

(Um-)Fragen über (Um-)Fragen: Die Entstehung von Umfragedaten am Beispiel des European Social Survey (ESS)

Christian Schnaudt, Michael Weinhardt und Stefan Liebig

Zusammenfassung

Immer häufiger werden wir im Alltag und den Medien mit Umfragen und deren Ergebnissen konfrontiert, wissen aber oftmals nicht, wie wir deren Qualität und Vertrauenswürdigkeit einschätzen können. Dieser Aufsatz verdeutlicht am Beispiel des Umfrageprogramms European Social Survey (ESS), welche Standards Umfragen erfüllen sollten, um ein verlässliches Abbild der Meinungen und Einstellungen innerhalb der Bevölkerung zu liefern.

1 Einführung

Wie benutzen Sie Klopapier üblicherweise?¹ Würden Sie unsere Bundesregierung grundsätzlich als fahrradfreundlich einstufen?² Halten Sie es für gerechtfertigt, unbewaffnete Flüchtlinge mit Waffengewalt am Grenzübertritt zu hindern oder halten Sie dies nicht für gerechtfertigt?³ Von skurril bis politisch brisant – kaum ein Tag vergeht, an dem wir nicht in der Fußgängerzone, am Telefon oder im Internet um die Teilnahme an einer Umfrage gebeten oder aber in den Medien über die Ergebnisse ebensolcher informiert werden. In der Regel sollen diese Umfragen ‚repräsentativ‘, d.h. in ihren Aussagen auf eine bestimmte Zielgruppe oder gar die gesamte Bevölkerung übertragbar sein. Oftmals bleibt jedoch unklar, worauf sich die (angebliche) Repräsentativität von Umfragen gründet, also auf Grundlage welcher methodischen Standards diese erhoben und nach welchen Kriterien Befragte für selbige ausgewählt werden. Für Au-



ßenstehende ist es somit schwierig, die Qualität und Vertrauenswürdigkeit von Umfragen zu bewerten und „gute“ von „schlechten“ Umfragen zu unterscheiden. Dies erscheint problematisch, da Umfragen zum einen darauf abzielen, ein Abbild der vorhandenen Meinungen, Einstellungen und Verhaltensweisen innerhalb der Bevölkerung zu schaffen, zum anderen aber faktisch immer auch Meinungen und Einstellungen beeinflussen, indem ihre Ergebnisse in der medialen Berichterstattung präsentiert und – oftmals kontrovers – diskutiert werden (vgl. Faas 2014). Dieser doppelte Charakter von Umfragen als *meinungsabbildend* in ihrer Intention und *meinungsbildend* in ihrer Wirkung wirft die Notwendigkeit auf, die Entstehung von Umfragedaten in ihren Details näher zu beleuchten, um sich so der Implikationen bestimmter Entscheidungen im Datengenerierungsprozess und deren Auswirkungen auf die Interpretation von Umfrageergebnissen bewusst zu werden. Nur wenn wir den Entstehungsprozess von Umfragedaten nachvollziehen können, verstehen wir, inwiefern Umfragen tatsächlich die vorhandenen Meinungen und Einstellungen innerhalb der Bevölkerung abbilden und welche konkreten Schlussfolgerungen – ungeachtet der medialen Darstellung – auf Grundlage der erhobenen Daten zulässig sind.

Damit Umfragen ein verlässliches Abbild der Meinungen und Einstellungen innerhalb der Bevölkerung liefern können, müssen diese bestimmte Standards erfüllen. Im vorliegenden Beitrag soll der Entstehungsprozess von Umfragedaten näher beleuchtet und am Beispiel des internationalen Umfrageprogramms European Social Survey (ESS) verdeutlicht werden. Auf Grund seiner innovativen methodischen Standards fungiert der ESS als Vorbild und Referenzstudie in der international vergleichenden Umfrageforschung und eignet sich daher hervorragend als Anwendungsbeispiel, um die relevanten methodischen und praktischen Probleme und Zielkonflikte im Zusammenhang mit dem Datenentstehungsprozess zu illustrieren.

2 Der European Social Survey im Überblick

Der European Social Survey (ESS) ist ein europaweites sozialwissenschaftliches Umfrageprogramm, welches seit 2002 in mehr als 30 verschiedenen Ländern durchgeführt wurde. Befragungen im Rahmen des ESS finden im zweijährlichen Turnus statt, so dass bis dato Umfragedaten und -ergebnisse von insgesamt sieben Erhebungswellen für die Jahre 2002 bis 2014 vorliegen. Die achte Welle wird im Herbst und Winter 2016 erhoben. Das Hauptziel des ESS besteht darin, Stabilität und Wandel in den sozialen, politischen, wirtschaftlichen, religiösen sowie moralischen Einstellungen und Verhaltensmustern europäischer Bürger und Bürgerinnen zu erfassen, um so Auskunft über den Zustand sowie die Entwicklungen europäischer Gesellschaften zu Beginn des 21. Jahrhunderts zu geben. Neben diesem inhaltlichen Hauptanliegen hat sich der ESS zum Ziel gesetzt, existierende Standards der international vergleichenden Umfrageforschung weiter zu verbessern sowie deren Anwendung und Verbreitung aktiv zu fördern. Diese Standards beziehen sich insbesondere auf die Entwicklung von Fragebögen und -instrumenten, Verfahren der Stichprobenziehung und Datenerhebung sowie qualitätssichernde Maßnahmen in der Datenaufbereitung. Für seine Errungenschaften in diesen Bereichen wurde der ESS im Jahr 2005 von der Europäischen Kommission mit dem Descartes-Preis für Wissenschaft ausgezeichnet.

Finanzierung und Organisationsstruktur des ESS orientieren sich klar an dessen Zielsetzungen und den hiermit verbundenen Problemen. Um ein identisches Umfrageprogramm sowie einheitliche Standards der Umfrageforschung in mehr als 30 teil-

nehmenden Ländern gewährleisten und durchsetzen zu können, finanziert sich der ESS durch nichtkommerzielle bzw. öffentliche Institutionen und folgt einer hierarchischen Organisationsform mit zentraler Koordinierungsstelle, welche die Richtlinien und Standards für alle teilnehmenden Länder entwickelt und deren Umsetzung überwacht. Im Detail erfolgte die Finanzierung des ESS für die ersten sechs Erhebungswellen durch die Forschungsrahmenprogramme der EU, die European Science Foundation (ESF) sowie nationale Förderinstitutionen in den jeweiligen Ländern (z.B. die Deutsche Forschungsgemeinschaft in Deutschland). Seit 2013 hat der ESS den Status als *European Research Infrastructure Consortium* (ERIC) inne. Die Etablierung als ESS ERIC impliziert eine weitere rechtliche und finanzielle Verstärkung des Umfrageprogramms, um dessen langfristigen Fortbestand zu gewährleisten. Organisatorisch gliedert sich das ESS ERIC in die Generalversammlung bestehend aus nationalen Repräsentanten der Mitgliedsländer sowie drei weiteren Unterausschüssen (wissenschaftlicher Beirat, Methodikausschuss und Finanzausschuss). Verantwortlich für das operative Geschäft ist der Direktor des ESS, welcher durch die Generalversammlung bestimmt und durch ein Team im ESS-Hauptsitz (City University, London) sowie sechs weitere Partnerinstitutionen und Universitäten (das *Core Scientific Team*) unterstützt wird. Die letzte Implementation und Umsetzung des zentral entwickelten ESS-Umfrageprogramms in den einzelnen Ländern übernehmen nationale Teams unter der Leitung der jeweiligen nationalen Koordinatoren (vgl. Schnaudt et al. 2014, 487-8).

Auch die Inhalte des ESS folgen in klarer Weise dessen Zielsetzungen. Um die Entwicklungen europäischer Gesellschaften konsistent erfassen zu können, müssen die erhobenen Inhalte sowohl über die Zeit konstant als auch zwischen den einzelnen Ländern vergleichbar sein. Diese angestrebte und notwendige Konstanz spiegelt sich in den sog. Kernmodulen des ESS wider. Die Kernmodule umfassen Themenkomplexe, welche von permanentem wissenschaftlichen und gesellschaftlichen Interesse sind und daher in jeder Erhebungswelle des ESS abgefragt werden. Zu diesen Themenkomplexen gehören neben den soziodemographischen Merkmalen der Befragten u.a. Fragen zum politischen Interesse, zur politischen Partizipation, zum Vertrauen in Mitmenschen, Politiker und politische Institutionen, zur wirtschaftlichen und persönlichen Zufriedenheit, zur sozialen Inklusion, zur Religionszugehörigkeit sowie zu den Wertvorstellungen der Befragten. Um neben dieser konstanten Komponente dennoch auch offen für neue oder spezifischere Themen zu sein, beinhaltet der ESS ebenfalls sog. Rotationsmodule. Die Inhalte der einzelnen Rotationsmodule wechseln in der Regel im zweijährlichen Turnus mit jeder neuen Erhebungswelle, können aber auch wiederholt werden. Die Rotationsmodule der ersten sieben Erhebungswellen beinhalten Fragen zu den Themen Immigration (2002 und 2014), Bürgerschaft und Demokratie (2002), Wirtschaftsmoral (2004), Arbeit, Familie und Wohlbefinden (2004 und 2010), Gesundheit und Pflege (2004), Lebensplanung (2006), persönliches und soziales Wohlbefinden (2006 und 2012), Altersdiskriminierung (2008), Wohlfahrtsstaat (2008), Vertrauen in Justiz und Polizei (2010), Verständnis und Bewertungen der Demokratie (2012) sowie gesundheitliche Ungleichheiten (2014). Die für die achte Erhebungswelle 2016 in Planung befindlichen Rotationsmodule umfassen Fragen rund um die Themen Wohlfahrtsstaat (Wiederholung von 2008) sowie Klimawandel und Energiesicherheit (vgl. Schnaudt et al. 2014, 490-3).

Sämtliche Umfragedaten des ESS inkl. der zugehörigen Dokumentationsmaterialien und Fragebögen sind nach einer kurzen Registrierung auf der internationalen Homepage des ESS frei zum Download verfügbar (www.europeansocialsurvey.org). Zusätzlich bietet die Homepage ein praktisches Online-Analyse Tool an, mit dessen Hilfe einfache Berechnungen und Analysen (z.B. Kreuztabellen, Mittelwerte, Korrelationen

und OLS-Regressionen) sowie Grafiken auf Grundlage der ESS-Daten erstellt werden können. Insbesondere im Vergleich zu kommerziellen Umfragen stellt die öffentliche Bereitstellung der ESS-Daten eine Besonderheit dar und zeugt nicht nur von deren Status als Allgemeingut, sondern ebenfalls von der im ESS umgesetzten Transparenz, welche es allen Interessierten ermöglicht, sich selbst ein Bild von der Umfrage und deren Qualität zu verschaffen (siehe hierzu auch www.europeansocialsurvey.de).

3 Wesentliche Aspekte und Probleme beim Entstehungsprozess von Umfragedaten am Beispiel von Wahlumfragen und des European Social Survey (ESS)

3.1 Wahlumfragen

Rückblick auf die britischen Unterhauswahlen vom 7. Mai 2015: Die konservativen Tories unter Führung des designierten Premierministers David Cameron erlangen mit 330 von insgesamt 649 Sitzen eine absolute Mehrheit. Während absolute Mehrheiten im Vereinigten Königreich nur wenig außergewöhnlich anmuten und lange Zeit den Regelfall bildeten, so stellt das offizielle Wahlergebnis für viele Beobachter dennoch eine Überraschung dar. Insbesondere unter den Mitarbeitern und Verantwortlichen der Umfrageinstitute im Land dürfte das eindeutige Wahlergebnis für einiges Kopfzerbrechen und Stirnrunzeln gesorgt haben, denn noch bis kurz vor den Wahlen hatten diese ein Kopf-an-Kopf-Rennen zwischen der *Conservative Party* und der *Labour Party* prognostiziert und die jeweiligen Stimmenanteile beider Parteien auf etwa 34% beziffert. Die Zahlen im amtlichen Wahlergebnis belaufen sich letztlich auf einen Stimmenanteil von 36,8% für die Tories und 30,4% für Labour (vgl. Hawkins, Keen und Nakatudde 2015) – eine Differenz von 6,4 Prozentpunkten und somit alles andere als das von den Umfrageinstituten prognostizierte Kopf-an-Kopf-Rennen.

Eine solch gravierende Diskrepanz zwischen den Prognosen der Umfrageinstitute und den tatsächlichen Ergebnissen der britischen Unterhauswahl 2015 lässt sich mit der immer gegebenen statistischen Unsicherheit bei Umfragen allein nicht erklären. In einem 2016 erschienenen Bericht führt der Wahlforscher John Curtice diese Diskrepanz auf die mangelnde Repräsentativität der von den Umfrageinstituten durchgeführten Befragungen zurück, wobei sich insbesondere zwei Quellen mangelnder Repräsentativität identifizieren lassen: Zum einen seien Personen, die sich relativ einfach für Wahlumfragen rekrutieren lassen, seltener Wähler der Tories. Dadurch seien in den Wahlumfragen weniger Wähler der *Conservative Party* befragt worden als dies ihrem Anteil in der Gesamtbevölkerung entsprochen hätte. Zum anderen zeigten jüngere Personen eine viel geringere Wahrscheinlichkeit der Wahlbeteiligung als ältere Personen, unterstützten aber gleichzeitig mit größerer Wahrscheinlichkeit die Labour Party. Indem die meisten Wahlumfragen die Wahlbeteiligung jüngerer Personen höher einschätzten als dies tatsächlich der Fall gewesen sei, überschätzten sie den Stimmenanteil für Labour (vgl. Curtice 2016). Diese simultane Unterschätzung des Stimmenanteils der Tories und Überschätzung des Stimmenanteils für Labour führte zur Prognose eines Kopf-an-Kopf-Rennens, welches letztlich in einer absoluten Mehrheit für die Tories und einer relativen Ratlosigkeit seitens der Umfrageinstitute endete.

Das hier geschilderte Beispiel bringt uns zurück zum zentralen Anliegen dieses Beitrags und der Frage, unter welchen Bedingungen Umfragen ein repräsentatives

Abbild der in der Gesamtbevölkerung vorhandenen Meinungen und Einstellungen liefern können. Für eine systematische Beantwortung dieser Frage orientieren wir uns im Folgenden an fünf Leitfragen, mit denen sich jeder Umfragemethodiker im Verlauf der Konzeption und Durchführung einer Befragung auseinandersetzen sollte (vgl. Groves et al. 2009, 33). Wir diskutieren diese Leitfragen nacheinander, weisen auf potenzielle Probleme beim Entstehungsprozess von Umfragedaten hin und veranschaulichen mögliche Lösungsansätze an Hand der konkreten Umsetzung im ESS.

Wer wird ausgewählt und befragt? Ausgangspunkt einer jeden Umfrage bildet die Frage, wie man von den Antworten einzelner Befragter zu verlässlichen Aussagen über eine vorher definierte Grundgesamtheit oder Zielpopulation gelangt. Dies ist die Gesamtheit an Personen, für welche die Aussagen der Umfrage letztlich Gültigkeit besitzen sollen. Hierbei ist zu beachten, dass Umfragen aus Effizienz- und Kostengründen nicht die gesamte Zielpopulation, sondern immer nur einen Teil derselben, die sog. Stichprobe, befragen können. Dies ist relativ unproblematisch, sofern die Stichprobe auf einer zufälligen Auswahl der zur Zielpopulation gehörenden Mitglieder basiert, bei der jedes der Mitglieder die gleiche oder eine bekannte Wahrscheinlichkeit besitzt, für die Stichprobe ausgewählt zu werden. Nur auf Basis einer solchen Zufallsstichprobe der Zielpopulation ist es möglich, mit Hilfe statistischer Verfahren von etwa 1000 Befragten auf eine Grundgesamtheit von beispielsweise mehreren Millionen Bürgern zu schließen. Zwar weichen auch Zufallsstichproben zu einem gewissen Grad von den ‚wahren‘ Werten (z.B. dem Stimmenanteil der SPD) in der Grundgesamtheit ab. Ihr großer Vorteil besteht jedoch darin, dass die Unsicherheit der auf Zufallsstichproben basierenden Ergebnisse berechnet und ein Bereich angegeben werden kann, innerhalb dessen sich die Werte in der Grundgesamtheit wahrscheinlich befinden (das sog. Konfidenzintervall). Wird z.B. bei der Veröffentlichung von Umfrageergebnissen ein Stimmenanteil von 33% für die SPD mit einem zugehörigen Stichprobenfehler von 2% angegeben, so ist zu erwarten, dass der ‚wahre‘ Stimmenanteil in der Grundgesamtheit in einem Bereich zwischen 31 und 35% liegt (vgl. Schoen und Greszki 2014: 21-2; Jessen 2014: 27). Die Auswahl der Befragten auf Basis einer Zufallsstichprobe ist eine der essenziellen Bedingungen für die Repräsentativität einer Umfrage und deren Fähigkeit, einem Abbild der in der Grundgesamtheit vorhandenen Meinungen und Einstellungen zu entsprechen. Diesem Anspruch können Umfragen, welche auf nicht zufälligen Auswahlmechanismen beruhen, gar nicht oder nur sehr eingeschränkt genügen. Dies zeigt sich deutlich am obigen Beispiel der britischen Unterhauswahlen: Hier wurden die Befragten ausschließlich auf Basis nicht zufälliger Mechanismen (u.a. sog. Quotenstichproben) ausgewählt – die Konsequenzen sind bekannt (vgl. Curtice 2016: 4-5).

3.2 Der European Social Survey

Wie wird nun in der wissenschaftlichen Umfrageforschung – hier im ESS – eine Zufallsstichprobe der Zielpopulation realisiert? Zunächst definiert der ESS seine Zielpopulation als Gesamtheit sämtlicher Individuen, welche mindestens 15 Jahre alt und permanent in einem privaten Haushalt wohnhaft sind. Auf Grundlage dieser Definition gilt es, eine vollständige Auflistung sämtlicher Elemente der Zielpopulation zu erlangen, aus welcher dann eine Teilmenge für die Stichprobe zufällig ausgewählt werden kann (Stichprobenplan). In den meisten Ländern des ESS wird dies mit Hilfe offizieller Listen von Individuen, Haushalten oder Adressen erreicht, welche von den zuständigen Verwaltungen und öffentlichen Einrichtungen zur Verfügung gestellt werden.

Um diesen Prozess genauer zu beleuchten, gehen wir im Folgenden auf die deutsche Teilstudie des ESS ein und verdeutlichen die Stichprobenziehung hier im Detail (vgl. Schnaudt et al. 2014, 501-2). Die Auswahl der Befragten in Deutschland entspricht einer geschichteten, zweistufigen Zufallsstichprobe. Hierbei werden zunächst aus dem nach Region und Größe der Gemeinden geschichteten Gemeindeverzeichnis zufällig Gemeinden ausgewählt. Die Schichtung der Gemeinden nach Region und Größe stellt sicher, dass sich in der finalen Stichprobe Befragte aus unterschiedlich großen und geografisch unterschiedlich gelegenen Regionen befinden und hilft, die Präzision der Stichprobe zu erhöhen (vgl. Groves et al. 2009: 113ff). Anschließend wird auf Grundlage der Melderegister und Personenverzeichnisse der ausgewählten Gemeinden eine zuvor festgelegte Anzahl an Personen zufällig ausgewählt, welche in ihrer Gesamtheit die Stichprobe für die Befragung bilden.

Der Vorteil der Auswahl von Befragten auf Grundlage der hier beschriebenen Mechanismen besteht darin, dass sämtliche Elemente der Grundgesamtheit durch den Stichprobenplan abgedeckt sind, da zumindest theoretisch alle mindestens 15 Jahre alten und permanent in privaten Haushalten wohnhaften Individuen von den offiziellen Meldelisten und Personenregistern abgedeckt sind. Auf diese Weise wird im ESS das Risiko eines potenziellen ‚undercoverage‘ der Stichprobe (also des Nichterfassens von Elementen der Zielpopulation; vgl. Groves et al. 2009: 54-6) minimiert und eine größtmögliche Übereinstimmung der Stichprobe mit der definierten Zielpopulation gewährleistet. Telefon- oder Internetumfragen, wie im obigen Beispiel der britischen Unterhauswahlen angewandt (vgl. Curtice 2016: 4), können dies abermals nur eingeschränkt leisten: Hier werden sämtliche zur Zielpopulation gehörenden Personen, die nicht über einen Telefon- oder Internetanschluss verfügen, systematisch von der Befragung ausgeschlossen, was eine mögliche Verzerrung der Ergebnisse nach sich zieht.

Wie werden die ausgewählten Personen von einer Teilnahme überzeugt? Während die Auswahl der Befragten auf Basis einer Zufallsstichprobe eines der essenziellen Qualitätsmerkmale einer Umfrage darstellt, so kann diese allein die Übertragbarkeit der Ergebnisse einer Stichprobe auf die zu Grunde liegende Zielpopulation nicht garantieren. Wie am vorangegangenen Beispiel der britischen Unterhauswahlen verdeutlicht, kommt es nicht nur darauf an, die Befragten einer Umfrage adäquat auszuwählen, sondern letztlich auch die gewünschten Informationen von diesen zu erhalten. Befragten wir lediglich solche Personen, welche leicht zu rekrutieren sind oder sich für das Thema der Umfrage interessieren, so erhielten wir sehr wahrscheinlich ein verzerrtes Abbild der in der Grundgesamtheit vorhandenen Meinungen und Einstellungen, da die Teilnahmebereitschaft an Umfragen oftmals mit den uns primär interessierenden Merkmalen wie beispielsweise politischem Interesse, Bildung oder sozialem Status zusammenhängt. Ziel einer jeden Umfrage muss es also sein, Personen für die Befragung nicht nur zufällig auszuwählen, sondern auch exakt diese tatsächlich zu befragen. Nur wenn die Nichtteilnahme oder Verweigerungsrate unter den ausgewählten Befragten ausreichend gering oder zufällig verteilt ist, sind die gewonnenen Erkenntnisse von der Stichprobe auf die Grundgesamtheit übertragbar.

Im ESS wird in allen Ländern eine Rücklaufquote von 70% angestrebt. Um diese auch in der Praxis realisieren zu können, sind im ESS verschiedene Maßnahmen implementiert, welche von der Kontaktaufnahme bis zur Realisierung der Befragung die Teilnahmebereitschaft der ausgewählten Befragten erhöhen und sicherstellen sollen. Diese Maßnahmen umfassen persönliche Kontaktbriefe (bei Auswahl der Befragten auf Basis von Personenregistern), in denen die Befragten auf die bevorstehende Um-

frage hingewiesen, über deren Inhalte informiert sowie über die Relevanz ihrer Teilnahme aufgeklärt werden. Während der Erhebung der Umfrage, die im ESS durch persönliche Interviews durchgeführt wird, sind die Interviewer zudem angehalten, mindestens vier persönliche Kontaktversuche über einen Zeitraum von zwei Wochen zu unternehmen (zu verschiedenen Zeiten, an verschiedenen Tagen, einmal abends, einmal am Wochenende), sofern sie die ausgewählten Interviewpartner nicht beim ersten Kontakt antreffen oder von einer Teilnahme überzeugen können. Zusätzlich zu den genannten Maßnahmen kann eine Teilnahme am ESS durch monetäre Incentives belohnt werden. In der siebten Erhebungswelle erhielten die Befragten in Deutschland in der Regel 20 Euro Entlohnung für ihre Teilnahme (vgl. Schnaudt et al. 2014: 502-3).

Wie werden die ausgewählten Personen befragt? Wie werden die Interviewer geschult und überwacht? Sämtliche Maßnahmen zur Ziehung und Realisierung einer für die Grundgesamtheit repräsentativen Stichprobe sind bedeutungslos, wenn der Befragungsprozess selbst Defizite aufweist. Wie bereits erwähnt, findet die Befragung im ESS in allen Ländern durch persönliche Interviews der Befragten statt. Der Vorteil von persönlichen Interviews gegenüber postalischen und telefonischen Umfragen oder Internetsurveys besteht zunächst in den besseren Rekrutierungsmöglichkeiten auch schwer zu erreichender oder nichtteilnahmewilliger Personen. Zudem kann durch persönliche Interviews gewährleistet werden, dass auch tatsächlich die ausgewählten und keine anderen Personen an der Befragung teilnehmen. Im Hinblick auf die Qualität der Umfragedaten kommt den Interviewern ebenfalls eine zentrale Rolle zu: sie führen die Befragung durch, dokumentieren die Antworten der Befragten und kontrollieren diese auf Vollständigkeit. Kommen Interviewer diesen Aufgaben nur unzureichend nach, hat dies direkte Konsequenzen für die Aussagekraft der gesammelten Daten (vgl. Groves et al. 2009: 291ff). Umso wichtiger ist es, den Befragungsprozess so wenig fehleranfällig wie möglich zu gestalten und die Interviewer grundlegend für die Durchführung der Umfrage zu schulen und deren Leistungen während der Erhebungsphase kontinuierlich zu überprüfen.

Zur Umsetzung dieser Ziele existiert im ESS eine Vielzahl standardisierter Richtlinien für die Datenerhebung, welche eine gemeinsame Erhebungsphase von vier Monaten in allen Ländern (September-Dezember), eine angestrebte Interviewdauer von etwa 60 Minuten sowie die ausschließliche Nutzung standardisierter, persönlicher Interviews bei den Befragten zu Hause vorsieht. Zur Verringerung der Fehleranfälligkeit bei der Dateneingabe werden u.a. in der deutschen Stichprobe sämtliche Interviews als CAPI (Computer Assisted Personal Interviewing) durchgeführt. CAPI führt die Interviewer automatisiert durch die Befragung und ermöglicht schon während des Interviews, die getätigten Eingaben auf Vollständigkeit und Plausibilität zu überprüfen. Um den Einfluss der Interviewer während der Interviews zu minimieren und sog. Interviewereffekte möglichst auszuschließen, erhalten sämtliche Interviewer harmonisierte Schulungsmaterialien zum Interviewablauf sowie kurz vor Beginn der Erhebungsphase persönliche Briefings durch die Mitglieder der nationalen ESS-Teams. Hinzu kommt die Bedingung, dass die Befragungen im ESS nur von Interviewern durchgeführt werden dürfen, welche bereits über Erfahrung mit persönlichen Interviews auf Basis von Zufallsstichproben verfügen. Zur weiteren Gewährleistung der Interviewqualität und Vorbeugung von Interviewereffekten, dürfen in der Erhebungsphase pro Interviewer maximal 24 Interviews geführt werden. Die endgültige Qualitätssicherung und Überprüfung der Interviewerleistung wird im ESS durch die nationalen Koordinatoren übernommen. Diese überwachen in Rücksprache mit den Umfra-

geinstituten die Einhaltung der Instruktionen der Interviewerbriefings, die Interviewdauer, die Anzahl der Kontaktversuche sowie die Ausschöpfungsquoten einzelner Interviewer. Hierzu gehört auch, dass ein Teil der ausgewählten Befragten nochmals kontaktiert wird, um zu überprüfen, ob die korrekte Person identifiziert wurde, ein Interview tatsächlich stattgefunden hat und dieses nach den Vorgaben durchgeführt wurde (sog. Interviewer ‚back-checks‘). Im Falle einer Verletzung der offiziellen Vorgaben oder einer unzureichenden Leistung seitens der Interviewer findet ein Austausch selbiger durch das Umfrageinstitut statt (vgl. Schnaudt et al. 2014: 498-9).

Wie wird die Qualität der Fragen bewertet und überprüft? In unseren bisherigen Ausführungen sind wir stillschweigend davon ausgegangen, dass die Fragen einer Umfrage bereits vorliegen. Ganz im Gegenteil jedoch stellt die Entwicklung von Fragen selbst einen eigenständigen und zentralen Bestandteil der Umfrageforschung dar. Im ESS durchlaufen neu entwickelte Fragen mehrstufige Pretests, bei denen auch sog. kognitive Interviews zum Einsatz kommen. Hierbei sollen Befragte die Fragen nicht nur beantworten, sondern ebenfalls erläutern, wie sie die Fragen interpretieren und welche Probleme sie u.U. bei der Beantwortung haben, um so die Fragequalität zu verbessern. Erschwert wird dieser Prozess, wenn es sich um eine internationale Umfrage wie den ESS handelt, bei dem die Fragen gleichzeitig in mehreren Ländern und in verschiedenen Sprachen erhoben werden. Im Folgenden werden wir uns also darauf konzentrieren, auf welche Weise im ESS gewährleistet wird, dass die gestellten Fragen in jedem der teilnehmenden Länder das gleiche interessierende Merkmal erheben und die Ergebnisse somit über alle Länder hinweg vergleichbar sind.

Ausgangspunkt der Umfrageinhalte des ESS ist der in englischer Sprache verfasste Quellfragebogen, welcher von allen nationalen ESS-Teams in sämtliche Sprachen übersetzt werden muss, die von mindestens 5% der Bevölkerung im jeweiligen Land als Erstsprache gesprochen wird. Der Übersetzungsprozess folgt hierbei der sog. TRAPD-Prozedur (Translation, Review, Adjudication, Pre-Testing, Documentation; vgl. Harkness 2007). Hierbei wird der Fragebogen von verschiedenen Personen mit unterschiedlicher Expertise (Sozialwissenschaften, Umfragemethodik, Sprachwissenschaften) zunächst übersetzt und die verschiedenen Versionen werden anschließend begutachtet. Für Länder, welche über eine gemeinsame Sprache verfügen (z.B. Deutschland, Österreich und die Schweiz) ergibt sich die weitere Notwendigkeit, die jeweilig gefundenen Übersetzungen miteinander zu vergleichen, um unnötige Abweichungen zwischen den Ländern zu vermeiden. Sobald eine einheitliche Übersetzungsversion gefunden ist, wird diese in einem zusätzlichen Verifikationsprozess von einem externen Dienstleister weiteren linguistischen Überprüfungen unterzogen und falls nötig angepasst.

Ein letzter Schritt zur Überprüfung der Fragequalität im ESS besteht im sog. Survey Quality Predictor (SQP) Coding, welches die gefundenen Übersetzungen auch in ihrer formalen Struktur mit den Formulierungen des englischen Quellfragebogens vergleicht, um auch hier unnötige Abweichungen zu vermeiden. Bis dato ist der ESS das einzige internationale Umfrageprogramm, in welchem SQP Coding Anwendung findet. Nach Durchlaufen der skizzierten Schritte werden die Übersetzungsversionen in nationalen Pretests vorab getestet, gegebenenfalls notwendige Änderungen der Frageformulierungen vorgenommen und der komplette Übersetzungsprozess von den nationalen Koordinatoren dokumentiert (vgl. Schnaudt et al. 2014: 496). Die Gesamtheit der mit dem Übersetzungsprozess verbundenen Schritte verdeutlicht, welche Maßnahmen im ESS möglich und nötig sind, um die Qualität der erhobenen Fragen bewert-

ten und überprüfen zu können und so letztlich zu verlässlichen Aussagen über die Grundgesamtheit zu gelangen.

Wie wird die Qualität der erhobenen Daten überprüft? Wie werden mögliche Fehler beseitigt? Sofern die bisher am Beispiel des ESS diskutierten Maßnahmen und Vorkehrungen der Umfrageforschung konsequent befolgt und angewandt werden, besteht kaum Anlass, Zweifel an der Qualität von Umfragedaten aufkommen zu lassen. Dies kann jedoch nicht einfach angenommen, sondern muss immer auch überprüft werden. Dies geschieht im ESS zunächst durch die Umfrageinstitute und nationalen Teams, welche die erhobenen Daten auf mögliche Fehler kontrollieren und deren Anonymität sicherstellen. Anschließend sehen es die offiziellen Regularien des ESS vor, sowohl die erhobenen Daten als auch sämtliche Dokumentationsmaterialien des kompletten Umfrageprozesses an das Datenarchiv des ESS (NSD, Norwegen) zu übermitteln. Dort werden erste Versionen der nationalen Datensätze erstellt, welche anschließend von den nationalen Koordinatoren (nochmals) auf ihre interne Konsistenz überprüft werden müssen.

Vor Veröffentlichung des vollständigen internationalen Datensatzes wird dieser weiteren Qualitätschecks durch das *ESS Core Scientific Team* sowie den ESS Hauptsitz in London unterzogen. Diese Qualitätschecks überprüfen, ob jedes teilnehmende Land die offiziellen Vorgaben des ESS hinsichtlich Fragebogenentwicklung, Stichprobenziehung und Datenerhebung befolgt und sämtliche vereinbarten Dokumentationsmaterialien eingereicht hat. Sollten sich hierbei Unstimmigkeiten ergeben, so kann dies im schlimmsten Fall zum Ausschluss der Daten des betroffenen Landes aus dem internationalen Datensatz führen. Die von den Ländern eingereichten Dokumentationsmaterialien zur Stichprobenziehung und den Rücklaufquoten werden außerdem dafür verwendet, sog. GewichtungsvARIABLEN zu berechnen, mit deren Hilfe sowohl Abweichungen der Stichprobe von der Zielpopulation als auch der realisierten von der ursprünglich angestrebten Stichprobe korrigiert und ausgeglichen werden können (vgl. Schnaudt et al. 2014: 497-8). Diese Maßnahmen zeigen abschließend noch einmal deutlich, welche komplexen Schritte und Vorkehrungen beim Entstehungsprozess von (internationalen) Umfragedaten nötig sind, damit diese tatsächlich ein repräsentatives Abbild der in der Grundgesamtheit vorhandenen Meinungen und Einstellungen liefern können.

4 Schlussbemerkungen

Ziel war es, angesichts der stetig zunehmenden Anzahl und Vielfalt von Umfragen im Alltag den Entstehungsprozess von Umfragedaten näher zu beleuchten und am Beispiel des internationalen Umfrageprogramms European Social Survey (ESS) zu verdeutlichen, welche vielfältigen und komplexen Anforderungen Umfragen erfüllen müssen, damit ihre Ergebnisse tatsächlich repräsentativ für die jeweils zu Grunde liegende Grundgesamtheit sein können. Kenntnisse über das Zustandekommen von Umfragedaten und die Interpretation von Umfrageergebnissen sind essenziell – gerade weil wir täglich mit diesen konfrontiert werden und uns oftmals in unseren Meinungen und Entscheidungen von diesen beeinflussen lassen (vgl. Faas 2014: 10).

Meinungsumfragen können als Quelle wichtiger Informationen über das politische, soziale oder auch kulturelle Klima einer Gesellschaft fungieren. Voraussetzung ist jedoch, dass sie in ihrer Konzeption klaren Regeln folgen, um so tatsächlich ein Abbild der Meinungen, Einstellungen und Verhaltensweisen innerhalb der Gesellschaft zu reprä-

sentieren. Insofern ist der inflationäre Anstieg von Umfragen, bei welchen offenbleibt, auf welcher Grundlage die Befragten ausgewählt wurden, welche Rücklaufquote sie aufweisen oder wie groß ihr Stichprobenfehler ist, für Umfrageprogramme mit hohen methodischen Standards wie dem ESS eher Fluch als Segen. Letztlich hat die tägliche Konfrontation mit Umfragen das Potenzial, die Teilnahmebereitschaft der Bürger an ebendiesen kontinuierlich zu senken. Angesichts solcher Entwicklungen lassen sich Fragen hinsichtlich der Zukunftsperspektiven von Meinungsumfragen und möglicher Alternativen nicht vermeiden. Bei all diesen Diskussionen und möglichen Abgesängen auf die Umfrageforschung darf jedoch letztlich nicht vergessen werden, dass der große Vorteil und die Stärke von Umfragen gerade in deren Repräsentativität besteht – vorausgesetzt, sie werden nach den in diesem Beitrag grundlegenden Prinzipien erhoben. Folglich sollten wir uns weniger mit Alternativen zu Umfragen beschäftigen, sondern die bestehenden Probleme vielmehr als Herausforderungen sehen, um auch in Zukunft ein repräsentatives Abbild der Meinungen und Einstellungen innerhalb der Bevölkerung liefern zu können. Das Beispiel des ESS zeigt deutlich, dass die Erhebung valider und qualitativ hochwertiger Daten möglich ist, jedoch Zeit, Aufwand und Ressourcen erfordert. Nicht alle Umfragen, die uns im Alltag begegnen oder deren Ergebnisse uns präsentiert werden, können oder wollen diese Ressourcen investieren. Daher ist es umso wichtiger, dass wir uns im Zusammenhang mit Umfragen und deren Ergebnissen stets vor Augen führen, wie diese entstanden sind und welche Schlussfolgerungen auf deren Basis tatsächlich zulässig sind. Der ESS kann hierbei als Orientierungshilfe dienen.

Anmerkungen

- 1 <http://research.mindtake.com/de/osterreicher-falten-gerne-klopapier> [zuletzt abgerufen am 24/02/2016]
- 2 http://www.bmvi.de/SharedDocs/DE/Anlage/VerkehrUndMobilitaet/Fahrrad/fahrrad-monitor-deutschland-2015.pdf?__blob=publicationFile [zuletzt abgerufen am 24/02/2016]
- 3 <https://yougov.de/news/2016/02/09/schiessbefehl-und-verfassungstreue-der-afd-informa/> [zuletzt abgerufen am 24/02/2016]

Literatur

- Curtice, J. (2016). *The Benefits of Random Sampling: Lessons from the 2015 UK General Election*. London: National Centre for Social Research.
- Faas, T. (2014). „Zur Wahrnehmung und Wirkung von Meinungsumfragen.“ *APuZ* 64 (43-45), 3-10.
- Groves, R./Fowler, F./Couper, M./Lepkowski, J./Singer, E./Tourangeau, R. (2009). *Survey Methodology*, 2. Auflage. Hoboken, NJ: Wiley.
- Harkness, J. (2007). „Improving the Comparability of Translation.“ In Jowell, R./Roberts, C./Fitzgerald, R./Eva, G. (Hrsg.): *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*. London: Sage, 79-94.
- Hawkins, O./Keen, R./Nakatudde, N. (2015). *General Election 2015*. London: House of Commons Research Briefings, Nr CPB7186, 28.07.2015.
- Jessen, A. (2014). „Was steckt hinter den Zahlen? Methoden der Demoskopie.“ *APuZ* 64 (43-45), 25-32.
- Schnaudt, C./Weinhardt, M./Fitzgerald, R./Liebig, S. (2014). „The European Social Survey: Contents, Design, and Research Potential.“ *Schmollers Jahrbuch* 134 (4), 487-506.
- Schoen, H./Grezski, R. (2014). „Politische Meinungsforschung in Deutschland.“ *APuZ* 64 (43-45), 18-24.