, die nach einem wissenschaftlichen Zufallsverfahren ausgewählt wurden. Die Interviews dauerten im Durchschnitt ca. 23 Minuten.

Grundgesamtheit und Stichprobenverfahren

Die sogenannte „Grundgesamtheit“ der Befragung umfasst die deutschsprachige Wohnbevölkerung in Deutschland ab dem Alter von 16 Jahren. Diese Definition soll sicherstellen, dass die Befunde am Ende auf alle Menschen verallgemeinert werden können, die dauerhaft in Deutschland leben – unabhängig von Staatsangehörigkeit, Herkunft etc. – und die damit auch Teil der deutschen Bürgerschaft sind und am gesellschaftlichen Leben teilhaben. Die Beschränkung auf den deutschsprachigen Teil der Bevölkerung schränkt die Aussagekraft etwas ein, weil z. B. erst kürzlich zugewanderte Personen mit noch geringen Sprachkenntnissen nicht befragt werden. Dies ist ein Kompromiss, der im unverhältnismäßigen Aufwand begründet ist, den eine Übersetzung des Fragebogens in mehrere Sprachen und die Rekrutierung und Ausbildung fremdsprachiger Interviewer*innen bedeuten würde.

Um diese Grundgesamtheit in einer Stichprobe repräsentativ abbilden zu können, wurde ein sogenannter dualer Auswahlrahmen aus Festnetz- und Mobilfunknummern festgelegt („Dual-Frame-Ansatz“; Callegaro et al., 2011; Sand, 2014). Angesichts des anhaltenden Trends hin zur mobilen Telefonie und der steigenden Anzahl von Personen in Deutschland, die vorrangig

oder nur auf dem Mobiltelefon zu erreichen sind (sogenannte „Mobile-Onlys“, siehe Schneiderat & Schlinzig, 2011; Busse & Fuchs, 2012), wurde das Verhältnis von Festnetz- und Mobilfunknummern, die angerufen werden sollten, auf 50:50 festgelegt. Diese Entscheidung geht bewusst über die übliche Empfehlung eines Mobilfunkanteils in Stichproben von 30 bis 40 Prozent (zuletzt Häder & Sand, 2019, S. 69-73) hinaus, da sich diese Empfehlung auf eine Simulationsstudie des *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute* (ADM, 2012) stützt, die zum Zeitpunkt der Befragung bereits acht Jahre zurücklag und damit den fortschreitenden Anstieg der Mobilfunknutzung seither (Europäische Kommission, 2018; Bundesnetzagentur, 2019) nicht berücksichtigt.

Die Bereitstellung der Bruttostichprobe aus Telefonnummern, die angerufen werden sollten, erfolgte durch das *GESIS – Leibniz-Institut für Sozialwissenschaften* nach dem von Gabler und Häder (1997) entwickelten Stichprobendesign („Gabler-Häder-Design“), nach dem die Festnetz- und Mobilfunknummern nach einem jeweils festgelegten Zufallsverfahren generiert werden (siehe auch Häder & Gabler, 1998; Häder, 2015). Hierbei ist wichtig, dass die Telefonnummern nicht aus einem Bestand an tatsächlich existierenden Nummern ausgewählt werden, sondern zufällig generiert werden, so dass zu Beginn der Befragung noch unbekannt ist, ob diese Nummern tatsächlich existieren oder nicht. Im Umgang mit den Nummern wurden alle gesetzlichen Bestimmungen zum [Datenschutz](#) eingehalten. Das Verfahren hat zur Folge, dass als Bruttostichprobe eine wesentlich größere Zahl an Nummern am Anfang der Befragung steht, als am Ende tatsächlich Interviews geführt werden.

Für die Studie wurden insgesamt 101.822 Festnetz- und 61.980 Mobilfunknummern zufällig erstellt. Diese Telefonnummern wurden gemäß der „Richtlinie für telefonische Befragungen“ der ADM (2016) mit einem Maximum von zehn Erstkontaktversuchen angerufen. Bei Anrufen über das Festnetz wurden die Zielpersonen dabei über die sogenannte „Last-Birthday-Methode“ bestimmt, da über Festnetznummern häufig mehrere Personen in einem gemeinsamen Haushalt erreichbar sind. Bei Mobilfunknummern wurde die antwortende Person ohne weiteres Auswahlverfahren direkt befragt.

Befragungsverlauf und Ausschöpfungsquote

Da die Teilnahme an der Befragung selbstverständlich freiwillig war, konnte nicht mit allen über das Zufallsverfahren angesprochenen Personen ein Interview geführt werden. Da die Repräsentativität des Datensatzes umso stärker ist, je besser die Stichprobe „ausgeschöpft“ wird, wurde während des Trainings der Interviewer*innen und während der Befragung unter Einhaltung forschungsethischer Regeln großer Wert auf die Gewinnung von Teilnehmer*innen gelegt.

Die folgende Tabelle dokumentiert das Verhältnis von erreichten Anschlüssen und geführten Interviews:

Tabelle 1

	Stichprobenrelevante Nummern ¹	Stichprobenrelevante Ausfälle ²	Realisierte Interviews (Nettostichprobe)	Ausschöpfungsquote
Festnetz	4179	3782	397	9,5 %
Mobil	8318	7417	901	10,8 %
Gesamt	12497	11199	1298	10,4 %

¹ Nummern, die den Anforderungen der Grundgesamtheit entsprachen: Anschlüsse, deren Nummern vergeben waren und hinter denen Privatanschlüsse standen, über die Menschen über 16 mit dauerhaftem Wohnsitz in Deutschland erreichbar waren. Die Differenz zur sehr viel größeren Zahl an Nummern in der Bruttostichprobe ergibt sich aus nichtexistierenden Nummern, Geschäftsanschlüssen etc.

² Über diese Nummern wurde im Befragungsverlauf mindestens einmal eine Person erreicht, mit der aber wegen Teilnahmeverweigerung, Problemen bei der Terminabsprache oder anderen Gründen kein Interview zustande kam.

Datengewichtung für die Auswertung

Das gewählte Zufallsverfahren ist zwar die beste Methode, um in der befragten Stichprobe ein möglichst gutes Abbild der Grundgesamtheit zu erstellen. Allerdings gibt es in der Forschungspraxis immer gewisse Einschränkungen, die sich u. a. in der Ausschöpfungsquote niederschlagen: nur ein kleinerer Teil der zufällig ausgewählten und angesprochenen Personen nimmt an der Befragung teil, ein größerer Teil lehnt ab oder ist nicht erreichbar. Dies geht zum Teil auf zufällige Faktoren zurück (z. B., wenn eine Person zum Zeitpunkt des Anrufs gerade keine Zeit hat oder nicht an den Apparat gehen kann), die eher wenig Einfluss auf die Richtigkeit der Ergebnisse haben. Aber es gibt auch „Verzerrungen“ in der Stichprobe, die dazu führen können, dass bestimmte Ergebnisse nicht ganz der Grundgesamtheit entsprechen. Dies kann daran liegen, dass Menschen mit bestimmten Eigenschaften häufiger die Teilnahme ablehnen oder dass bestimmte Personengruppen grundsätzlich schwerer erreichbar sind oder eine geringere Chance haben, für die Teilnahme ausgewählt zu werden.

Kennt man die Ursachen solcher Verzerrungen, kann man diese – zumindest für rein beschreibende, deskriptive Auswertungen – durch eine nachträgliche Gewichtung der Antworten ausgleichen. Personengruppen, von denen man weiß, dass sie in der Stichprobe überrepräsentiert sind (etwa Menschen mit höherem Bildungsgrad und mittlerem Einkommen), gehen dann mit geringerem Gewicht z. B. in Prozentberechnungen ein als Gruppen, die unterrepräsentiert sind. Solche Gewichtungen sind in der angewandten Umfrageforschung zwar üblich, tragen aber gewisse Risiken in sich, denn es ist – gerade in wissenschaftlichen Untersuchungen, bei denen man sich mit neuen, bisher noch kaum erforschten Phänomenen beschäftigt – oft unklar, ob eine Gewichtung auch tatsächlich die Qualität des gesamten Datensatzes verbessert oder nicht auch Verzerrungen verstärken kann (Alt & Bien 1994). Aus Sicht der wissenschaftlichen Statistik ist einem transparenten Umgang mit gewissen Verzerrungen von Daten der Vorzug zu geben gegenüber zu umfangreichen und komplexen Nachgewichtungsverfahren, die evtl. die Repräsentativität der Daten noch weiter verschlechtern können, zumindest aber eine Beurteilung von möglichen Verzerrungen erschweren, weil diese durch multiple Gewichtungen verschleiert werden

können. Wenn eine Gewichtung eingesetzt werden soll – und zumindest bei deskriptiven Auswertungen, bei denen Befunde beschreibend auf die Grundgesamtheit verallgemeinert

werden sollen, kann dies sinnvoll sein – sollte sie möglichst zurückhaltend, nachvollziehbar und auf Grundlage theoretischer Argumente durchgeführt werden (Rothe 1994).

Ein solcher Weg wurde für die vorliegende Studie gewählt. Zwar zeigten sich bestimmte Verzerrungen in verschiedenen Variablen, z. B. in der Verteilung der Befragten nach Haushaltsgröße; da es aber wenig bekannte Zusammenhänge zwischen der Haushaltsgröße der befragten Personen und den in dieser Studie relevanten Variablen – digitale Mediennutzung und politisches Engagement – gibt, wurde auf eine Gewichtung nach solchen Faktoren verzichtet. Dagegen ist die Verzerrung der Daten hinsichtlich des höchsten Bildungsabschlusses relativ deutlich, und von diesem sind klarere Zusammenhänge mit den hier relevanten Variablen bekannt. Es ist also mit höherer Wahrscheinlichkeit zu erwarten, dass eine Gewichtung dieses Faktors die Repräsentativität der Antworten tendenziell erhöht.

Als Grundlage für die Gewichtung wurden Daten des Mikrozensus herangezogen, die aufgrund der besonderen Erhebungsbedingungen als verlässlichste Quelle für die tatsächlichen Verteilungen von demografischen Variablen in der deutschen Bevölkerung dienen können.

Tabelle 2 dokumentiert das Verhältnis in Mikrozensus und Befragungsdaten und die auf dieser Basis festgelegten Gewichtungsfaktoren:

Tabelle 2

Abschluss	Mikrozensus 2017 ¹	Befragung 2019	Gewichtungsfaktor
Noch Schüler*in	3,6 %	2,0 %	1,77
Hauptschulabschluss	30,5 %	13,8 %	2,39
Realschulabschluss	29,9 %	25,8 %	1,16
Fachhochschulreife/Abitur	32,0 %	57,7 %	0,55

¹ Das Jahr 2017 ist das jüngste Jahr, für das Daten des Mikrozensus vorliegen. Die vom Statistischen Bundesamt veröffentlichten Daten sind leicht anders gruppiert, sie umfassen z. B. leicht andere Altersgruppen (ab 15 Jahren) und sind damit gegenüber der vorliegenden Erhebung (ab 16) zeitlich leicht verschoben. Auch die Bildungsgruppen unterscheiden sich z. T. vom Mikrozensus, etwa in der getrennten Ausweisung der gymnasialen Oberstufe, die hier der Gruppe „Abitur“ zugeschlagen wurde. Andere, kleinere Gruppen wurden wegen ihrer geringen Relevanz für den Gesamtdatensatz aus der Gewichtung ausgeschlossen und verblieben mit dem Gewicht 1 im Datensatz (z. B. 8 Personen ohne Schulabschluss); deshalb summieren sich die hier dokumentierten Gruppen auch nicht auf 100 % auf.

Es wurde ein Gewichtungsverfahren gewählt, das die Gesamtzahl der Fälle in der Summe nicht verändert, um über die gesamte Auswertung eine grundsätzliche Vergleichbarkeit der Daten herzustellen. Dies bedeutet, dass überrepräsentierte Personengruppen ein geringeres Gewicht erhalten, unterrepräsentierte dagegen ein höheres. Das mit diesem Vorgehen theoretisch verbundene Problem der Inflation möglicher Stichprobenfehler wurde aufgrund der relativ hohen Fallzahl in der Gruppe der größer gewichteten Abiturienten als relativ gering eingeschätzt. Hält man dieses Risiko für substanzieller, so wäre eine ausschließliche Kleingewichtung aller anderen Gruppen gegenüber der überproportional vertretenen Gruppe angeraten. Dieses Verfahren kann dann allerdings wiederum zu relativ geringen Fallzahlen in einzelnen Gruppen führen, was ebenfalls Probleme für die Datenanalyse nach sich ziehen kann.



Kontakt

Nadja Schaetz: nadja.schaetz@fu-berlin.de

Prof. Dr. Martin Emmer: martin.emmer@fu-berlin.de

Quellen

ADM (2012). ADM-Forschungsprojekt „Dual-Frame-Ansätze“ 2011/2012. Forschungsbericht. Abgerufen von: <https://www.adm-ev.de/wp-content/uploads/2018/07/Forschungsbericht-Dual-Frame-Ansätze.pdf>

ADM (2016). Richtlinie für telefonische Befragungen. Abgerufen von: <https://www.adm-ev.de/wp-content/uploads/2018/07/RL-Telefonbefragung.pdf>

Alt, C., & Bien, W. (1994). Gewichtung, ein sinnvolles Verfahren in der Sozialwissenschaft? In S. Gabler, J. H. P. Hoffmeyer-Zlotnik, & D. Krebs (Hrsg.), *Gewichtung in der Umfragepraxis* (S. 124-140). Opladen: Westdeutscher Verlag.

Bundesnetzagentur (2019). Teilnehmerentwicklung im Mobilfunk. Abgerufen von: https://www.bundesnetzagentur.de/cln_1422/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Marktbeobachtung/Deutschland/Mobilfunkteilnehmer/Mobilfunknehmer.html?nn=268208

Busse, B., & Fuchs, M. (2012). The components of landline telephone survey coverage bias. The relative importance of no-phone and mobile-only populations. *Quality & Quantity*, 46(4), 1209–1225. doi:10.1007/s11135-011-9431-3

Callegaro, M., Ayhan, Ö., Gabler, S., Häder, S., & Villar, A. (2011). Combining landline and mobile phone samples: a dual frame approach. *GESIS-Working Papers*, 2011|13. Köln: GESIS.

Europäische Kommission (2018). *Spezial-Eurobarometer 462: Elektronische Kommunikation und digitaler Binnenmarkt*. Bericht. Abgerufen von: <https://publications.europa.eu/en/publication-detail/-/publication/57889a55-8fb6-11e8-8bc1-01aa75ed71a1/language-de>

Gabler, S., & Häder, S. (1997). Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. *ZUMA-Nachrichten*, 21(41), 7–18.

Häder, S. (2015). Stichproben in der Praxis (Version 1.1). GESIS Survey Guidelines. doi:10.15465/gesis-sg_014

Häder, S., & Gabler, S. (1998). Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In S. Gabler, S. Häder, & J. H. P. Hoffmeyer-Zlotnik (Hrsg.), *Telefonstichproben in Deutschland* (S. 69–88). Wiesbaden: VS.

Häder, S., & Sand, M. (2019). Telefonstichproben. In S. Häder, M. Häder, & P. Schmich (Hrsg.), *Telefonumfragen in Deutschland* (S. 113–151). Wiesbaden: Springer VS.



Rothe, G. (1994). Wie (un)wichtig sind Gewichtungen? Eine Untersuchung am ALLBUS 1986. In S. Gabler, J. H. P. Hoffmeyer-Zlotnik, & D. Krebs (Hrsg.), *Gewichtung in der Umfragepraxis* (S. 62-87). Opladen: Westdeutscher Verlag.

Sand, M. (2014). Dual-Frame-Telefonstichproben: Entwicklung, Handhabung und Gewichtung. *GESIS-Technical Reports 2014|02*. Mannheim, Köln: GESIS.

Schneiderat G., & Schlinzig T. (2011). Mobile- and Landline-Onlys in Dual-Frame-Approaches: Effects on Sample Quality. In S. Häder, M. Häder, & M. Kühne (Hrsg.), *Telephone Surveys in Europe: Research and Practice* (S. 121–143). Berlin, Heidelberg: Springer.