

# Informationen & Recherchen

## Die Eignung von Umfragemethoden

**Methodische Einschätzung**

*Sabine Pokorny und Jochen Roose*



## Inhalt

<b>Quantitative Umfragemethoden</b>	<b>4</b>
<b>Was bedeutet „repräsentativ“?</b>	<b>4</b>
Auswahlverfahren .....	5
Realisierung der Auswahl.....	7
Interviewdurchführung .....	9
Bedingungen repräsentativer Umfragen im Überblick.....	9
<b>Face-to-face-Umfragen</b>	<b>10</b>
Vorgehen bei Face-to-face-Umfragen und Auswahlverfahren.....	10
Vorteile der persönlichen Interviews.....	10
Nachteile der persönlichen Interviews.....	11
<b>Telefonumfragen</b>	<b>12</b>
Vorgehen bei telefonischen Umfragen und Stichprobenziehung.....	12
Vorteile von telefonischen Umfragen.....	13
Nachteile von telefonischen Umfragen.....	13
<b>Online-Umfragen</b>	<b>14</b>
Vorgehen bei Online-Umfragen und Stichprobenziehung .....	14
Vorteile von Online-Umfragen.....	16
Nachteile von Online-Umfragen.....	16
Online-Access-Panels und Repräsentativität .....	17
<b>Fazit</b>	<b>19</b>
<b>Literatur</b>	<b>20</b>
<b>Impressum</b>	<b>24</b>

## Auf einen Blick

- › Die Repräsentativität von Umfragen hängt maßgeblich vom realisierten Auswahlverfahren und weniger von der Anzahl an Befragten ab.
- › Basis für eine repräsentative Umfrage ist eine Zufallsstichprobe. Insbesondere Stichproben, die auf einer willkürlichen Auswahl beruhen (z.B. wenn sich die Befragten aus eigener Motivation beteiligen können), produzieren keine repräsentativen Ergebnisse. Die Erhebungsform (face-to-face, telefonisch, online) und die damit verbundenen Möglichkeiten der Auswahl sind maßgeblich für die Qualität der erhobenen Daten.
- › Gerade für politische Entscheidungen ist es wichtig, dass korrekte und belastbare Ergebnisse in die Entscheidungen eingehen. Die Qualität von Umfragen und die Passung des Vorgehens zu den benötigten Informationen sind wesentlich für die Frage, ob korrekte Ergebnisse vorliegen. Eine sorgfältige Prüfung dieser Frage sollte Teil der Rezeption von Umfrageergebnissen und ihrer Beauftragung sein.

## Quantitative Umfragemethoden

Wie die Bevölkerung über diese oder jene Frage denkt, wird vielfältig in Umfragen erhoben und publiziert. Mit der Markierung von Ergebnissen als „repräsentativ“ geht die Behauptung einher, die Umfrage habe das Stimmungsbild in der Bevölkerung richtig erhoben. Allerdings sind Ergebnisse von Umfragen, je nach Art der durchgeführten Umfrage, unterschiedlich belastbar. Nicht jedes Umfrageergebnis lässt sicher den Schluss zu, dass in der Bevölkerung oder in anderen befragten Gruppen tatsächlich die Einstellungen in der dargestellten Weise verteilt sind. Die Beurteilung der Qualität von Umfragen erfordert einiges Detailwissen, ist aber unerlässlich für die richtige Einschätzung von Umfrageergebnissen. Wer Umfragen liest und mehr noch, wer sie beauftragt, kommt nicht umhin, sich mit Fragen der Qualität von Umfrageergebnissen zu befassen.

Das Angebot an Umfragen und Umfragemethoden ist groß, vor allem im Bereich der quantitativen Umfragen. Die Mehrheit der Interviews in der quantitativen Markt- und Sozialforschung wird inzwischen online durchgeführt. Laut Erhebung des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM) unter seinen Mitgliedern 2018 wurden 40 Prozent der quantitativen Interviews online durchgeführt, 28 Prozent der Interviews wurden telefonisch geführt und 23 Prozent waren persönlich-mündliche, also Face-to-face-Interviews. Der Anteil der schriftlichen Interviews spielte mit 5 Prozent kaum noch eine Rolle (ADM 2020).

Diese verschiedenen Erhebungsmethoden haben unterschiedliche Vor- und Nachteile, derer sich ein Leser von Umfrageergebnissen bewusst sein sollte. Nicht jede Umfragemethode ist für jeden Auswertungszweck geeignet. Zusätzlich gibt es auch innerhalb der Umfragemethoden Unterschiede. Die Unterscheidung nach Erhebungsart, also telefonisch, online oder face-to-face, gibt noch keine Auskunft über das Verfahren der Stichprobenziehung, das aber für die Qualität von Umfragen von großer Bedeutung ist.

Im Folgenden soll ein kurzer Überblick über die drei häufigsten quantitativen Erhebungsmethoden mit ihren jeweiligen Vor- und Nachteilen, den verschiedenen Stichprobenverfahren und der Frage, inwiefern die Ergebnisse als repräsentativ gelten können, gegeben werden. Qualitative Umfragen wie Tiefeninterviews oder Fokusgruppen sind ebenfalls von großer Bedeutung, erheben aber grundsätzlich nicht den Anspruch, repräsentativ zu sein. Aus diesem Grund werden sie hier nicht berücksichtigt.

### Was bedeutet „repräsentativ“?

Umfragen behaupten gern von sich, repräsentativ zu sein. Es wird als Qualitätskriterium einer Umfrage angesehen. Richtig angewendet trifft dies auch zu.

Repräsentativität bezieht sich auf das Verhältnis von Ergebnissen, die aus einer Stichprobe gewonnen wurden, zu den wahren Werten in der Grundgesamtheit. Mit Grundgesamtheit sind alle Fälle gemeint, über die eine Aussage gemacht werden soll. Bei Meinungsumfragen sind dies in der Regel Personen, wobei der Personenkreis eingegrenzt ist, beispielsweise auf alle wahlberechtigten Menschen in Deutschland oder auch alle in Deutschland lebenden Jugendlichen oder ähnliches. Die Definition der Grundgesamtheit ist abhängig davon, über wen die Umfrage etwas aussagen soll. Um die Repräsen-

tativität einer Umfrage beurteilen zu können, ist demnach zunächst die Kenntnis nötig, für welche Grundgesamtheit eine Aussage gemacht werden soll.<sup>1</sup>

Repräsentativität im Sinne der sozialwissenschaftlichen Umfrageforschung bedeutet, dass eine Stichprobe „ein verkleinertes Abbild einer angebbaren Grundgesamtheit“ ist (Kromrey/Roose/Strübing 2016: 264; genauso z.B. Diaz-Bone/Weischer 2015: 350 f. oder die Einträge zu *representativeness* bei Porta 2014 oder Scott 2014). Repräsentativität ist also dann gegeben, wenn man aus einer Stichprobe korrekte Schlüsse auf die Grundgesamtheit ziehen kann. Sie bezeichnet das Verhältnis von Stichprobe zur Grundgesamtheit für eine Datenerhebung.

Um aus einer Stichprobe Informationen zu ermitteln, die entsprechende Merkmale in der Grundgesamtheit richtig wiedergeben, sind drei Schritte entscheidend:

1. das Auswahlverfahren, wie also die Stichprobe aus der Grundgesamtheit gezogen wird.
2. die Realisierung der Auswahl, in welchem Maße also die geplante Auswahl tatsächlich erfolgreich umgesetzt wird.
3. die Erhebung von Informationen in der realisierten Auswahl, ob also die eigentliche Datenerhebung die Informationen richtig sammelt.

## Auswahlverfahren

Das Auswahlverfahren gibt an, wie die Fälle der Grundgesamtheit in die Stichprobe gelangen. Verschiedene Auswahlverfahren führen mit unterschiedlicher Wahrscheinlichkeit zu einer repräsentativen Stichprobe.<sup>2</sup> Dabei ist es nicht am wichtigsten, wie groß die Stichprobe ist, wie viele Personen also befragt wurden. Eine Umfrage mit 6.000 Teilnehmenden ist nicht zwingend besser als eine Umfrage mit 1.000 Teilnehmenden. Viel entscheidender ist, wie die Befragten in die Stichprobe gelangen, also welches Auswahlverfahren angewendet wurde.

Der Königsweg zu einer repräsentativen Stichprobe ist die einfache Zufallsauswahl (Langer 2018). Die Besonderheit der Zufallsauswahl besteht darin, dass sie keine Kenntnisse über Verteilungen von Merkmalen oder Zusammenhänge zwischen Merkmalen in der Grundgesamtheit voraussetzt. Mit der Logik der Wahrscheinlichkeitsrechnung lässt sich zeigen, dass bei einer tatsächlichen Zufallsauswahl der wahre Wert eines Merkmals in der Grundgesamtheit mit einer angebbaren Wahrscheinlichkeit in der Nähe des Wertes liegen muss, der in der Stichprobe gefunden wird (Kromrey/Roose/Strübing 2016: 278).<sup>3</sup>

---

<sup>1</sup> Entsprechend lässt sich die Behauptung der Repräsentativität einer Umfrage solange nicht beurteilen, wie die gemeinte Grundgesamtheit unbekannt ist. Eine Umfrage kann immer nur für eine angegebene Grundgesamtheit mehr oder weniger repräsentativ sein.

<sup>2</sup> Rein logisch kann jedes Auswahlverfahren zu einer repräsentativen Stichprobe führen. Gleichzeitig lässt sich für eine Stichprobe nur sehr selten herausfinden, ob sie tatsächlich repräsentativ ist, weil dazu eine wahre Kenntnis über die Merkmale in der Grundgesamtheit notwendig wäre. Aber genau diese Kenntnis fehlt, sonst bräuchte es keine Untersuchung. Entsprechend interessiert, auf welche Weise eine Stichprobe gezogen werden kann, die mit einer hohen Wahrscheinlichkeit repräsentativ für die Grundgesamtheit ist. Die logische Möglichkeit einer tendenziell repräsentativen Stichprobe kann aber missbraucht werden, um das Konzept selbst zu verwässern.

<sup>3</sup> Dieser Bereich, in dem der wahre Wert in der Grundgesamtheit nach den Berechnungen der Wahrscheinlichkeitstheorie mit einer angebbaren Wahrscheinlichkeit liegen muss, wird Vertrauensintervall (Konfidenzintervall) genannt und wird manchmal bei Studien mit angegeben. In der Regel wird der Bereich angegeben, in dem der wahre Wert mit einer Wahrscheinlichkeit von 95 Prozent liegt. Korrekt ist diese Aussage allerdings nur für Zufallsauswahlen, während bei anderen Auswahlformen eine solche

Bei der Zufallsauswahl wird in einem kontrolliert zufälligen Verfahren aus der gesamten Grundgesamtheit eine festgelegte Anzahl von Fällen gezogen.<sup>4</sup> Die Auswahl kann auf die Fälle der Grundgesamtheit selbst (zum Beispiel Personen) oder auch eine Listenrepräsentation aller Fälle (beispielsweise Listen aller Personen im Einwohnermeldeamt) angewendet werden. Die Zufallsauswahl ist auch möglich, wenn bei dem Auswahlverfahren zwar nicht alle Fälle die identische Chance haben, gezogen zu werden, aber eine bekannte Chance. Das Ergebnis einer einfachen Zufallsauswahl ist nicht eine garantiert repräsentative Stichprobe, aber die Stichprobe ist mit angegebbarer Wahrscheinlichkeit eine strukturtreue Abbildung der Grundgesamtheit, also repräsentativ.

Tatsächlich lässt sich eine solche einfache Zufallsauswahl häufig nicht realisieren. Ein Problem des Auswahlverfahrens liegt in der Zugänglichkeit der Fälle der Grundgesamtheit bzw. Verfügbarkeit von Listen über alle Personen der Grundgesamtheit. Oftmals stimmt der Auswahlrahmen, also die Menge von Fällen, aus denen tatsächlich ausgewählt wird, nicht exakt mit allen Fällen der Grundgesamtheit überein. Bei einer Telefonumfrage werden Menschen ohne Telefon nicht erreicht, bei einer Erhebung auf Basis von Einwohnermeldedaten können Wohnungslose oder nicht gemeldete Personen nicht in die Stichprobe kommen. Bei einer Auswahl auf der Basis von Verzeichnissen, beispielsweise Mitglieder-Verzeichnissen, kann die Abweichung zwischen angezielter Grundgesamtheit und Auswahlrahmen noch deutlich größer sein. Abhängig vom Auswahl- und Erhebungsverfahren gibt es bessere oder schlechtere Annäherungen. In welchem Maße eine Annäherung an die Zufallsauswahl möglich ist, beeinflusst die Wahrscheinlichkeit, eine repräsentative Stichprobe zu erreichen. Dies wird für die verschiedenen Befragungsarten in den folgenden Abschnitten diskutiert.

Es gibt weitere Auswahlverfahren, die nicht dem Zufallsprinzip folgen (Kromrey/Roose/Strübing 2016: 265-277). Hierunter fallen die bewusste Auswahl und die Willkürauswahl. Eine bewusste Auswahl entscheidet über Auswahlkriterien und identifiziert die entsprechenden auszuwählenden Fälle. Ob damit Repräsentativität erreicht wird, ist abhängig von der Qualität der Kriterien, die zur bewussten Auswahl herangezogen werden. Grundsätzlich wird die Repräsentativität der bewussten Auswahl umso besser, je besser das Wissen über die interessierenden Merkmale ist. Da aber Studien in der Regel durchgeführt werden, weil die interessierenden Merkmale eben nicht bekannt sind, erreicht die bewusste Auswahl nur in Sonderfällen Repräsentativität.

Bei mehrstufigen Auswahlen kann eine Kombination von bewusster Auswahl und Zufallsauswahl eine sehr gute Annäherung an eine repräsentative Auswahl ermöglichen. Bei Face-to-face-Befragungen werden beispielsweise oft Regionen bewusst ausgewählt und die zu befragenden Personen sind durch eine Zufallsauswahl innerhalb der Regionen ermittelt. Weil über die Regionen relativ viele Informationen zur Verfügung stehen, ist hier oft eine gute bewusste Auswahl möglich (siehe auch den Abschnitt zu Face-to-face-Umfragen). Eine andere Form der bewussten Auswahl innerhalb eines mehrstufigen Auswahlverfahrens ist die Quotenauswahl. Hierbei wird für bestimmte Gruppen innerhalb der Grundgesamtheit jeweils die Anzahl von Befragten festgelegt, die in der Stichprobe sein sollen (Quoten). So kann festgelegt sein, dass in der Stichprobe gleich viele Männer und Frauen sein sollen oder die Anzahl der Befragten in bestimmten Altersgruppen festgelegt sein. Innerhalb der quotierten

---

Aussage nicht möglich ist. Werden dennoch entsprechende Aussagen getroffen, sind sie nicht durch die Wahrscheinlichkeitsrechnung abgesichert.

<sup>4</sup> Für die Zufallsauswahl ist ein kontrolliertes Verfahren notwendig, das den Einfluss der handelnden Personen ausschließt. Würde man einer Person den Auftrag geben, „zufällig“ Teilnehmer einer Befragung auszuwählen, würden bewusste und unbewusste Kriterien (z.B. Sympathie, Erreichbarkeit) die Auswahl beeinflussen. Das Ergebnis wäre eine Willkürauswahl, die zu besonders schlechten, mit hoher Wahrscheinlichkeit nicht-repräsentativen Stichproben führt.

Merkmale wird dann per Zufall ausgewählt. So könnte festgelegt sein, dass der Anteil (Quote) von Männern zwischen 18 und 30 Jahren in der Stichprobe ihrem bekannten Anteil in der deutschen Wohnbevölkerung entsprechen soll. Innerhalb dieser Gruppe der 18- bis 30-jährigen Männer würden dann aber durch Zufall konkrete Befragte ausgewählt.

Bei der Willkürauswahl werden Fälle ohne Kriterien in die Stichprobe aufgenommen. Dies ist keine Zufallsauswahl, sondern es gehen implizit Kriterien in die Auswahl ein, ohne dass klar wäre, welches diese Kriterien sind. Das *convenience sampling*, also eine Auswahl nach praktischer Erreichbarkeit, ist eine Form der Willkürauswahl. Die Selbstrekutierung der Befragten ist eine weitere Art der Willkürauswahl. Die Willkürauswahl hat die geringste Wahrscheinlichkeit, zu einer repräsentativen Stichprobe zu führen. Insbesondere kann nur darüber spekuliert werden, nach welchen Kriterien die Stichprobe entstanden ist.

## Realisierung der Auswahl

Neben dem Auswahlverfahren ist die tatsächliche Realisierung der Auswahl, also die tatsächliche Befragung der ausgewählten Personen, ein Problem, das die Repräsentativität maßgeblich beeinflusst.

Die Realisierung der Auswahl ist abhängig von der Umsetzung durch die dafür zuständigen Personen sowie der Bereitschaft der ausgewählten Befragungspersonen, auf die ihnen gestellten Fragen wahrheitsgemäß zu antworten. Für die Umsetzung des Auswahlverfahrens sind die technischen Gegebenheiten, praktischen Schwierigkeiten und Anreize für die Auswählenden entscheidend. Abhängig vom Erhebungsverfahren können manche Schritte eines Auswahlverfahrens automatisiert und so einem möglicherweise verfälschenden Einfluss der Auswählenden entzogen werden (zum Beispiel in der Telefonumfrage mit automatischen Wählsystemen). Andere Auswahlsschritte müssen Personen übernehmen. Dann ist entscheidend, welche Anreize für die Auswählenden bestehen, die Regeln etwas „großzügig zu interpretieren“, also zu verfälschen.

Realisierte Auswahlen weichen oftmals von reinen Zufallsauswahlen ab. Dies liegt zum einen an unterschiedlichen, aber bekannten Wahrscheinlichkeiten, in die Stichprobe zu gelangen. Diese unterschiedliche Wahrscheinlichkeit kann durch eine Gewichtung der Daten (Design-Gewicht oder Redressement-Gewicht) korrigiert werden. So wird bei Telefonumfragen mit erhoben, wie viele Personen unter der zufällig ermittelten Telefonnummer erreicht werden können oder andersherum, unter wie vielen Nummern eine befragte Person erreicht werden kann. Aus diesen Informationen lässt sich dann ermitteln, mit welcher Wahrscheinlichkeit die befragte Person bei zufällig ermittelten Telefonnummern ausgewählt werden konnte. Wenn die unterschiedliche Auswahlwahrscheinlichkeit korrekt bekannt ist, kann das Design-Gewicht korrekt berechnet werden.

Bei der Erhebung selbst können ebenfalls Fehler entstehen, die zu einer Verzerrung führen. Auf der einen Seite entstehen Fehler durch Versehen, Missverständnisse, Übertragungsfehler oder Ähnliches. Solange sich diese Fehler zufällig verteilen (*Missing at Random*), stellen sie kein prinzipielles Problem dar. Sie geben also den Ergebnissen keine falsche Richtung. Anders sieht es aus, wenn Fehler nur in eine Richtung entstehen, die erhobenen Informationen also in eine bestimmte Richtung verfälschen.

Bei der Realisierung von Interviews kann es zu fehlenden Informationen kommen, wenn die Kontaktaufnahme scheitert, ein Interview explizit abgelehnt wird oder beim Interview falsche Angaben gemacht werden (Engel/Schmidt 2019; Schnell 2019: 145-174). Die Kontaktaufnahme für ein Interview kann aus den unterschiedlichsten Gründen scheitern. Manche dieser Gründe können sich gleichmäßig über die Bevölkerung verteilen. Vielfach sind es aber Lebensumstände, die Kontakte wahrscheinlicher oder unwahrscheinlicher machen. Mobile oder viel beschäftigte Menschen sind zum Beispiel schwerer zu

erreichen als wenig beschäftigte. Wie wahrscheinlich solche Interviewausfälle sind, ist von der Art der Umfrage abhängig.

Ein erhebliches Problem in der Umfrageforschung ist die Antwortverweigerung. Bei verschiedenen Formen der allgemeinen Bevölkerungsumfrage ist die Antwortbereitschaft deutlich gesunken (Aust/Schröder 2009; Schnell 2019: 151-153), wie im Folgenden für die einzelnen Umfrageformen diskutiert wird. Die Sensibilität für den Schutz persönlicher Informationen scheint zuzunehmen. Auch wenn genaue Informationen über Personen, die keine Information preisgeben möchten, kaum zu bekommen sind, dürften bestimmte Menschen bzw. Menschen mit bestimmten Einstellungen stärker zur Antwortverweigerung tendieren als andere. Dabei sind die konkreten Motive aber durchaus vielfältig. Eine empfundene Verletzung der Privatsphäre kann genauso die Antwortbereitschaft beeinflussen wie ein Desinteresse am Thema der Studie, eine Ablehnung des Auftraggebers der Studie, Ablehnung von Forschung allgemein, Angst vor Betrug oder andere Gründe (Schnell 2019: 146-147). Je nach Studieninhalt können durch diese Gründe relevante und interessante Teile der Bevölkerung in der Umfrage nicht enthalten sein.

In den vorliegenden Informationen aus einer Befragung sind für bestimmte Aspekte Vergleiche zwischen der Stichprobe und der Grundgesamtheit möglich. Dies gilt für Merkmale, die in der Umfrage erhoben werden und für die, bezogen auf die Grundgesamtheit, recht genaue Informationen verfügbar sind. So ist sehr genau bekannt, wie viele Männer und Frauen in Deutschland leben, wie sich die Menschen auf die Bundesländer verteilen oder wie die Altersverteilung der Menschen ist. Auch über die Schulabschlüsse in der Bevölkerung liegen sehr genaue Informationen vor. Der Vergleich dieser Merkmale zwischen realisierter Stichprobe und Grundgesamtheit weist regelmäßig nennenswerte Unterschiede auf.

In Reaktion auf Unterschiede dieser Art ist es übliche Praxis, eine Anpassung zwischen Stichprobe und Grundgesamtheit mithilfe von Gewichtung vorzunehmen (Kiesl 2019). Dabei handelt es sich um ein Struktur-Gewicht (*post-stratification*-Gewicht). Das Gewicht gleicht die Verteilung der Strukturmerkmale (für Deutschland in der Regel Geschlecht, Alter und Schulabschluss, zum Teil auch Bundesland) an die bekannten Werte der Grundgesamtheit an.

Die Anwendung des Struktur-Gewichts ist allerdings an Voraussetzungen gebunden. Mit dem Struktur-Gewicht erhalten Befragte aus jenen Personengruppen, die eine geringere Wahrscheinlichkeit haben, Teil der Stichprobe zu werden, ein höheres Gewicht. Wären in einer Stichprobe beispielsweise nur halb so viele junge Menschen mit niedrigerem formalem Schulabschluss wie in der Grundgesamtheit, kann jede Person in der Stichprobe mit diesen Merkmalen doppelt gewichtet werden. In der gewichteten Stichprobe entspricht dann das Verhältnis dieser Personengruppe zu anderen wieder dem Verhältnis, wie es in der Grundgesamtheit zu finden ist.

Dieses Vorgehen setzt aber eine wichtige Annahme voraus: Es geht davon aus, dass die Verteilung der interessierenden Merkmale (zum Beispiel politische Einstellungen oder Ähnliches) unter den Befragten dieser Personengruppe ähnlich ist wie in dieser Personengruppe in der Grundgesamtheit. Sind also Merkmale zu selten in einer Stichprobe, können sie durch das Struktur-Gewicht stärker in die Gesamtergebnisse eingehen und so die Ergebnisse korrigieren. Sind allerdings bestimmte Befragte gar nicht in der Stichprobe vertreten, weil sie grundsätzlich an einer Befragung dieser Art nicht teilnehmen (oder nicht teilnehmen können), kann dieses Problem auch nicht durch ein Struktur-Gewicht korrigiert werden.

## Interviewdurchführung

Für gültige, repräsentative Ergebnisse ist schließlich die Richtigkeit der jeweils gemachten Angaben auf die einzelnen Fragen wichtig. Auch hier lauert eine Fehlerquelle (Jedinger/Michael 2019). In manchen Fällen wissen Befragte nicht die richtige Antwort, geben aber dennoch eine, um sich nicht in der Situation zu blamieren. Je nach Art der Frage kann den Befragten die eigentlich richtige Angabe unangenehm sein oder sie wollen gegenüber der interviewenden Person einen bestimmten Eindruck machen und geben Antworten, von denen sie annehmen, die interviewende Person würde sie gern hören. Antworten dieser Art werden als „sozial erwünschte Antworten“ bezeichnet. In welchem Maße dieses Problem auftritt, hängt von den Inhalten der Fragen ab, aber auch von der Situation der Befragung. Erscheinung und Verhalten der Interviewer können die Antworten mehr oder weniger beeinflussen.

Zu unrichtigen Angaben kann es schließlich kommen, wenn Befragte das Ergebnis einer Umfrage gezielt in eine Richtung beeinflussen wollen. Dieses Problem stellt sich allerdings nur, wenn das Auswahlverfahren es Interessierten erlaubt, sich selbst als Befragungsteilnehmer zu rekrutieren. Bei Zufallswahlen ist dieses Problem ausgeschlossen.

## Bedingungen repräsentativer Umfragen im Überblick

Tabelle 1 gibt einen Überblick über mögliche Probleme bei der Durchführung von Umfragen, die zu geringerer Repräsentativität, also Abweichungen von Merkmalen der Stichprobe von den wahren Merkmalen in der Grundgesamtheit, führen können.

Tabelle 1: Mögliche Probleme für repräsentative Ergebnisse

Durchführungsschritt der Umfrage	Mögliche Entstehung von eingeschränkt repräsentativen Ergebnissen
Auswahlverfahren	<ul style="list-style-type: none"><li>• Unvollständiger Auswahlrahmen</li><li>• Nicht-repräsentative Stichprobe durch Willkür Auswahl</li><li>• Nicht-repräsentative Stichprobe durch falsch definierte bewusste Auswahl</li></ul>
Auswahldurchführung	<ul style="list-style-type: none"><li>• Interviewvermeidung durch Interviewer</li><li>• Nicht-Erreichbarkeit</li><li>• Verweigerung</li></ul>
Interviewdurchführung	<ul style="list-style-type: none"><li>• Antwort trotz Nichtwissen</li><li>• Sozial erwünschtes Antwortverhalten</li><li>• Manipulation durch Befragte</li></ul>

Die zahlreichen möglichen Beeinträchtigungen der Repräsentativität von Umfragen mögen den Eindruck erwecken, Umfragen seien kein geeigneter Weg, um sich angemessen über die Bevölkerungsmeynung zu informieren. Dies trifft allerdings so nicht zu. Zum einen ist für die Repräsentativität der Ergebnisse die Qualität der durchgeführten Umfrage maßgeblich. Die Ergebnisse einer Umfrage können mehr oder weniger von den tatsächlichen Werten in der Grundgesamtheit abweichen und somit mehr oder weniger repräsentativ sein. Zum anderen sind alternative Verfahren, um sich zu informieren, ebenfalls von diesen Problemen betroffen. Wenn ich mich im Gespräch mit anderen Bürgern selbst informiere, bin ich aller Wahrscheinlichkeit nach noch deutlich stärker mit den Problemen einer nicht-repräsentativen Auswahl oder sozial erwünschten Antworten konfrontiert. Die Alternative zu schlechten Umfragen sind daher gute Umfragen und nicht der Verzicht auf Umfragen.

## Face-to-face-Umfragen

### Vorgehen bei Face-to-face-Umfragen und Auswahlverfahren

Bei persönlichen Interviews werden Menschen von Angesicht zu Angesicht, also face-to-face, befragt.<sup>5</sup> In der Regel kommt ein Interviewer zu den Menschen nach Hause und stellt ihnen die standardisierten Fragen. Die Antworten werden von dem Interviewer direkt in Papierbögen oder in einen Computer (*Computer Assisted Personal Interview* – CAPI) eingetragen.

In der Interviewführung können die Fragen vorgelesen werden, es ist aber auch möglich, eine größere Anzahl von Antwortmöglichkeiten in Papierform vorzulegen. Genauso können den Befragten in der Situation Bilder oder anderes Material vorgelegt werden. Die gemeinsame Anwesenheit von Interviewer und befragter Person ermöglicht eine große Vielfalt von Befragungsarten.

Die Auswahl der zu befragenden Personen kann im Prinzip aus den Registern des Einwohnermeldeamtes erfolgen. Dies könnte eine sehr gute Annäherung an eine Zufallsauswahl unter der Wohnbevölkerung sein. Tatsächlich ist dieses Vorgehen, Befragte an den unterschiedlichsten Orten aufzusuchen, ausgesprochen aufwendig (Schnell 2019: 184) und damit letztlich kaum durchführbar.

Stattdessen gibt es bei persönlichen Befragungen eine Beschränkung auf ausgewählte Orte. Das Auswahlverfahren ist entsprechend mehrstufig. In einer ersten Stufe erfolgt die Festlegung von Orten, an denen befragt werden soll. Diese Auswahl kann als Zufallsauswahl erfolgen, meist ist es aber eine bewusste Auswahl von unterschiedlichen Startpunkten, mit denen die Vielfalt der zu befragenden Bevölkerung möglichst gut abgedeckt ist. Da für kleinräumige Einheiten (z. B. Straßenzüge oder Wohnblocks) recht gute Informationen vorliegen, kann in dieser Stufe eine bewusste Auswahl für das Ziel einer repräsentativen Auswahl sinnvoll sein.

In einem zweiten Schritt werden vor Ort Befragte durch ein Random-Route-Verfahren (Zufallsweg-Verfahren) ausgewählt (Häder 2015: 150-157; Schnell 2019: 196-197). Die Auswahl folgt vor Ort nach Anweisungen, von welchem Startpunkt welcher Weg zu gehen ist und beim wievielten Haushalt ein Interview mit welcher Person zu führen ist. Auf diese Weise soll sich die Auswahl der Zielpersonen einer reinen Zufallsauswahl größtmöglich annähern.

In der Praxis erweist sich dieses Verfahren als schwierig. Schon eine präzise Formulierung von Anweisungen für den Zufallsweg ist kompliziert, weil sie alle vorkommenden Besonderheiten (z.B. Sackgassen, Hofeinfahrten, Verwaltungsgebäude, mehrere Eingänge zu einem Haus) berücksichtigen muss. In der Anwendung müssen dann all diese Sonderfälle korrekt erkannt und abgearbeitet werden, was wiederum schwierig ist.

### Vorteile der persönlichen Interviews

Zwei Aspekte stechen als Vorteile der persönlichen Interviews hervor. In der Befragungssituation ist das persönliche Interview sehr flexibel. Unterschiedlichste Frageformen sind durch die direkte Anwesenheit von Interviewer und befragter Person möglich. Vergleichsweise komplizierte und lange Interviews sind auf diese Weise durchführbar.

---

<sup>5</sup> Allgemein zu Face-to-face-Umfragen siehe Stocké (2019) und Schnell (2019: 177 ff.).

Ein weiterer erheblicher Vorteil liegt in den Möglichkeiten, mit dem Auswahlverfahren Repräsentativität zu erreichen. Einwohnermeldedaten sind von hoher Qualität und im Prinzip für eine Zufallsauswahl besonders gut geeignet. Doch auch das Random-Route-Verfahren, das nach einem Zufallsprinzip real existierende Häuser und Wohnungen auswählt, profitiert von dieser Vollständigkeit. Alle Wohnungen können im Prinzip durch ein solches Verfahren ausgewählt werden und der Anteil von Menschen, die durch dieses Verfahren prinzipiell nicht in die Stichprobe gelangen können, ist ausgesprochen klein. Dieser Vorteil findet allerdings seine Grenzen in der praktischen Umsetzung.

## Nachteile der persönlichen Interviews

Ein erheblicher Nachteil von persönlichen Interviews sind die hohen Durchführungskosten. Diese sind stark abhängig von der Anzahl der Orte, an denen Interviews durchgeführt werden. Allerdings trägt eine größere Zahl verschiedener Orte zu einer besseren Stichprobe bei. Mehr verschiedene Erhebungsorte erhöhen die Chance auf eine repräsentative Stichprobe.

Die Qualität der Stichprobe ist davon abhängig, wie gut der Stichprobenplan umgesetzt wird. Dies wird beeinflusst von der Klarheit der Regeln, aber auch von der Umsetzung der Regeln vor Ort. Unsystematische Fehler wären für die Qualität der Stichprobe nicht relevant, doch es ist durchaus von systematischen Fehlern auszugehen. Identifizieren die Interviewer selbst ihre Gesprächspartner, so können zudem die Vorlieben der Interviewer einflussreich werden.

Ein weiteres Problem ergibt sich, wenn die zu befragende Person beim Besuch des Interviewers nicht anwesend ist. In diesem Fall ist ein zusätzlicher Besuch mit allem verbundenen Aufwand notwendig. Kommt es zu mehreren Kontaktversuchen, so finden diese nicht selten in schneller Folge statt.<sup>6</sup> Da in Deutschland Interviewer in der Regel pro realisiertem Interview bezahlt werden, ist der Anreiz groß, eine andere Person zu befragen oder anderweitig das Vorgehen abzukürzen (Schnell 2019: 183-184). Sodeur (2007) konnte zeigen, dass der Anteil von befragten Frauen in gemischtgeschlechtlichen Partnerschaften mit den Einflussmöglichkeiten der Interviewer bei der Auswahl der Zielperson zunimmt. In diesem Fall wurde also die leichter erreichbare Person im Haushalt überdurchschnittlich häufig befragt, wenn dem Interviewer ausreichend Einfluss auf den Auswahlprozess zur Verfügung stand.

Ein ähnliches Problem ergibt sich für die Interviewführung selbst. So konnten beispielsweise Matschinger/Bernert/Angermeyer (2005, vgl. Schnell 2019: 229) zeigen, dass Antworten, die umfangreiche Nachfragen und damit ein längeres Interview bedingen, mit der Dauer der Feldzeit seltener wurden. Dies scheint ein Beispiel zu sein, wie Interviewer sich in der Befragung die Arbeit „unauffällig“ ein wenig leichter machten und so die Ergebnisse verzerrten.

Die Komplexität des Erhebungsverfahrens beeinflusst auch die Geschwindigkeit, mit der Daten erhoben werden können. Insbesondere für kurzfristige Umfragen, die schnell Ergebnisse liefern sollen, ist diese Methode kaum geeignet.

Die Kontrolle der Interviewdurchführung ist bei persönlichen Interviews nur begrenzt möglich (Details bei Schnell 2019: 217). Während bei telefonischen Interviews für den Interviewer unbemerkt ein Mithören (zumindest der Fragen) möglich ist, entfällt eine solche Möglichkeit bei persönlichen Interviews. Wählt der Interviewer die Befragten selbst aus, ist die Umsetzung der Auswahlregeln

---

<sup>6</sup> In einer älteren Studie zeigt Schnell (1997: 231-234), dass bei der renommierten wissenschaftlichen Bevölkerungsumfrage ALLBUS die zusätzlichen Kontaktversuche nach einer Abwesenheit meist innerhalb einiger Stunden bis weniger Tage stattfanden.

ebenfalls nur in kleinen Stichproben möglich. Die Tatsache, dass Kontrollen nur begrenzt möglich sind, kann bereits Auswirkungen auf das Verhalten der Interviewer und ihre Zuverlässigkeit haben.

Schließlich sind im Vergleich der Erhebungsmethoden bei den persönlichen Interviews die stärksten Interviewereffekte zu vermuten (Schnell 2019: 197-199, Jedinger/Michael 2019, West/Blom 2017). Durch die persönliche Begegnung nimmt die befragte Person den Interviewer in seiner ganzen Person wahr. Seine Kleidung, sein Auftreten, sein Mienenspiel, all dies ist für den Befragten sichtbar und kann ihn beeinflussen. Bei allen anderen Erhebungsmethoden tritt der Interviewer deutlich weniger in Erscheinung. Beeinflussungen sind zudem von anwesenden dritten Personen möglich. Wenn das Interview, wie üblich, beim Befragten stattfindet, lässt sich nur sehr begrenzt beeinflussen, ob weitere Personen beim Interview zugegen sind. Dies kann Einfluss auf das Antwortverhalten haben, insbesondere bei Fragen, die als sensibel angesehen werden, was auf Einstellungen schnell zutreffen kann.

## Telefonumfragen

### Vorgehen bei telefonischen Umfragen und Stichprobenziehung

Bei telefonischen Umfragen werden die Befragten telefonisch am Festnetzapparat oder am Mobiltelefon angerufen. Die Interviewer sind bei ihrer Arbeit in einem Telefonstudio oder telefonieren von zu Hause. Die Verbindung wird in der Regel computergesteuert hergestellt, wobei zunächst automatisch festgestellt wird, ob die angewählte Nummer eine gültige Telefonnummer ist und ob ein Mensch (und nicht beispielsweise ein Anrufbeantworter) abgenommen hat. Erst nach diesem Prüfschritt werden Interviewer und Befragter verbunden. Bei dem Interview werden die Fragen abgelesen und die Antworten direkt an einem Computer eingegeben. Die übliche Bezeichnung CATI steht für *Computer Assisted Telephone Interview*.<sup>7</sup>

Grundsätzlich können Telefonumfragen mit unterschiedlichen Stichprobenverfahren durchgeführt werden. In aller Regel handelt es sich bei telefonischen Umfragen jedoch um Zufallsauswahlen der Befragten. Früher wurde eine solche Zufallsauswahl anhand von Eintragungen im Telefonverzeichnis vorgenommen. Das war sinnvoll, da alle Telefonnummern eingetragen werden mussten. Seit die Eintragungspflicht abgeschafft ist, sind jedoch immer weniger Telefonanschlüsse eingetragen, sodass dieses Verfahren nicht mehr für eine belastbare zufallsbasierte Stichprobenziehung geeignet ist. Stattdessen erfolgt die Auswahl der Befragten in der Regel auf Basis von Rufnummernblöcken der Bundesnetzagentur, in denen auch nicht eingetragene Rufnummern generiert werden (Häder 2015: 160; Häder/Sand 2019: 50 ff.). Dieser Auswahlrahmen wird vom ADM seinen Mitgliedsinstituten zur Verfügung gestellt.

Inzwischen ist jedoch nicht mehr jede Person der Zielgruppe über einen Festnetzanschluss zu erreichen. Stattdessen steigt die Zahl an Personen, die lediglich ein Mobiltelefon besitzen. Sie werden als Mobile-onlys bezeichnet. Diese Mobile-only-Haushalte sind nicht gleichmäßig über die Bevölkerung verteilt, sondern sind im Mittel jünger und häufiger erwerbstätig, leben überdurchschnittlich in Einpersonenhaushalten und verfügen über ein geringeres Nettoeinkommen (Heinz 2014). Nun könnte man die Umfrage statt über den Festnetzanschluss komplett über Mobilfunk führen. Eine reine Erhebung über Mobilfunk ohne Berücksichtigung von Festnetzanschlüssen würde jedoch auch zu Verzerrungen führen, da es immer noch einen gewissen Anteil an sogenannten Landline-onlys gibt, die also nur über das Festnetz erreichbar sind. Sie sind im Schnitt älter und leben in Westdeutschland und in ländlichen Regionen (Häder/Sand 2019: 61 f.). Aus diesem Grund werden heutzutage in den meisten Telefonumfragen in einem sogenannten Dual-Frame-Ansatz sowohl Festnetz- als auch Mobilfunknummern einbezogen (Häder/Gabler 2006; Hüfken 2019: 760). Da es für Mobilfunknummern keinerlei

---

<sup>7</sup> Allgemein zu telefonischen Umfragen siehe Hüfken (2019) und Schnell (2019: 257-275).

Verzeichnisse gibt, kommt bei deren Auswahl ein sogenanntes Random-Digit-Dialing (RDD) zur Anwendung, d.h. für die bekannten Vorwahlbereiche werden alle möglichen Nummernkombinationen generiert, woraus dann die Stichprobe gezogen wird (Häder/Sand 2019: 62).

## Vorteile von telefonischen Umfragen

Einer der wichtigsten Vorteile telefonischer Umfragen liegt in der beschriebenen Möglichkeit für eine zuverlässige und belastbare Zufallsstichprobe. Sie ist ein wichtiger Baustein für repräsentative Ergebnisse. Daneben gibt es aber noch weitere Vorteile, vor allem im Vergleich zu Face-to-face-Umfragen.

Die Durchführung von Telefonumfragen ist kostengünstiger und schneller als persönlich-mündliche Befragungen. Des Weiteren sind Sonderstichproben bestimmter Zielgruppen leicht durchzuführen (Hüfken 2019: 761), da ein Screening (z.B. von Befragten zwischen 18 und 30 Jahren) mit vertretbarem Aufwand durchführbar ist. Zudem sind Telefonumfragen deutschlandweit ohne regionale Klumpungen möglich, was sich bei persönlichen Befragungen kaum realisieren lässt (Häder 2015: 161). Darüber hinaus fällt der Effekt der sozialen Erwünschtheit geringer aus als bei Face-to-face-Interviews, da der Interviewer dem Befragten nicht gegenüber sitzt während der Befragung. Auch Interviewereffekte sind bei Telefonumfragen meist geringer als bei Face-to-face-Interviews (von Hermanni 2019b: 197 ff.). Zusätzlich fällt die Wahrscheinlichkeit von Interviewfälschungen bei telefonischen Umfragen deutlich geringer aus, da die Interviewer während der gesamten Feldphase besser kontrolliert werden können, z.B. durch das Mithören einzelner Interviews oder durch die Verwendung von Computersoftware, die die Eingabe in den Fragebogen erst ermöglicht, wenn auch ein Gespräch stattfindet. Im Gegensatz zu Face-to-face-Befragungen sind bei telefonischen Umfragen problemlos mehrere Kontaktversuche sowie eine Unterbrechung des Interviews und Fortführung zu einem späteren Zeitpunkt möglich (Häder 2015: 254 ff.).

## Nachteile von telefonischen Umfragen

Telefonbefragungen weisen neben den vielen Vorteilen aber auch Nachteile im Vergleich mit Face-to-face- oder Online-Umfragen auf. Dazu gehört eine begrenzte Interviewdauer (Hüfken 2019: 762). Face-to-face-Umfragen dauern häufig 60 Minuten pro Befragtem, bei telefonischen Umfragen sind 20 bis 30 Minuten üblich. Im Gegensatz zu persönlichen oder Online-Umfragen können telefonisch keine Bildvorlagen oder eine größere Anzahl von Antwortmöglichkeiten genutzt werden (Hüfken 2019: 762). Zudem kommt es am Telefon häufiger zu Interviewabbrüchen, als dies im persönlichen Gespräch der Fall ist, wo der Interviewer gebeten werden muss, die Wohnung zu verlassen. Bei Rückerinnerungs- oder offenen Fragen ist die Antwortqualität zum Teil geringer und telefonische Befragungen sind langweiliger für die Befragten als Face-to-face-Umfragen (Häder 2015: 255 f.).

Zu den Vorteilen von Telefonumfragen gehört die geringere soziale Erwünschtheit im Vergleich zu persönlichen Interviews. Im Vergleich mit Online-Umfragen wiederum fallen Effekte der sozialen Erwünschtheit stärker aus. Auch Interviewereffekte sind bei Telefonumfragen größer als bei Online-Umfragen, die ohne Interviewer auskommen. Dem kann aber begegnet werden, indem ausreichend viele Interviewer eingesetzt werden, denn je mehr Interviews einer Umfrage ein einzelner Interviewer führt, desto stärker sind die Interviewereffekte (von Hermanni 2019b: 197 ff.).

Ein viel zitierter Nachteil von Telefonumfragen sind die sinkenden Ausschöpfungsquoten<sup>8</sup> (Aust/Schröder 2009; Baur 2006; Häder 2015). Selbst mit einem Dual-Frame-Ansatz liegen die Ausschöpfungsquoten häufig unter 20 Prozent. Dabei kommt es nicht nur zu unsystematischer, also zufälliger, sondern auch zu systematischer Verweigerung: ältere und formal niedrig gebildete Personen sowie Stadtbewohner sind in vielen Telefonumfragen unterrepräsentiert (Engel/Schmidt 2019: 390). Allerdings muss eine niedrige Ausschöpfung nicht zwangsläufig mit einer schlechteren Datenqualität einhergehen (Yeager et al. 2011: 730 f.). Entscheidend ist, die systematische Verweigerung möglichst gering zu halten. Unsystematische Verweigerung der Interviewteilnahme wirkt sich nicht negativ auf die Datenqualität aus.

## Online-Umfragen

### Vorgehen bei Online-Umfragen und Stichprobenziehung

Online-Umfragen erfreuen sich zunehmender Beliebtheit. Nicht nur professionelle Umfrageinstitute bieten Online-Umfragen an, sondern auf den verschiedensten Wegen werden Online-Umfragen erstellt bis hin zu Webseiten, die ohne oder für geringe Kosten die eigene Erstellung einer Online-Umfrage anbieten. Bei der Online-Umfrage wird ein Fragebogen von den Befragten am Computer ohne Hilfe eines Interviewers beantwortet.<sup>9</sup>

Online-Umfrage ist nicht gleich Online-Umfrage. Ganz im Gegenteil: Inzwischen gibt es eine Vielzahl an unterschiedlichen Arten von Online-Befragungen, die sich vor allem darin unterscheiden, wie sie die Teilnehmer der Umfrage rekrutieren.

*Convenience samples* laden über Banner auf einer Homepage zur Teilnahme an der Umfrage ein. Diese Art der Umfrage ist jedoch mit bestimmten Problemen behaftet. Nicht nur unterscheiden sich Internetnutzer systematisch von Nichtnutzern, sondern zusätzlich unterscheiden sich Nutzer einer Internetseite systematisch von Nutzern anderer Seiten. Zudem entscheiden die Teilnehmer selbst über die Teilnahme, können mehrmals teilnehmen und unterscheiden sich systematisch von Nicht-Teilnehmern. Daher kommen einige Autoren zu dem Urteil, „[s]olche Befragungen [...] [seien] in keiner Weise repräsentativ und dienen allenfalls zu Unterhaltungszwecken“ (Wagner-Schelewsky/Hering 2019: 791). Mancher sieht in den „verharmlosend auch als *convenience sample* bezeichneten Umfragen [...] eine echte Bedrohung für die seriöse Umfrageforschung“ (Maurer/Jandura 2009: 63; genauso Schnell 2019: 303; Langer 2018).

Bei Intercept-Befragungen beantworten Besucher einer Homepage Fragen, wobei jeder n-te Nutzer zur Befragung eingeladen wird. Es werden damit nur Nutzer der Homepage innerhalb eines bestimmten Zeitraumes erfasst und Nutzer, die die Seite häufiger besuchen, werden mit höherer Wahrscheinlichkeit ausgesucht. Zudem sind die Teilnahmequoten sehr gering (Maurer/Jandura 2009: 63 f.).

---

<sup>8</sup> Für Telefonumfragen wird zum Beginn der Feldzeit eine ausreichend große Bruttostichprobe an Telefonnummern gezogen. Hiervon fällt ein großer Teil der Nummern weg, da sie nicht geschaltet sind, nach zehn Kontaktversuchen nicht erreicht wurden, es sich um Faxnummern handelt, es keine Privathaushalte sind oder mit der Zielperson keine Verständigung möglich ist. In einer Umfrage der Konrad-Adenauer-Stiftung aus dem Jahr 2019 fielen 82 Prozent der Telefonnummern aus der Bruttostichprobe in diese Kategorie. Die restlichen Telefonnummern bilden dann die Nettostichprobe, aus der die Interviews geführt werden und die als Basis für die Berechnung der Ausschöpfungsquote dient.

<sup>9</sup> Zu Online-Umfragen allgemein siehe Couper/Coutts (2004), Schnell (2019: 277-303) und Wagner-Schelewsky/Hering (2019).

Eine weitere Variante ist das *river sampling*, das 1995 in den USA aufkam (Callegaro et al. 2014a: 5). Ähnlich wie bei *convenience samples* wird über verschiedene Homepages für die Teilnahme an einer Umfrage geworben. Die Teilnehmer müssen zusätzlich bestimmte Informationen angeben, die es ermöglichen, Zielgruppen zu identifizieren. Auf Basis dieser Informationen wird ihnen der entsprechende Fragebogen zugeteilt. Zusätzlich können diese Befragten am Ende des Fragebogens für ein Online-Access-Panel (s. unten) rekrutiert werden. *River sampling* gehört wie auch die bisher genannten Online-Umfragen zu den nicht-probabilistischen, also nicht zufallsbasierten, sondern selbstrekrutierten Auswahlverfahren (Callegaro et al. 2014a: 7). Der Berufsverband Deutscher Markt- und Sozialforscher e.V. (BVM) weist darauf hin, dass die Verzerrungen durch *river-sampling* „durch eine wie auch immer geartete Gewichtung statistisch kaum ausgeglichen werden können“ und die Ergebnisse daher „nicht ‚repräsentativ‘ für die (z.B. deutsche) Bevölkerung“ sind (BVM 2020; ebenso z. B. Couper/Coutts 2004: 230; Schnell 2019: 299).

Eine andere Art der Online-Umfrage sind Befragungen einer bekannten und begrenzten Zielgruppe, deren E-Mail-Adressen vollständig vorliegen, z.B. Mitarbeiter einer Organisation, Studenten einer Hochschule (Wagner-Schelewsky/Hering 2019: 791; Maurer/Jandura 2009: 65). Da hier die Zielgruppe nicht die Allgemeinbevölkerung ist, sondern eine klar begrenzte Gruppe, deren Adressen vorhanden sind, kann eine Zufallsstichprobe gezogen oder sogar die gesamte Zielgruppe zur Befragung eingeladen werden. Diese Umfragen sind ein Sonderfall und nicht mit den anderen Arten der Online-Umfrage vergleichbar. Daher werden sie im Folgenden nicht weiter betrachtet, sondern nur Online-Umfragen, bei denen zumindest teilweise Rückschlüsse auf die Allgemeinbevölkerung gezogen werden sollen.

Online-Access-Panels arbeiten in der Regel mit einer Selbstrekrutierung der Teilnehmer durch offene Einladungen (Wagner-Schelewsky/Hering 2019: 791). Hier melden sich die Teilnehmer jedoch für das gesamte Access-Panel an und nehmen wiederholt an verschiedenen Umfragen teil. Solche Access-Panels sind nur geeignet für spezielle Fragestellungen wie Pretest, explorative Untersuchungen oder Testungen psychologischer Modelle, beispielsweise zur Aufmerksamkeitsmessung (Häder 2015: 167). „In der Regel wird nicht davon ausgegangen, dass Access-Panels dazu geeignet sind, Befragungen durchzuführen, deren Ergebnisse auf die Allgemeinbevölkerung übertragen werden können“ (Häder 2015: 297).

Eine Variante des Online-Panels sind Invitation-only-Panels, die ihre Teilnehmer über vorhandene Listen rekrutieren, z.B. Vielflieger- oder andere Bonusprogramme. Diese Panels weisen keine Selbstrekrutierung auf, sondern ziehen eine Zufallsstichprobe aus der Liste. Teilnehmen darf nur, wer dazu eingeladen wurde. Diese Panels erlauben Rückschlüsse auf die Gesamtheit der Liste, also beispielsweise alle Nutzer des Bonusprogrammes, aber keine Rückschlüsse auf die Allgemeinbevölkerung (Callegaro et al. 2014a: 8).

Möglich ist aber auch eine Offline-Rekrutierung per Zufallsstichprobe in der Allgemeinbevölkerung (Häder 2015: 293). Hierzu werden Internetnutzer per Zufallsstichprobe ausgesucht und telefonisch, schriftlich oder persönlich-mündlich zur Teilnahme an einem Online-Panel eingeladen (Wagner-Schelewsky/Hering 2019: 792). Dann können die Ergebnisse repräsentativ für die Gruppe der Internetnutzer sein. Um auch Offliner in das Panel zu integrieren, kann Personen ohne Internetzugang die nötige Hardware bereitgestellt werden (Blom/Herzing 2017; Blom et al. 2017). Dazu werden Internetnutzer und Haushalte ohne Internetzugang telefonisch oder face-to-face nach einer Zufallsstichprobe zur Teilnahme an einem Online-Panel eingeladen. Haushalten ohne Internetzugang wird für die Teilnahme ein Internetzugang sowie ein Tablet oder Computer kostenfrei zur Verfügung gestellt. Dies bedeutet aber einen hohen Zeit- und Kostenaufwand, der von den meisten kommerziellen Anbietern nicht betrieben wird, sondern überwiegend von wissenschaftlichen Instituten, deren Online-Panel nur der wissenschaftlichen Forschung zur Verfügung stehen.

Eine weitere Variante ist ein sogenanntes Mixed-Mode-Verfahren, bei dem die Teilnahmemethode freiwillig ausgewählt werden kann. Hierbei haben die Befragten die Wahl, ob sie die Umfrage online, postalisch oder telefonisch durchführen wollen. Eine solche Kombination verschiedener Umfragemethoden ist jedoch mit einem hohen Kostenaufwand verbunden (Wagner-Schelewsky/Hering 2019: 792) und führt zu Methodeneffekten, die die Ergebnisse beeinträchtigen.

## Vorteile von Online-Umfragen

Online-Umfragen erfreuen sich deshalb großer Beliebtheit, weil sie eine Reihe von Vorteilen gegenüber anderen Umfragemethoden mitsichbringen. Anders als Face-to-face-Umfragen sind sie zeitgleich deutschlandweit durchführbar ohne regionale Klumpungen. Gleichzeitig können aber im Gegensatz zu Telefonbefragungen Bilder, multimediale und komplexe Erhebungsinstrumente genutzt werden. Zudem reicht meist eine kurze Feldzeit, aufgrund der fehlenden Interviewer gibt es auch keine Interviewereffekte und nur geringe bis keine soziale Erwünschtheit. Darüber hinaus sind die Daten sofort verfügbar. Vor allem aber sind die Erhebungskosten gering, da keine Interviewer benötigt werden (Wagner-Schelewsky/Hering 2019: 788 f.; Bandilla 2015: 1). Access-Panels bieten zusätzlich eine schnelle und kostengünstige Auswahl einer klar definierten Zielgruppe (z.B. Brillenträger) (Häder 2015: 167).

## Nachteile von Online-Umfragen

Allerdings haben Online-Umfragen nicht nur Vor-, sondern auch Nachteile, einer davon ist die fehlende Abdeckung von Haushalten ohne Internetzugang. Inzwischen verfügen zwar 95 Prozent der deutschen Haushalte über einen Internetanschluss (Eurostat 2020). Die grundsätzliche Verfügbarkeit eines Anschlusses bedeutet aber noch nicht, dass dieser auch von allen Haushaltsmitgliedern gleichermaßen genutzt wird. Laut Initiative D21 nutzen 86 Prozent der deutschen Bevölkerung das Internet (Initiative D21 2020: 12). 14 Prozent werden hingegen zu den Offlinern gezählt und können mit Online-Umfragen grundsätzlich nicht erreicht werden. Je nach Bevölkerungsgruppe liegt der Anteil an Onlinern deutlich niedriger und damit der Anteil an Offlinern entsprechend höher. Unter den formal niedrig Gebildeten nutzen nur 64 Prozent das Internet, unter den Einpersonenhaushalten sind es 70 Prozent und unter den über 70-Jährigen 52 Prozent (gegenüber 99 Prozent der 20- bis 39-Jährigen). Zudem nutzen immer noch weniger Frauen (82 Prozent) als Männer (90 Prozent) das Internet (Initiative D21 2020: 14). Damit unterscheiden sich Offliner systematisch von Onlinern. Auch eine hohe Fallzahl an Befragten kann das Problem nicht ausgleichen, ganz im Gegenteil: Der Einfluss von systematischen Fehlern, wie sie z.B. durch die Nichtberücksichtigung von Offlinern entstehen, werden mit zunehmender Fallzahl sogar größer (von Hermann 2019a: 21).

Darüber hinaus sind „professionelle Befragte“ in Access-Panels ein Problem, also Befragte, die sehr häufig an Umfragen teilnehmen. In einer österreichischen Studie zu den Motiven von Panelteilnehmern ist jeder zweite Teilnehmer Mitglied in mindestens einem weiteren Panel, jeder Zehnte nimmt sogar an drei oder mehr Access-Panels teil (Keusch et al. 2014: 180). Hinzu kommt, dass Mitglieder mit mehreren Panelmitgliedschaften häufiger an den Umfragen des Panels teilnehmen (Keusch et al. 2014: 182).

Zusätzlich sind vor allem Access-Panels auch von geringen Ausschöpfungsquoten betroffen. Die Ausschöpfung der Ausgangsstichprobe liegt in Online-Access-Panels teilweise bei unter 5 Prozent (Maurer/Jandura 2009: 66). Für selbstrekrutierte Online-Panels kann wiederum keine Ausschöpfungsquote berechnet werden, stattdessen werden Teilnahmequoten (Anteil der Umfrageteilnehmer an allen Panelmitgliedern, die zu der Umfrage eingeladen wurden) angegeben, die sich allerdings häufig im einstelligen Bereich bewegen (Baker et al. 2013: 21 f.).

Hinzu kommt gewissermaßen die Kehrseite der geringen sozialen Erwünschtheit, die sogenannte soziale Entkontextualisierung. Die Anonymität der Umfrage führt hier zu Verzerrungen bei Fragen des sozialen Alltags, z.B. in Bezug auf Einstellungen zu sozialen Normen (Wagner-Schelewsky/Hering 2019: 789; Taddicken 2009: 100 ff.).

Schließlich kann es bei Umfragen mit einer Selbstrekrutierung zum Versuch einer bewussten Manipulation der Umfrageergebnisse kommen. So ist gerade bei politischen Umfragen denkbar, dass eine interessierte Gruppe in bestimmten weltanschaulichen Kreisen zur Teilnahme mobilisiert, um das abgebildete Meinungsbild in ihre Richtung zu beeinflussen.<sup>10</sup> Bei großen Online-Access-Panels, die regelmäßig auch zu politischen Fragen ein Stimmungsbild erheben, sind längerfristig angelegte Manipulationsversuche ebenfalls denkbar.

## Online-Access-Panels und Repräsentativität

Laut Umfrage des ADM unter seinen Mitgliedern rekrutierten 77 Prozent der Online-Umfragen im Jahr 2018 ihre Teilnehmer über Access-Panels, 6 Prozent Onsite, 14 Prozent aus Kundenlisten und lediglich 2 Prozent offline (ADM 2020). Damit spielen Access-Panels bei Online-Umfragen mit Abstand die größte Rolle und sollen an dieser Stelle noch einmal ausführlicher diskutiert werden.

Für die meisten Online-Access-Panels können sich die Teilnehmer aus eigenem Antrieb registrieren. Innerhalb des Panels werden die Teilnehmer dann häufig durch ein Zufallsprinzip oder eine Quotenstichprobe für eine konkrete Umfrage eingeladen. Dies kann jedoch nicht ausgleichen, dass die Rekrutierung für das Panel „durch offene Einladungen [erfolgt] und [...] daher hoch selektiv“ (Wagner-Schelewsky/ Hering 2019: 791) ist. Der ADM schreibt dazu in seinen Standards zur Qualitätssicherung für Online-Befragungen: „Die Selbstselektion der Teilnehmer macht passiv rekrutierte Panels für repräsentative Befragungen ungeeignet“ (ADM 2001: 7). Und auch in der gemeinsamen Richtlinie für Online-Befragungen der einschlägigen Berufsverbände ADM, ASI, BVM und DGOF<sup>11</sup> steht: „Die Selbstrekrutierung der Befragten ist häufig nicht geeignet, verallgemeinerungsfähige Untersuchungsergebnisse zu gewährleisten“ (Rat der Deutschen Marktforschung 2007: 1).<sup>12</sup>

Häufig versuchen Anbieter von Access-Panels die durch die Rekrutierung entstandenen Verzerrungen mithilfe einer Gewichtung auszugleichen. Studien können jedoch zeigen, dass gewichtete Zufallsstichproben zuverlässiger sind als gewichtete Umfragen ohne Zufallsstichprobe und der Verteilung in der Bevölkerung deutlich näherkommen (Cornesse et al. 2020: 14 ff., Yeager et al. 2011: 718 ff.), selbst wenn die Ausschöpfungsquoten der Zufallsstichproben nicht besonders hoch sind (Yeager et al. 2011: 737). Die gewichteten, nicht zufallsbasierten Stichproben sind dagegen teilweise „strikingly inaccurate, regardless of their completion rates“ (Yeager et al. 2011: 737). Es gibt zwar inzwischen verschiedene Gewichtungsverfahren, für drei dieser Verfahren (Raking, Propensity-Gewichtung und Matching) kann jedoch gezeigt werden, dass sie die Verzerrungen nur um maximal 30 Prozent verringern (Mercer et al. 2018: 3). Auch wenn Zufallsstichproben nicht frei von Messfehlern

---

<sup>10</sup> Für eine Abstimmung über den zu wiederholenden „Tatort“ wurde beispielsweise eine Manipulation berichtet (Spiegel online 2020). Auch für andere Fälle dürfte es Manipulationen oder strategische Beeinflussungen von selbstrekrutierten Abstimmungen gegeben haben.

<sup>11</sup> ASI: Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V.; BVM: Berufsverband Deutscher Markt- und Sozialforscher e.V.; DGOF: Deutsche Gesellschaft für Online-Forschung e.V.

<sup>12</sup> Zu demselben Urteil kommt die American Association for Public Opinion Research (2010): „Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values.“

sind, können in diesem Fall die Verzerrungen durch Gewichtung erfolgreich ausgeglichen werden (Cornesse et al. 2020: 21).

Es gibt verschiedene Gründe dafür, dass die Gewichtung von Online-Panels nicht erfolgreich ist. Ein wichtiger Grund liegt in der Selbstselektion vieler Online-Panels, die dazu führt, dass kein Auswahlrahmen für die Rekrutierung der Panelteilnehmer existiert, wodurch keine Auswahlwahrscheinlichkeiten angegeben werden können.

Alle Formen der Anpassung setzen voraus, dass für alle Personen der Population, über die Aussagen getroffen werden sollen, eine Teilnahmewahrscheinlichkeit größer null vorliegt. Für Personen ohne Internetzugang beträgt die Teilnahmewahrscheinlichkeit jedoch null, womit eine wichtige Voraussetzung für die Gewichtung und anschließende Aussage über die Gesamtbevölkerung verletzt ist (Valliant 2019: 8). Auch Aussagen über Personen mit Internetzugang können nur sehr eingeschränkt getroffen werden, da sich durch die Selbstselektion die Teilnehmer von Online-Panels systematisch von anderen Internetnutzern unterscheiden (Callegaro et al. 2014b: 46 f.).

Die Gewichtung anhand soziodemographischer Merkmale wie Geschlecht, Alter und Bildung setzt zudem voraus, dass es keine Selektionseffekte innerhalb dieser Gruppen gibt (Cornesse et al. 2020: 12). Es werden also beispielsweise Unterschiede zwischen Männern und Frauen, aber nicht innerhalb dieser Gruppen in der Gewichtung berücksichtigt. Das unterstellt die Annahme, dass alle Männer dieselbe Wahrscheinlichkeit haben, an der Online-Umfrage teilzunehmen und ebenso alle Frauen. Es ist aber wahrscheinlich, dass teilnehmende Männer in bestimmten Bereichen andere Einstellungen aufweisen als nicht teilnehmende Männer.<sup>13</sup>

Eine weitere Variante besteht darin, parallel zur Online-Erhebung eine Befragung mit Zufallsauswahl durchzuführen. Diese Daten dienen dann zur Gewichtung der Online-Umfrage (Maurer/Jandura 2009: 63). Allerdings relativiert dieses Vorgehen die zeit- und kostenökonomischen Vorteile von Online-Befragungen. Erfolgreich ist das Vorgehen allerdings nur dann, wenn die zum Vergleich durchgeführte Befragung mit Zufallsauswahl die Variablen enthält, mit denen die Verfügbarkeit eines Internetanschlusses und die Teilnahmewahrscheinlichkeit an der betreffenden Online-Umfrage aufklären kann. Dies ist anscheinend bislang durchweg nicht der Fall (Schnell 2019: 301). Zudem ist fraglich, inwiefern eine solche Gewichtung erfolgreich sein kann, da Gewichtungen mit den vorhandenen Befragten arbeiten. Nicht befragte Personen, die sich systematisch von den befragten Personen unterscheiden, können nicht durch Gewichtung in den Datensatz aufgenommen werden (Baur/Florian 2009: 114). Wenn also beispielsweise Personen über 70 Jahre nicht in der Stichprobe enthalten sind, kann für sie kein Gewicht erstellt werden. Gleiches gilt für Personen ohne Internetzugang.

Selbst ohne Selbstselektion, also bei Online-Umfragen mit einer Zufallsstichprobe, gibt es deutliche Verzerrungen durch die Nichtberücksichtigung von Personen, die nicht an Online-Umfragen teilnehmen können, weil sie kein Internet haben (Blom et al. 2017; Schaurer/Weiß 2020). Das liegt daran, dass sich Personen ohne Internet deutlich von Personen mit Internet unterscheiden (s. oben). „Stichproben aus der Auswahlgesamtheit der Internetnutzer können deshalb keine Repräsentativität für die Zielgesamtheit der Gesamtbevölkerung beanspruchen“ (Wagner-Schelewsky/Hering 2019: 791). Dem kann theoretisch durch eine Offline-Rekrutierung begegnet werden, bei der die ausgewählten Personen mit den technischen Mitteln versehen werden, um an einer Online-Umfrage teilzunehmen. Allerdings

---

<sup>13</sup> Das Problem wiederholt sich in gleicher Weise mit der unterschiedlichen Teilnahmechance zwischen Menschen, die häufig das Internet nutzen, und Menschen, die selten online sind. Beide Gruppen sind zwar prinzipiell für Online-Umfragen erreichbar, haben aber eine sehr unterschiedliche Wahrscheinlichkeit der Teilnahme.

zeigt sich auch hier, dass Personen mit Internetzugang eher an einer Online-Umfrage teilnehmen als Personen ohne Internetzugang, selbst wenn Letzteren die technischen Mittel dazu bereitgestellt werden. Besonders gering fällt die Teilnahme unter älteren Offlinern aus (Blom et al. 2017: 507 ff.). Dennoch kann auf diese Art die Repräsentativität einer Online-Umfrage, die auf einer Zufallsstichprobe basiert, zumindest bezüglich Alter, Geschlecht, Bildung und Haushaltsgröße, verbessert werden (Blom et al. 2017: 509 ff.).

## Fazit

Das Vorgehen bei Umfragen ist entscheidend für die Qualität und Gültigkeit der Ergebnisse. Vor allem die Auswahl der befragten Personen ist von überragender Bedeutung für die Frage, ob die erhobenen Informationen tatsächlich Rückschlüsse auf die Grundgesamtheit (in der Regel Gesamtbevölkerung) zulassen, also repräsentativ sind. Es ist die Qualität der Auswahl und nicht die Anzahl der Befragten, die über die Repräsentativität von Umfrageergebnissen entscheidet.

Die Beurteilung, ob Ergebnisse tatsächlich repräsentativ sind oder dies nur behaupten, bedarf einiger Informationen über das Vorgehen. In der Presseberichterstattung sind diese Informationen oft nur zum Teil verfügbar, die Umfrageinstitute informieren aber zum Teil durchaus über Details ihres Vorgehens und können so eine Einschätzung ermöglichen.

Umfrageergebnisse können eine wichtige Information für politische Entscheidungen sein. Umso wichtiger ist, dass es korrekte und belastbare Ergebnisse sind, die in die Entscheidungen eingehen. Die Qualität von Umfragen und die Passung des Vorgehens zu den benötigten Informationen sind wesentlich für die Frage, ob korrekte und belastbare Ergebnisse vorliegen. Eine sorgfältige Prüfung dieser Frage sollte Teil der Beauftragung von Umfragen und der Rezeption ihrer Ergebnisse sein.

## Literatur

ADM e.V., 2001: Standards zur Qualitätssicherung für Online-Befragungen. <https://www.adm-ev.de/wp-content/uploads/2018/07/Standards-zur-Qualit%C3%A4tssicherung-bei-Online-Befragungen.pdf> [Zugriff am 31.07.2020].

ADM e.V., 2020: Die Marktforschung in Zahlen. <https://www.adm-ev.de/die-branche/mafo-zahlen/> [Zugriff am 15.07.2020].

American Association for Public Opinion Research, 2010: Report on Online Panels. <https://www.aapor.org/Education-Resources/Reports/Report-on-Online-Panels.aspx> [Zugriff am 24.09.2020].

Aust, Folkert/Schröder, Helmut, 2009: Sinkende Stichprobenausschöpfung in der Umfrageforschung. Ein Bericht aus der Praxis. In: Weichbold, Martin/Bacher, Johann/Wolf, Christof (Hrsg.): Umfrageforschung. Herausforderungen und Grenzen. Österreichische Zeitschrift für Soziologie. Sonderheft 9. Wiesbaden: Springer VS, S. 195-212.

Baker, Reg/Brick, Michael J./Bates, Nancy A./Battaglia, Mike/Couper, Mick P./Dever, Jill A./Gile, Krista J./Tourangeau, Roger, 2013: Report of the AAPOR Task Force on Non-Probability Sampling. [https://www.aapor.org/AAPOR\\_Main/media/MainSiteFiles/NPS\\_TF\\_Report\\_Final\\_7\\_revised\\_FNL\\_6\\_22\\_13.pdf](https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/NPS_TF_Report_Final_7_revised_FNL_6_22_13.pdf) [Zugriff am 20.08.2020].

Bandilla, Wolfgang, 2015: Online-Befragungen. GESIS Survey Guidelines. Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften.

Baur, Nina, 2006: Ausfallgründe bei zufallsgenerierten Telefonstichproben am Beispiel des Gabler-Häder-Designs. In: Faulbaum, Frank/Wolf, Christof (Hrsg.), Stichprobenqualität in Bevölkerungsumfragen. Bonn: Informationszentrum Sozialwissenschaften, S. 159-183.

Baur, Nina/Florian, Michael J., 2009: Stichprobenprobleme bei Online-Umfragen. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.), Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: Springer VS, S. 109-128.

Blom, Annelies G./Herzing, Jessica M.E./Cornesse, Carina/Sakshaug, Joseph W./Krieger, Ulrich/Bossert, Dayana, 2017: Does the Recruitment of Offline Households Increase the Sample Representativeness of Probability-Based Online Panels? Evidence From the German Internet Panel. *Social Science Computer Review* 35 (4): 498-520.

Blom, Annelies G./Herzin, Jessica M. E., 2017: Face-to-face Rekrutierung für ein probabilistisches Onlinepanel. Einfluss auf die Repräsentativität. In: Eifler, Stefanie/Faulbaum, Frank (Hrsg.). Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung. Wiesbaden: Springer VS.

BVM e.V., 2020: River Sampling – valide Basis für Markt- und Sozialforschung? <https://www.bvm.org/themen-positionen/detailseite-news/river-sampling-valide-basis-fuer-markt-und-sozialforschung/> [Zugriff am 27.08.2020].

Callegaro, Mario/Baker, Reg/Bethlehem, Jelke/Görizt, Anja S./Krosnick, Jon A./Lavrakas, Paul J., 2014a: Online Panel Research. History, concepts, applications and a look at the future. In: Callegaro, Mario/Baker, Reg/Bethlehem, Jelke/Görizt, Anja S./Krosnick, Jon A./Lavrakas, Paul J. (Hrsg.), Online Panel Research. A Data Quality Perspective. Chichester: Wiley, S. 1-22.

Callegaro, Mario/Villar, Ana/Yeager, David/Krosnick, Jan A., 2014b: A critical review of studies investigating the quality of data obtained with online panels based on probability and nonprobability

samples. In: Callegaro, Mario/Baker, Reg/Bethlehem, Jelke/Göritz, Anja S./Krosnick, Jon A./Lavrakas, Paul J. (Hrsg.), *Online Panel Research. A Data Quality Perspective*. Chichester: Wiley, S. 23-53.

Cornesse, Carina/Blom, Annelies G./Dutwin, David/Krosnick, Jon A./De Leeuw, Edith D./Legleye, Stéphane/Pasek, Josh/Panney, Darren/Phillips, Benjamin/Sakshaug, Joseph W./Struminskaya, Bella/Wenz, Alexander, 2020: A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research. *Journal of Survey Statistics and Methodology* 8: 4-36.

Couper, Mick P./Coutts, Elisabeth, 2004: Online-Befragung. Probleme und Chancen verschiedener Arten von Online-Erhebungen. In: Diekmann, Andreas (Hrsg.): *Methoden der Sozialforschung. Sonderheft 44 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 217-243.

Diaz-Bone, Rainer/Weischer, Christoph (Hrsg.), 2015: *Methoden-Lexikon für die Sozialwissenschaften*. Wiesbaden: Springer VS.

Engel, Uwe/Schmidt, Björn Oliver, 2019: Unit- und Item-Nonresponse. In: Baur, Nina/Blasius, Jörg (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 385-404.

Eurostat, 2020: Haushalte – Internet-Zugangsdichte.

[https://appsso.eurostat.ec.europa.eu/nui/show.do?query=BOOKMARK\\_DS-053754\\_QID-5E887D4A\\_UID-3F171EB0&layout=TIME,C,X,0;GEO,L,Y,0;UNIT,L,Z,0;HHTYP,L,Z,1;INDICATORS,C,Z,2;&zSelection=DS-053754INDICATORS,OBS\\_FLAG;DS-053754HHTYP,TOTAL;DS-053754UNIT,PC\\_HH;&rankName1=HHTYP\\_1\\_2\\_-1\\_2&rankName2=UNIT\\_1\\_2\\_-1\\_2&rankName3=INDICATORS\\_1\\_2\\_-1\\_2&rankName4=TIME\\_1\\_0\\_0\\_0&rankName5=GEO\\_1\\_0\\_0\\_1&sortR=ASC\\_-1\\_FIRST&sortC=ASC\\_-1\\_FIRST&rStp=&cStp=&rDCh=&cDCh=&rDM=true&cDM=true&footnes=false&empty=false&wai=false&time\\_mode=NONE&time\\_most\\_recent=false&lang=DE&cfo=%23%23%23.%23%23%23%2C%23%23%23](https://appsso.eurostat.ec.europa.eu/nui/show.do?query=BOOKMARK_DS-053754_QID-5E887D4A_UID-3F171EB0&layout=TIME,C,X,0;GEO,L,Y,0;UNIT,L,Z,0;HHTYP,L,Z,1;INDICATORS,C,Z,2;&zSelection=DS-053754INDICATORS,OBS_FLAG;DS-053754HHTYP,TOTAL;DS-053754UNIT,PC_HH;&rankName1=HHTYP_1_2_-1_2&rankName2=UNIT_1_2_-1_2&rankName3=INDICATORS_1_2_-1_2&rankName4=TIME_1_0_0_0&rankName5=GEO_1_0_0_1&sortR=ASC_-1_FIRST&sortC=ASC_-1_FIRST&rStp=&cStp=&rDCh=&cDCh=&rDM=true&cDM=true&footnes=false&empty=false&wai=false&time_mode=NONE&time_most_recent=false&lang=DE&cfo=%23%23%23.%23%23%23%2C%23%23%23)  
[Zugriff am 15.07.2020].

Faas, Thorsten/Schoen, Harald, 2009: Fallen Gewichte ins Gewicht? Eine Analyse am Beispiel dreier Umfragen zur Bundestagswahl 2002. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.), *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*. Wiesbaden: Springer VS, S. 145-157.

Jedinger, Alexander/Michael, Tobias, 2019: Interviewereffekte. In: Baur, Nina/Blasius, Jörg (Hrsg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 365-376.

Häder, Michael, 2015: *Empirische Sozialforschung. Eine Einführung*. 3. Auflage. Wiesbaden: Springer VS.

Häder, Sabine/Gabler, Siegfried, 2006: Neue Entwicklungen bei der Ziehung von Telefonstichproben in Deutschland. In: Faulbaum, Frank Faulbaum/Wolf, Christof (Hrsg.), *Stichprobenqualität in Bevölkerungsumfragen*. Bonn: Informationszentrum Sozialwissenschaften, S. 11-17.

Häder, Sabine/Sand, Matthias, 2019: Telefonstichproben. In: Häder, Sabine/Häder, Michael/Schmich, Patrick (Hrsg.), *Telefonumfragen in Deutschland*. Wiesbaden: Springer VS, S. 45-80.

Heinz, Luise, 2014: Festnetzsubstitution durch Mobilfunk. Entwicklung & Tendenzen der soziodemografischen Merkmale bei exklusiven Mobil- und Festnetznutzern. In: Häder, Michael/Neumann, Robert (Hrsg.), *Auswahlprobleme und Antwortverhalten bei Telefonbefragungen, Dresdner Beiträge zur Soziologie, Band 5*. Dresden: MV-Verlag, S. 65-87.

Hüfken, Volker, 2019: Telefonische Befragung. In: Baur, Nina/Blasius, Jörg (Hrsg.), Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 757-768.

Initiative D21, 2020: D21-Digital-Index 2019/2020. Jährliches Lagebild zur digitalen Gesellschaft. [https://initiatived21.de/app/uploads/2020/02/d21\\_index2019\\_2020.pdf](https://initiatived21.de/app/uploads/2020/02/d21_index2019_2020.pdf) [Zugriff am 03.11.2020].

Keusch, Florian/Batinic, Bernad/Mayerhofer, Wolfgang, 2014: Motives for joining nonprobability online panels and their association with survey participation behavior. In: Callegaro, Mario/Baker, Reg/Bethlehem, Jelke/Görizt, Anja S./Krosnick, Jon A./Lavrakas, Paul J. (Hrsg.), Online Panel Research. A Data Quality Perspective. Chichester: Wiley, S. 171-191.

Kiesl, Hans, 2019: Gewichtung. In: Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 405-412.

Kromrey, Helmut/Roose, Jochen/Strübing, Jörg, 2016: Empirische Sozialforschung. Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung. Berlin, Boston: de Gruyter.

Langer, Gary, 2018: The Importance of Probability-Based Sampling Methods for Drawing Valid Inferences. In: Vannette, David L./Krosnick, Jon A. (Hrsg.): The Palgrave Handbook of Survey Research. Cham: Palgrave Macmillan, S. 7-12.

Matschinger, Herbert/Bernet, Sebastian/Angermeyer, Matthias C., 2005: An Analysis of Interviewer Effects on Screening Questions in a Computer Assisted Personal Mental Health Interview. *Journal of Official Statistics* 21(4): 657-674.

Maurer, Marcus/Jandura, Olaf, 2009: Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.), Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: Springer VS, S. 61-73.

Mercer, Andrew/Lau, Arnold/Kennedy, Courtney, 2018: For Weighting Online Opt-In Samples, What Matters Most? Pew Research Center.

Porta, Miquel, 2014: A Dictionary of Epidemiology. Oxford: Oxford University Press.

Rat der Deutschen Marktforschung e.V., 2007: Richtlinie für Online-Befragungen. [http://rat-marktforschung.de/fileadmin/user\\_upload/pdf/R08\\_RDMS.pdf](http://rat-marktforschung.de/fileadmin/user_upload/pdf/R08_RDMS.pdf) [Zugriff am 31.07.2020].

Schaurer, Ines/Weiß, Bernd, 2020: Investigating selection-bias of online surveys on coronavirus-related behavioral outcomes. *Survey Research Methods* 14 (2): 103-108.

Schnell, Rainer, 2019: Survey-Interviews. Methoden standardisierter Befragungen. 2. Auflage. Wiesbaden: Springer VS.

Schnell, Rainer, 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.

Scott, John, 2014: A Dictionary of Sociology. Oxford: Oxford University Press.

Sodeur, Wolfgang, 2007: Entscheidungsspielräume von Interviewern bei der Wahrscheinlichkeitsauswahl: ein Vergleich von ALLBUS-Erhebungen. In: *Methoden Daten Analysen*, Jg. 1, Heft 2, S. 107-130.

Spiegel online, 2020: Voting zu Wunsch-"Tatort" aus Münster manipuliert. 26.06.2020. <https://www.spiegel.de/kultur/tv/tatort-voting-fuer-wunschfolge-auf-muenster-wurde-manipuliert-a-ba738cb8-77ee-4a1b-a2e0-23044f2226b9> [Zugriff am 24.06.2020].

Stocké, Volker, 2019: Persönlich-mündliche Befragung. In: Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 745-756.

Taddicken, Monika, 2009: Die Bedeutung von Methodeneffekten der Online-Befragung: Zusammenhänge zwischen computervermittelter Kommunikation und erreichbarer Datengüte. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.), Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: Springer VS, S. 91-107.

Valliant, Richard, 2019: Comparing Alternatives for Estimation from Nonprobability Samples. [https://www.researchgate.net/profile/Richard\\_Valliant/publication/337950671\\_COMPARING\\_ALTERNATIVES\\_FOR\\_ESTIMATION\\_FROM\\_NONPROBABILITY\\_SAMPLES/links/5df7bc15a6fdcc2837268747/COMPARING-ALTERNATIVES-FOR-ESTIMATION-FROM-NONPROBABILITY-SAMPLES.pdf?origin=publication\\_detail](https://www.researchgate.net/profile/Richard_Valliant/publication/337950671_COMPARING_ALTERNATIVES_FOR_ESTIMATION_FROM_NONPROBABILITY_SAMPLES/links/5df7bc15a6fdcc2837268747/COMPARING-ALTERNATIVES-FOR-ESTIMATION-FROM-NONPROBABILITY-SAMPLES.pdf?origin=publication_detail) [Zugriff am 03.08.2020].

von Hermanni, Hagen, 2019a: Das Total Survey Error Modell in telefonischen Befragungen. In: Häder, Sabine/Häder, Michael/Schmich, Patrick (Hrsg.), Telefonumfragen in Deutschland. Wiesbaden: Springer VS, S. 17-34.

von Hermanni, Hagen, 2019b: Rolle und Einfluss von Interviewern in telefonischen Umfragen. In: Häder, Sabine/Häder, Michael/Schmich, Patrick (Hrsg.), Telefonumfragen in Deutschland. Wiesbaden: Springer VS, S. 193-239.

Wagner-Schelewsky, Pia/Hering, Linda, 2019: Online-Befragung. In: Baur, Nina/Blasius, Jörg (Hrsg.), Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 787-800.

West, Brady T./Blom, Annelies G., 2017: Explaining Interviewer Effects. A Research Synthesis. In: *Journal of Survey Statistics and Methodology*, Jg. 5, Heft 2, S. 175-211.

Yeager, David S./Krosnick, John A./Chang, Linchiat/Javitz, Harold S./Levendusky, Matthew S./Simpser, Alberto/Wang, Rui, 2011: Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly* 75 (4): 709-747.

## Impressum

### Die Autoren

Dr. Sabine Pokorny studierte Soziologie, Englische und Italienische Philologie in Köln und Florenz. Von 2007 bis 2011 war sie wissenschaftliche Mitarbeiterin an der Freien Universität Berlin. Seit September 2011 war sie Koordinatorin für Empirische Sozialforschung in der Hauptabteilung Politik und Beratung und seit Januar 2020 ist sie Referentin in der Abteilung Wahl- und Sozialforschung in der Hauptabteilung Analyse und Beratung der Konrad-Adenauer-Stiftung e.V.

PD Dr. Jochen Roose studierte an der Freien Universität Berlin Soziologie. Nach einer Promotion als Mitarbeiter des Wissenschaftszentrums Berlin für Sozialforschung und einer Station an der Universität Leipzig habilitierte er an der FU Berlin in Soziologie. Als Professor war er an der Universität Hamburg, der FU Berlin und der Universität Wrocław (Breslau) beschäftigt, bevor er 2018 in die Hauptabteilung Politik und Beratung zur Konrad-Adenauer-Stiftung e. V., Berlin, wechselte. Seit Januar 2020 arbeitet er in der Wahl- und Sozialforschung für die Hauptabteilung Analyse und Beratung.

## Konrad-Adenauer-Stiftung e. V.

### **Dr. Sabine Pokorny**

Abteilung Wahl- und Sozialforschung  
Hauptabteilung Analyse und Beratung  
T +49 30 / 26 996-3544  
[sabine.pokorny@kas.de](mailto:sabine.pokorny@kas.de)

### **PD Dr. Jochen Roose**

Abteilung Wahl- und Sozialforschung  
Hauptabteilung Analyse und Beratung  
T +49 30 / 26 996-3798  
[jochen.roose@kas.de](mailto:jochen.roose@kas.de)

Postanschrift: Konrad-Adenauer-Stiftung e. V., 10907 Berlin

Herausgeberin: Konrad-Adenauer-Stiftung e. V. 2020, Berlin  
Gestaltung: yellow too Pasiak Horntrich GbR



Der Text dieses Werkes ist lizenziert unter den Bedingungen von „Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 4.0 international“, CC BY-SA 4.0 (abrufbar unter: <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>).