

Rüdiger Jacob, Andreas Heinz, Jean Philippe Décieux

Umfrage

Rüdiger Jacob, Andreas Heinz,
Jean Philippe Décieux

Umfrage

Einführung in die Methoden der Umfrageforschung

4., überarbeitete und ergänzte Auflage

DE GRUYTER
OLDENBOURG

ISBN 978-3-11-059731-8

e-ISBN (PDF) 978-3-11-059738-7

e-ISBN (EPUB) 978-3-11-059959-6

Library of Congress Control Number: 2019948122

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

© 2019 Walter de Gruyter GmbH, Berlin/Boston

Umschlaggestaltung: marchmeena / iStock / Getty Images Plus

Satz: le-tex publishing services GmbH, Leipzig

Druck und Bindung: CPI books GmbH, Leck

www.degruyter.com

Vorwort zur vierten Auflage

Auch die dritte Auflage des Lehrbuchs hat sich großer Nachfrage erfreut, so dass wir nun die vierte Auflage veröffentlichen. In einer Rezension hat ein Leser über die dritte Auflage geschrieben: „Das Buch bietet eine mehr als ausführliche Einführung in die Welt der Umfrageforschung. Ich habe es für meine Bachelor-Thesis bestellt und wurde nicht enttäuscht, es bietet selbst auf Fragen, die ich mir vorher nicht gestellt hatte, eine Antwort.“ Wir freuen uns über dieses Kompliment und sehen es als Bestätigung unserer Beobachtung, dass sich viele Initiatoren von Befragungen (zunächst) nicht genug Gedanken darüber machen, wie sie fragen sollen. Stattdessen stehen häufig Fragen der statistischen Datenanalyse im Vordergrund, obwohl Fehler in der Analyse so gut wie immer repariert werden können, Fehler in der Datenerhebung aber so gut wie nie.

Daher ist das Ziel des Buches auch in der vierten Auflage unverändert geblieben: Wir wollen die theoretischen Grundlagen und praxisnahe Kenntnisse zu quantitativen Umfragen vermitteln, damit die Leser die Qualität von fremden Befragungen beurteilen, aber auch eigene Befragungen durchführen können.

Neben wohl immer notwendigen Korrekturen haben wir einige Abschnitte aktualisiert. So haben wir eine Reihe zusätzlicher Praxisbeispiele aus zumeist eigenen Studien integriert, gehen stärker auf neuere Befragungstechniken für heikle Themen ein und vertiefen u. a. das Thema „Incentives“, welches im Zuge sinkender Rücklaufquoten immer relevanter wird. Auch befassen wir uns intensiver als vorher mit den Themen Antwortverhalten sowie Umfragen auf mobilen Endgeräten. Den Abschnitt zum Datenschutz haben wir aufgrund vielfältiger und für Nichtjuristen teilweise auch unübersichtlicher Änderungen und Regelungen – wir nennen hier die DSGVO – gestrichen. Wir empfehlen unseren Lesern hier, im Zweifelsfall den Rat einschlägig ausgewiesener Ethikkommissionen oder Juristen einzuholen.

Als praktische Arbeitshilfe stellen wir weiterhin eine Checkliste zum Buch zur Verfügung, in der die wichtigsten Regeln zur Erstellung eines Fragebogens zusammengefasst sind. Zu finden ist diese Liste im Anhang dieses Lehrbuchs und online auf der Seite orbilu.uni.lu/handle/10993/39403 (letzter Abruf: 26.04.2019). Leider konnten wir das Word-Makro zur Erstellung von Fragebögen (mit vorgefertigten Skalen, Intervieweranweisungen, Einleitungstexten und weiterem Feldmaterial) nicht weiterentwickeln, da es in den neuen Versionen von Word nicht mehr möglich ist, diese Makrofunktionen zu editieren. Die alte Version ist allerdings weiterhin erhältlich und funktionsfähig und kann unter orbilu.uni.lu/handle/10993/21421 (letzter Abruf: 26.04.2019) angefordert werden.

Vorwort zur zweiten Auflage

Nach 10 Jahren – mittlerweile in der Forschung eine sehr lange Zeit – war eine Überarbeitung und Erweiterung des vorliegenden Lehrbuchs dringend erforderlich. Wir haben diese Neuauflage zum einen genutzt, um die Fehler und Unklarheiten der ersten Auflage zu korrigieren. Ob uns das vollständig gelungen ist, muss der Leser entscheiden, für entsprechende Rückmeldungen sind wir in jedem Fall sehr dankbar. Zum anderen bedurfte das Buch aber auch einiger Ergänzungen, denn gerade in der Umfrageforschung ist die Entwicklung seit der Jahrtausendwende beileibe nicht stehen geblieben. Die Onlineumfrage als neue Erhebungsform, die in den 1990er-Jahren noch kaum eine Rolle gespielt hat, prägt in zunehmendem Maß die Forschungslandschaft. Dagegen haben sich die Erwartungen, die mit der damals neuen Methode der telefonischen Befragung verbunden wurden, nicht erfüllt, der Telefonsurvey ist wohl nicht „der Königsweg“ der Umfrageforschung. Bedenkt man die methodischen Restriktionen dieser Erhebungsform und die zusehends schlechter werdenden Response-Raten, dann spricht Vieles sogar für eine Renaissance der klassischen Face-to-Face-Interviews (früher PAPI – *paper-assisted personal interview* oder *paper-pencil interview* genannt, jetzt im neuen Gewand des *computer-assisted personal interview* – CAPI), wenn man an für die Bevölkerung repräsentativen Daten interessiert ist. Zu neuen Erhebungsformen kommen Weiterentwicklungen bei den Fragetechniken, und auch in der Forschungsinfrastruktur gab es in den letzten zehn Jahren deutliche Veränderungen – nicht zuletzt auch durch die Weiterentwicklung der Informationstechnologie und des Internets. Wir haben die entsprechenden Abschnitte deutlich gekürzt, da vor dem Hintergrund der sehr dynamischen Entwicklung gedruckte Texte recht schnell veralten. Wir beschränken uns nunmehr auf eine knappe Darstellung zentraler Institutionen und Datenquellen und verweisen ansonsten auf die Webseiten der jeweiligen Anbieter sowie auf eine Linkliste zu den Themen Datenanalyse und Statistik auf unserer Webseite (www.uni-trier.de/index.php?id=61912; letzter Abruf: 26.04.2019). Zur Erstellung eigener Umfragen stellen wir auf der Webseite www.uni-trier.de/index.php?id=34328 (letzter Abruf: 26.04.2019) weitere Hilfsmittel zur Verfügung:

- ein Word-Makro zur Erstellung von Fragebögen mit vorgefertigten Skalen, Intervieweranweisungen, Einleitungstexten u. Ä.;
- eine Kurzanleitung zur Software EFS Survey inklusive einer Kurzanleitung zur Programmierung von HTML-Befehlen sowie eine Formatvorlage in Excel zum Import von Adressdaten in EFS Survey.

Umfrageforschung ist Teamarbeit, bei der Entwicklung von Fragebögen ist eine Gruppe von Forschern immer effektiver und effizienter als ein einzelner Autor. Diese Einsicht manifestiert sich auch in der Zusammensetzung der Autorenschaft. Lehrbücher wie das vorliegende Buch sind sinnvollerweise in einem Team zu erstellen. Auch dabei gab es Veränderungen – erfreuliche, aber leider auch sehr traurige. Mit Andreas Heinz

und Jean Philippe Décieux wurde das Autorenteam um zwei junge Wissenschaftler erweitert, die beide intensiv auf dem Gebiet der Umfrageforschung arbeiten und aufgrund ihrer fachspezifischen Ausbildung – der eine ist Volkswirt, der andere Betriebswirt – und ihrer spezifischen Interessen zu einer deutlichen Bereicherung des Themenspektrums beigetragen haben.

Leider ist Willy Eirmbter nicht mehr dabei. Er verstarb nach schwerer Krankheit im Jahr 2001. Die Lücke, die durch seinen viel zu frühen Tod gerissen wurde, spüren wir noch heute. Ihm ist dieses Buch gewidmet.

Vorwort zur ersten Auflage

Es ist schon eine sozialwissenschaftliche Trivialaussage: Wir leben an der Schwelle zur Informationsgesellschaft, in der dem theoretischen Wissen der Primat zukommt und in der Information und Wissen als „neue“ Produktionsfaktoren eine herausragende Bedeutung haben. Damit kommt auch der empirischen Grundlage von Information und Wissen, den Daten sowie dem Prozess der Datengewinnung eine ganz besondere Bedeutung zu. In den Naturwissenschaften werden die Prozesse der Datengewinnung und die Qualität der Daten kontinuierlich und systematisch verbessert. Die Beobachtungsverfahren in den Naturwissenschaften sind heute problemspezifisch hoch spezialisiert und differenziert, und entsprechend spezialisiert sind auch die Methodenentwicklung und die Methodenanwendung. In den Kulturwissenschaften wird die Datenqualität dagegen häufig (immer noch) als zweit- oder gar drittrangig hinter Theorieflexion und Dateninterpretation angesehen.

Dennoch: Die Vermittlung empirischer Forschungsmethoden und das Wissen über deren erkenntnistheoretische Grundlagen sind für alle Sozialwissenschaften und alle Sozialwissenschaftler immer bedeutsamer geworden. Analysen von Stellenanzeigen für sozialwissenschaftliche Berufsfelder bestätigen dies, und die Zunahme umfassender Lehrbücher zur Methodenlehre empirischer Forschung, die von der wissenschaftstheoretischen Grundlegung bis zur Datenanalyse jedes Thema behandeln, unterstreichen dies mit Nachdruck.

Für die Sozialwissenschaften stellen Daten, die über die Anwendung von Befragungstechniken gewonnen werden, immer noch die gebräuchlichste und wichtigste Datenquelle in der Primärdatenanalyse dar, wenngleich auch andere Daten, insbesondere prozessproduzierte Daten, eine zunehmend größere Anwendung finden. Trotz der Bedeutung, die den Befragungsdaten zukommt, gibt es aber so gut wie keine lehrbuchhafte Einführung, die sich ganz speziell und entsprechend detailliert um die Vermittlung der Methoden und Techniken der quantitativen Befragung bemüht, die es praxisnah und gleichzeitig methodenkritisch insbesondere Anfängern erleichtert, selbst ein Umfrageprojekt durchzuführen, und die es den Konsumenten von Befragungsergebnissen leichter macht, die Methodik der Datengewinnung nachzuvollziehen und die Qualität der Ergebnisse zu beurteilen.

Zwar behandeln die schon genannten allgemeinen Einführungen in die Methoden der empirischen Sozialforschung natürlich auch Befragungen, und es existieren auch einige fundierte und umfassende Übersichtsbände speziell zur Umfrageforschung, was aber auf dem deutschen Markt (anders als in den USA) bisher fehlt, ist ein Lehrbuch für die Formulierung von Fragen und die Konstruktion von Fragebögen.

Dabei ist insbesondere dieser Arbeitsschritt der Konstruktion des Fragebogens, also der Erstellung des Mess- und Datenerhebungsinstruments, in einem empirischen Forschungsprojekt von zentraler Bedeutung. Fehler, die hier gemacht werden, lassen sich später nicht mehr korrigieren. Man kann gar nicht oft genug betonen, dass nichts

für die Ergebnisse und die Qualität einer Untersuchung entscheidender ist als der Fragebogen. Der beste Auswahlplan und die ausgefeiltesten Analysetechniken nützen überhaupt nichts, wenn man die falschen Fragen gestellt und ein fehlerhaftes Messinstrument konstruiert hat.

Trotzdem – oder vielleicht auch gerade weil diese zentrale Arbeitsphase in einem empirischen Forschungsprojekt bislang in der Literatur ein wenig zu stiefmütterlich behandelt wurde – lässt sich nicht nur bei Anfängern in der Umfrageforschung immer wieder die Vorstellung finden, dass die Formulierung der Fragen sowie die Konstruktion des Fragebogens die kleinsten Probleme einer Untersuchung darstellen und hierfür weder viel Zeit noch besonderes Können notwendig sind. Viele unserer Studenten in Lehrforschungsprojekten oder bei empirischen Examensarbeiten und Dissertationen mussten allerdings die mitunter recht leidvolle Erfahrung machen, dass sie in ihrem wissenschaftlichen Arbeiten mit wenigen Einschätzungen so falsch gelegen haben wie mit dieser. Aber nicht nur Studenten der Sozialwissenschaften zeigen diese Fehleinschätzung, auch im Kollegenkreis der Soziologen, Psychologen, Pädagogen, Politikwissenschaftler, Volks- und Betriebswirte herrscht oft die Meinung vor, dass Frageformulierung, Fragebogenkonstruktion und die Befragung selbst „vorwissenschaftliche“ Bereiche wissenschaftlichen Tuns seien. Diesen „handwerklichen“ Abschnitt der Forschungsarbeit überlässt man irgendwelchen Hilfskräften oder zuarbeitenden Institutionen, eine Differenzierung zwischen Wissenschaft und Handwerk, die heute ohnehin fragwürdig geworden ist, man denke z. B. an die zunehmende Professionalisierung der Berufe.

Dementsprechend wenig Zeit, Engagement und Wissen werden für diese Forschungsphase der Instrumententwicklung aufgewendet. Damit kommt der Datengewinnung durch Anwendung von Befragungstechniken in den Sozialwissenschaften häufig eine Geringschätzung zu, die sich ein Naturwissenschaftler nicht erlauben würde und dürfte. Die Folgen dieser mangelnden methodischen Qualität bei Befragungen sind evident:

Zweifel an der Gültigkeit der Messergebnisse, Widersprüche in den Messergebnissen und – wohl am schlimmsten – Scheingenauigkeiten, weil elaborierte Datenanalysen mit zweifelhaften Daten und mit Koeffizienten, die bis zur dritten Stelle hinter dem Komma ausgewiesen werden, dies suggerieren.



Dass die Schwierigkeiten bei der Konstruktion von Fragebögen häufig massiv unterschätzt werden, dürfte auch damit zusammenhängen, dass sich die Umfrageforschung bei oberflächlicher Betrachtung kaum von der gängigen Informationsbeschaffung des Alltagslebens unterscheidet: Wenn man etwas vom anderen wissen will, dann fragt man ihn eben – eine Handlungsroutine, die man schon in frühester Kindheit parallel zur Sprachentwicklung erwirbt.

Nur: Eine Befragung ist eben kein „natürliches“ Gespräch zwischen zwei Personen, sondern eine spezielle Technik der Erhebung von Daten, die gelernt werden muss

und nicht lediglich eine kurze „Anlernphase“ oder ein kurzes Einlesen erfordert. Nur zu häufig bekommen wir Fragebögen in die Hand, deren Anwendung zu allen möglichen Messungen, nur nicht zu den eigentlich intendierten führt, oder die gar Datensätze zur Konsequenz haben, die man besser einstampfen denn analysieren und interpretieren sollte.

Mit diesem Buch wollen wir versuchen, durch eine praxisnahe Vermittlung von Kenntnissen zu wissenschaftlichen Umfragen diesen Problemen mangelnder Gültigkeit und Zuverlässigkeit von Befragungsdaten entgegenzuwirken. Der Schwerpunkt der Darstellung liegt auf Hinweisen zur Formulierung von Fragen und zur Konstruktion von Fragebögen, aber natürlich werden auch viele andere Probleme eines Projekts zur Umfrageforschung behandelt, wie z. B. Fragen des Forschungsdesigns, der Planung und Durchführung der Feldarbeit etc. Lediglich die Phase der Datenaufbereitung und -analyse wird nur sehr knapp dargestellt. Wir verweisen hier auf die Vielzahl der Lehrbücher, die sich mit Fragen der Datenanalyse und Statistik befassen. Genau die Lücke zwischen allgemeinen und sehr umfassenden Lehrbüchern zur empirischen Sozialforschung einerseits und speziellen Lehr- und Handbüchern zur Datenanalyse andererseits wollen wir mit dem vorliegenden Band schließen helfen.

Wir haben diese Einführung ganz bewusst als Lehrbuch im positiven Sinne eines „Kochbuchs“ mit vielen praktischen Beispielen konzipiert und zugunsten einer besseren Lesbarkeit auf einen ausführlichen wissenschaftlichen Apparat weitgehend verzichtet. Stattdessen finden sich Verweise auf weiterführende und vertiefende Literatur am Ende der Hauptkapitel, geordnet nach den jeweiligen in den Kapiteln behandelten Themen.



Wir wollen damit studierenden und praktizierenden Sozialforschern eine Einführung in die Hand geben, die bei Lehrforschungsprojekten, Forschungspraktika, empirischen Examensarbeiten und Dissertationen, Forschungsprojekten und Befragungen in der beruflichen Praxis zu Rate gezogen werden kann und bei der Planung und Durchführung von Befragungen direkte praktische Hilfestellung geben soll.

Dieses Lehrbuch hätte ohne die Unterstützung vieler Kollegen nicht geschrieben werden können. Für die vielfältigen Hilfen bei Literatur- und Beispielrecherchen, Korrekturlesen, der Endredaktion, für viele kritische Anmerkungen und wichtige Hinweise danken wir insbesondere unseren Trierer Kollegen Alois Hahn, Paul Hill, Inge Jansen, Stephanie Kern, Rainer Meckes und Alexandra Uhly. Außerdem möchten wir Jürgen Hoffmeyer-Zlotnik, Achim Koch, Peter Prüfer, Margit Rexroth und Michael Schneid vom Zentrum für Umfragen, Methoden und Analysen (ZUMA), Sabine Haars und Barbara von Harder von GFM/GETAS (heute IPSOS) für Anschauungsmaterial aus Befragungen und eine langjährige konstruktive Zusammenarbeit im Rahmen drittmittelfinanzierter Umfrageforschung sowie Jörg Blasius, Heidi Dorn und Maria Rohlinger vom Zentralarchiv für empirische Sozialforschung (ZA) recht herzlich für die verschiedensten Anregungen, Hinweise und Korrekturen danken. Unser ganz besonderer

Dank aber gilt Isabel Dittgen, Sven Hinrichs, Nina Jakoby und Michael Kern, die uns bei der Gestaltung des Typoskripts, der Module für den Fragebogen und der Erstellung der Diskette eine kompetente und stets zuverlässige, unermüdliche und unersetzliche Hilfe gewesen sind.

Wir sind uns bewusst: Dieses Lehrbuch stellt einen vielschichtigen, möglicherweise recht prekären Kompromiss dar zwischen einführendem und detailbetontem Charakter, zwischen Praxisnähe und Methodenkritik, zwischen Anwendungsorientierung und reflektierender Wissenschaftlichkeit. In Ansehung dieser Gratwanderung bitten wir um Anregungen und Kritik, aber auch um Nachsicht für nicht ganz gelungene Kompromisse.

Trier, im Juni 1999

Dr. phil. Rüdiger Jacob, Akademischer Rat für Soziologie, insbes. Empirische Sozialforschung, Fachbereich IV/Abt. Soziologie, Universität Trier

Dr. rer. pol. Willy H. Eirmbter, Universitätsprofessor für Soziologie, insbes. Empirische Sozialforschung, Fachbereich IV/Abt. Soziologie, Universität Trier

Inhalt

Vorwort zur vierten Auflage — V

Vorwort zur zweiten Auflage — VI

Vorwort zur ersten Auflage — VIII

- 1 Einleitung: Vom Nutzen der Umfrageforschung — 1**
 - 1.1 Ergebnisse und Aussagekraft von Umfragen — 2
 - 1.2 Zur Funktion der Umfrageforschung — 3
 - 1.3 Zur Geschichte der Umfrageforschung — 6
 - 1.4 Weiterführende Literatur — 23

- 2 Theoretische Grundlagen: Befragung als Datenabrufprozess? — 25**
 - 2.1 Messtheorie — 26
 - 2.2 Das Interview als soziale Situation — 41
 - 2.3 Weiterführende Literatur — 56

- 3 Der Forschungsprozess: Die Planungsphase — 58**
 - 3.1 Worum geht es? Die Forschungsfrage — 60
 - 3.2 Art der Forschung — 66
 - 3.3 Grundgesamtheit und Stichprobe — 70
 - 3.4 Auswahlverfahren I: Zufallsauswahlen — 73
 - 3.5 Auswahlverfahren II: Nicht-zufällige Auswahlen — 86
 - 3.6 Access-Panels — 92
 - 3.7 Querschnitt oder Längsschnitt? — 94
 - 3.8 Auswahlplan — 96
 - 3.9 Zeit- und Kostenplan — 97
 - 3.10 Der Forschungsantrag — 102
 - 3.11 Weiterführende Literatur — 104

- 4 Der Forschungsprozess: Die Erhebung der Daten — 105**
 - 4.1 Befragungsarten — 106
 - 4.2 Fragen — 133
 - 4.3 Aufbau des Fragebogens — 200
 - 4.4 Feldarbeit — 207
 - 4.5 Weiterführende Literatur — 234

5	Der Forschungsprozess: Eingabe, Analyse und Präsentation der Daten — 238
5.1	Datenaufbereitung und Datenanalyse — 238
5.2	Der Forschungsbericht: Checkliste — 247
5.3	Weiterführende Literatur — 249
6	Forschungsinfrastruktur — 251
6.1	Infrastruktureinrichtungen der empirischen Sozialforschung — 251
6.2	Markt- und Meinungsforschungsinstitute — 253
7	Replikative Surveys und amtliche Statistik — 254
7.1	Sozio-oekonomisches Panel (SOEP) — 256
7.2	Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) — 257
7.3	International Social Survey Programme (ISSP) — 258
7.4	Eurobarometer — 258
7.5	European Social Survey (ESS) — 259
8	So nicht: Beispiele für schlechte Fragen — 260
	Anhang: Checkliste für Fragebögen — 275
	Glossar — 279
	Quellenverzeichnis — 287
	Weiterführende Literatur — 293
	Stichwortverzeichnis — 299

1 Einleitung: Vom Nutzen der Umfrageforschung

Sozialforscher nutzen Befragungen als Verfahren der Datengewinnung in zwei Varianten: In der qualitativen Sozialforschung werden u. a. offene, narrative oder Leitfadeninterviews geführt, wobei selten mehr als 50 Personen befragt werden. Demgegenüber werden in der quantifizierenden Forschung mit standardisierten Befragungen meist Daten über eine Vielzahl von Personen oder Institutionen erhoben.¹ Mit dem Begriff „Umfrageforschung“ und dem meist synonym verwendeten Begriff „Demoskopie“ wird dabei die quantitative Variante bezeichnet. Mit Umfragen werden unterschiedliche und in den Sozialwissenschaften zumeist qualitative Sachverhalte so aufbereitet, dass sie unter Verwendung von Zahlen als Variable behandelt und numerischen Operationen unterworfen werden können. Nur solche quantitativen Verfahren eignen sich zur Beschreibung und Analyse von gesellschaftlichen Massenphänomenen – wobei Umfrageforschung inzwischen selbst zu einem Massenphänomen geworden ist.

Meinungsumfragen und die alsbaldige Verkündigung ihrer zahlenmäßigen „Ergebnisse“ gehören heute so selbstverständlich zu unserem Alltag und scheinen uns vielfach so folgenlos oder auch so banal zu sein, dass wir sie kaum noch zur Kenntnis nehmen und schon gar nicht mehr darüber nachdenken, welche wichtige Bedeutung Umfragen für demokratisch verfasste Gesellschaften in vielerlei Hinsicht haben.

Es vergeht ja auch wirklich kein Tag, an dem man nicht in den Medien sehen, hören oder lesen kann, wer regieren würde, wenn am nächsten Sonntag Wahl wäre, wie viele Deutsche gegen Kampfeinsätze der Bundeswehr oder für die LKW-Maut auch auf Landstraßen sind, wie viele fürchten, Opfer eines Verbrechens zu werden, oder den Zuzug von Ausländern begrenzen wollen, Angst vor der Zukunft haben, den Missbrauch sozialer Leistungen beklagen oder kein Bier einer bestimmten Marke trinken, aber gerne Volksmusik hören. Über 90 % aller sozialwissenschaftlichen Daten stammen aus Umfragen – auch wenn sich durchaus darüber streiten lässt, ob die Umfrageforschung der Königsweg der empirischen Sozialforschung ist: Die Hauptstraße ist sie allemal.

Trotz oder vielleicht gerade wegen dieser Inflation und Allgegenwart von „Ergebnissen“ der Umfrageforschung, bei der bisweilen eine gewisse Beliebigkeit der Themen einhergeht mit trivialen Befunden, erfreut sich dieser Zweig der Forschung in der Bevölkerung allerdings keines allzu guten Rufs. Dies hängt nicht zuletzt mit der Qualität dieser Umfragen zusammen: Jeder – vorausgesetzt er hat das notwendige Geld – kann heute zu jedem Thema Umfragen durchführen oder durchführen lassen. Dabei sind die Qualität der einzelnen Fragen, des gesamten Fragebogens, die Durchführung

¹ Probleme qualitativer Interviews behandeln wir in diesem Lehrbuch nicht, sondern verweisen auf die weiterführende Literatur am Ende dieses Kapitels. Wenn in diesem Buch von Befragungen oder Interviews die Rede ist, dann sind damit stets standardisierte, quantitative Befragungen gemeint, sofern nicht ausdrücklich anders vermerkt.

der Befragung, die Techniken zur Sicherstellung der Repräsentativität, die Codierung der Befragungsergebnisse, die Datenbereinigung etc. nahezu immer von nachrangiger Bedeutung.

Diese zentralen Bereiche der Forschung werden nicht nur häufig unterschätzt, sie werden oft auch als weniger wichtig angesehen und – als vermeintlich der Wissenschaft vorgelagert – an externe Organisationen oder an Hilfskräfte delegiert. Selbst die Meinung, bei der Praxis der Umfrageforschung handele es sich um eine besondere „handwerkliche Kunst“, die man am besten an spezialisierte Institute vergibt, zeugt meist nicht davon, dass die Datengewinnung im Prozess der empirischen Forschung ein hohes Ansehen genießt.

Von Interesse ist das möglichst schnell erzielte und möglichst spektakuläre „Ergebnis“, die Qualität der Daten ist kaum ein Thema. Übersehen wird dabei stets, dass jede Datenanalyse nur so gut sein kann wie die zuvor erhobenen Daten, auf denen sie basiert. Folgen dieser Art von schlechter Umfrageforschung sind fehlerhafte oder widersprüchliche Ergebnisse, Ergebnisse von einer relativen Beliebigkeit, interessen-gesteuerte Manipulation. Die eher harmlose Konsequenz dieser auf Umfragen basierenden Sozialforschung ist die Übersättigung der Öffentlichkeit mit „Daten“ über Trivialitäten. Problematischer ist es aber, wenn mit unprofessionellen, inflationär durchgeführten Befragungen eine allgemeine Geringschätzung der Umfrageforschung, eine Anzweiflung der Seriosität aller Sozialforschung und nicht zuletzt auch die Ablehnung der Kooperation bei Befragungen verbunden sind, welche die Feldbedingungen für seriöse und wichtige Forschung verschlechtern.

1.1 Ergebnisse und Aussagekraft von Umfragen

Zu diesen Zweifeln an der Seriosität der Umfrageforschung hat sicherlich auch die übliche „Ergebnispräsentation“ beigetragen: Man konfrontiert die Öffentlichkeit mit eindimensionalen Tabellen, in denen die univariaten Verteilungen als prozentuale Häufigkeiten bestimmter Merkmale dokumentiert werden, also z. B. wie viele Befragte am nächsten Sonntag die SPD, eine der Unionsparteien, die Grünen, die FDP oder sonstige Parteien wählen würden.

Auf der Grundlage – tatsächlicher oder behaupteter – repräsentativer Stichproben suggerieren die üblicherweise verwendeten Prozentzahlen ein genaues Ergebnis im Sinne eines Punktwertes, der den „wahren“ Wert für die Grundgesamtheit darstellen soll. Diese Insinuation blendet zum einen aus, dass wir stets mit Messfehlern bei Befragungen rechnen müssen und schon deshalb Gewissheit über „wahre“ Werte nicht zu erlangen ist. Zum anderen widerspricht die Präsentation von Umfrageergebnissen in Form von sogenannten „Punktschätzern“ (d. h., die Ergebnisse der Befragung werden mit den jeweiligen Anteilswerten in der Grundgesamtheit gleichgesetzt) der Logik des zugrundeliegenden induktiven Schließens. Auf diese Einschränkungen wird in

der Regel aber nicht hingewiesen, weil sie angeblich der interessierten Öffentlichkeit nicht vermittelbar sind. Tatsächlich können die Ergebnisse von Befragungen, die auf zufällig ausgewählten Stichproben basieren, nur als ungefähre Hinweise interpretiert werden, keinesfalls aber als exakte Größen.

Abgesehen von Wahlprognosen ist dies im Grunde auch nicht so wichtig, da Einzelbefunde vergleichsweise uninteressant sind. Interessant sind dagegen Antwortunterschiede, die aber wiederum statistisch signifikant sein müssen. Dies ist gerade bei Wahlprognosen mit einem Kopf-an-Kopf-Rennen der großen oder der kleinen Parteien oder aber bei Parteien nahe der Fünf-Prozent-Grenze problematisch. Wenig erhellend ist auch ein Ergebnis, wonach – um ein Beispiel aus unseren eigenen Forschungen zu nennen – 8,7 % der Befragten glauben, eine Ansteckung mit HIV drohe auch bei flüchtigen Körperkontakten (Eirmbter, Hahn und Jacob 1993), auch wenn dies tatsächlich exakt dem wahren Wert der Grundgesamtheit entspräche. Soziologisch bedeutsam ist hingegen, dass dieser Prozentsatz erstens gruppenspezifisch erheblich differiert und zweitens mit anderen Merkmalsausprägungen in typischer Weise korreliert und theoretisch gut erklärt werden kann.

Eindimensionale Häufigkeitstabellen sind deshalb auch nicht das *Ergebnis* seriöser Umfrageforschung, sondern der *Ausgangspunkt* der eigentlichen Datenanalyse unter Einbeziehung mehrerer Merkmale, wobei komplexe Muster und Zusammenhänge aufgedeckt werden sollen. Das Ergebnis eines Forschungsprojekts ist die Interpretation, die Deutung der in den Tabellen dokumentierten Daten vor dem Hintergrund bestimmter Annahmen. Diese sind im Idealfall in einer dem Analyseteil vorangestellten theoretischen Begründung der Forschungsfrage dokumentiert.

1.2 Zur Funktion der Umfrageforschung

Wenn man Umfrageforschung betreibt, ist man nicht an der Meinung konkreter einzelner Personen interessiert, obwohl naturgemäß immer nur Einzelpersonen befragt werden können. Dies mag widersprüchlich anmuten, ist aber eine unabdingbare Voraussetzung für Befragungen bzw. für jede Form von quantifizierender Forschung.

Die Abstraktion von Individuen ist eine formale Notwendigkeit, wenn man soziale Sachverhalte zählen oder messen will. Dazu muss man eine gemeinsame Basis herstellen. Dies gelingt dadurch, dass man nicht die gesamte, komplexe und vermutlich auch mehr oder weniger widersprüchliche Persönlichkeit von Menschen zu erfassen versucht, sondern nur bestimmte Ausschnitte, bestimmte Merkmale misst. Diese Merkmale werden noch dazu künstlich eingeengt, indem man sie in der Regel einigen wenigen vorformulierten und deshalb für alle Befragten gleichen Kategorien zuordnet, sogenannten standardisierten Merkmalsausprägungen, die wiederum mit Zahlen codiert werden – zum Zweck quantifizierender Darstellung und numerischer Kalkulationen.

Nicht Individuen in ihrer ganzen Komplexität werden also gemessen (was im Übrigen auch ein hoffnungsloses Unterfangen wäre), sondern ausgewählte Eigenschaften von „Merkmalsträgern“ eines klar definierten Kollektivs. Diese Merkmalsträger werden anonym erfasst und gezählt und sind für die Zwecke der Forschung prinzipiell austauschbar. Quantitative Befragungen richten sich immer an eine Mehrzahl, an ein Aggregat von Personen mit dem Ziel der Beschreibung und Analyse von Merkmalsverteilungen, von komplexen Mustern oder Konfigurationen, von sozialen Regelmäßigkeiten. Abgesehen von Vollerhebungen basieren Ergebnisse quantifizierender empirischer Forschung somit nicht auf den Angaben von allen Mitgliedern eines Kollektivs, verfolgen zumeist aber den Anspruch, Aussagen über alle Mitglieder einer solchen, durch bestimmte Merkmale definierten Grundgesamtheit machen zu können.

Untersucht werden stets nur einige Merkmale und Merkmalskombinationen unter bewusster Ausblendung aller sonstigen Charakteristika, die die befragten Personen noch aufweisen, die aber für die anstehende Forschungsfrage als irrelevant angesehen werden. Schon aus diesem Grund ist es unerlässlich, eine Forschungsfrage zu formulieren und zumindest in Grundzügen theoretisch zu begründen: Man muss entscheiden und begründen, welche Merkmale man für die Untersuchung eines bestimmten Sachverhaltes als wichtig ansieht. Bereits hier wird die Gültigkeit des allgemeinen Anspruchs an jede empirische Forschung offensichtlich:



Keine Empirie ohne Theorie!

Die Beschränkung des Interesses auf standardisierte Merkmale erklärt auch, warum ein Interview kein Gespräch ist. Interviewer sind üblicherweise Fremde, die plötzlich und unerwartet vor der Tür stehen oder anrufen und dann mitunter recht intime Themen ansprechen. Dazu stellen sie vorformulierte Fragen und liefern die Antworten gleich mit, wechseln das Thema anscheinend nach Gusto (tatsächlich bestimmt der Fragebogen die Reihenfolge der Fragen), gehen auf ihr Gegenüber aber kaum ein (wann sie dies dürfen, wird ebenfalls durch den Fragebogen bestimmt). Die Leistung der Befragten reduziert sich in der Regel auf die Wahl einer Antwortalternative, ein „Einbringen“ der ganzen Persönlichkeit in das „Gespräch“ ist gerade nicht erwünscht.

Diese hochgradig künstliche und asymmetrische soziale Situation ist ein wesentlicher Grund für eine relativ weitverbreitete Kritik an quantitativen Datenerhebungen mit standardisierten Fragebögen. Dabei wird insbesondere moniert, dass solche Fragebögen komplexe Strukturen über Gebühr reduzieren und vereinfachen und deshalb weniger zur Erhebung von Fakten als vielmehr zur Produktion von Artefakten beitragen, also zu Ergebnissen, die nicht der Realität entsprechen.

Solche Einwände sind berechtigt und sehr ernst zu nehmen. In der Tat besteht bei Datenerhebungen mit standardisierten Fragebögen die Gefahr, dass man Artefakte produziert. Denn es ist für Befragte leicht, sich für irgendeine der vorgegebenen Antwortalternativen zu entscheiden – auch wenn sie zu dem erfragten Sachverhalt

keine Kenntnisse oder keine Meinung haben bzw. ihre tatsächliche Meinung nicht äußern wollen. Befragungen sind psychische Low-Cost-Situationen: Die Wahl inhaltlicher Antworten auch bei Meinungslosigkeit oder Unkenntnis der untersuchten Sachverhalte bleibt sozial folgenlos, vermittelt aber in der Situation den Eindruck, ein informierter und kompetenter Zeitgenosse zu sein.

Die durchaus berechtigte Kritik kann aber kein Grund sein, auf quantitative Befragungen zu verzichten. Ganz im Gegenteil ist diese Kritik Anlass, auf die Konstruktion des Messinstruments viel Sorgfalt und Mühe zu verwenden, um solche Fehlerquellen zu minimieren. Zudem darf man nicht übersehen, dass beide Varianten der Befragung – qualitative wie quantitative – versuchen, latente Phänomene zu erfassen, die der direkten Beobachtung entzogen sind, nämlich Einstellungen, Werthaltungen, Wünsche, Hoffnungen, Kenntnisse, Verhaltensdispositionen usw. All dies ist nur indirekt durch Befragungen zu erfahren. Dabei sind sowohl Fragen eines standardisierten Fragebogens als auch solche aus einem Leitfadeninterview stets nur Indikatoren für das eigentlich interessierende Phänomen. Damit besteht in beiden Fällen auch die Gefahr der Artefakt-Produktion.

Es kann denn auch nicht darum gehen, die eine Form der Datenerhebung durch die andere zu ersetzen. Beide sind notwendig und haben je spezifische Vor- und Nachteile. Das unbestrittene Verdienst aller qualitativen Befragungsformen liegt darin, bestimmte Sachverhalte und Zusammenhänge, bestimmte Einstellungsensembles wesentlich detaillierter und tiefergehend zu erforschen, als dies mit quantitativen Befragungen möglich wäre. Dabei werden oft auch Zusammenhänge aufgedeckt, die verblüffend oder kontraintuitiv sind und bei quantitativen Befragungen so nie aufgefallen wären.² Wenn man aber wissen will, wie verbreitet solche Konstellationen von Einstellungen in einer bestimmten Population sind, ob und wenn ja für welche Gruppen sie typisch sind, ob sie sich im Zeitverlauf wandeln, dann lässt sich dies nur mit quantitativen repräsentativen Befragungen ermitteln.

Bei qualitativen Befragungen erkaufte man sich den Vorteil der inhaltlichen Tiefe und der Nähe zur Lebenswelt ganz konkreter Interviewpartner durch den Nachteil eines Mangels an Reproduzierbarkeit und Generalisierbarkeit der Ergebnisse. Umgekehrt ist eben dies nur um den Preis eines vereinheitlichten Messkonzeptes, vorgegebener Fragen mit wenig Spielraum für die Interviewpartner und der völligen Abstraktion von Einzelpersonen zu haben. Welche Form der Datenerhebung angezeigt ist, hängt von den jeweiligen Forschungsinteressen und den Zielen der Untersuchung ab.

² Beispielsweise hat Jutta Dornheim in qualitativen Interviews durch konkrete Verhaltensfragen zum Umgang mit Krebskranken herausgefunden, dass Krebs von den meisten der von ihr befragten Personen als ansteckend eingeschätzt wird, obwohl diese die direkte Frage nach der Infektiosität von Krebs (wie sie als dichotome Frage mit den Antwortvorgaben „ja“ und „nein“ auch in quantitativen Befragungen eingesetzt würde) verneint haben (Dornheim 1983).

Zur Erforschung sozialen Wandels, als Datengrundlage für Politik und Verwaltung oder als Stimmungsbarometer in einer pluralistischen Gesellschaft sind quantitative Befragungen unverzichtbar. Pointiert haben Noelle-Neumann und Petersen dies zusammengefasst:

In unserer Zeit ist Demoskopie ein unentbehrliches Hilfsmittel geworden, weil man sich über die so groß und abstrakt gewordene Gesellschaft, die der unmittelbaren Beobachtung entrückt ist, anders nicht zuverlässig informieren kann. Die Demoskopie liefert dabei keine Informationen über bestimmte Personen, sondern nur über die Gesellschaft oder bestimmte, durch Merkmale zusammengefaßte Gruppen der Gesellschaft. Unsere Sprache besitzt schon diese Unterscheidungsfähigkeit: Das Forschungsobjekt sind alle, nicht jeder. (Noelle-Neumann und Petersen 2005, S. 625)

Analysegegenstand der quantifizierenden empirischen Forschung sind also Kollektive, strukturelle Effekte und soziale Regelmäßigkeiten, Untersuchungseinheiten sind aber meist Individuen als Träger spezifischer Merkmale. Träger gleicher Merkmale und Merkmalskombinationen werden zu Gruppen zusammengefasst. Damit geht zwangsläufig eine gewisse Egalisierung einher. Personen mit gleichen Merkmalsausprägungen bei bestimmten Merkmalen werden ungeachtet aller sonstigen (möglichen) Unterschiede als gleich und austauschbar behandelt. Sie werden bei weitergehenden Analysen zudem auch noch mit Gruppen, die andere Merkmalsausprägungen aufweisen, verglichen, indem man Anteilswerte, Mittelwerte oder sonstige Kennziffern in diesen Gruppen unterscheidet: Haben Männer andere Ansichten zu Auslandseinsätzen der Bundeswehr als Frauen? Haben Personen aus der Unterschicht ein höheres Herzinfarktrisiko als Personen aus der Mittelschicht? Haben Studenten verhaltensbedingt ein höheres AIDS-Risiko als ältere Landbewohner, und haben sie deshalb auch mehr Angst vor AIDS?

Jede Form der Datenanalyse basiert im Grunde auf solchen Gruppenvergleichen, mittels derer Gesamtheiten von Merkmalsträgern immer wieder neu untergliedert und zum Gegenstand numerischer Kalkulationen gemacht werden können.

1.3 Zur Geschichte der Umfrageforschung

Ein solches Unterfangen wäre Menschen der ständisch geprägten Gesellschaft des Mittelalters wahrscheinlich nie in den Sinn gekommen und mutet ja auch heute noch für viele fremdartig an. Die Idee der Quantifizierung und die damit verbundene gedankliche Abstraktion von konkreten Menschen sind an einen bestimmten gesellschaftlichen Entwicklungsstand gekoppelt. Die agrarisch geprägte Gesellschaft des Mittelalters mit relativ kleinen und ähnlichen, aber weitgehend autarken Dörfern und nur wenigen (ebenfalls nach heutigen Maßstäben nicht sehr großen) Städten bot für Quantifizierungsversuche wenig Raum und wenig Anlass. Die Produktion war weitgehend als Subsistenzwirtschaft zu bezeichnen, Märkte existierten nur in Ansätzen

bzw. ganz im Sinne heutiger Wochenmärkte und waren lokal begrenzt. Der überlokale Warenverkehr blieb wenigen Handelshäusern oder Händlergilden wie der Hanse vorbehalten.

Erst mit der zunehmenden Entwicklung von marktvermittelten Tauschbeziehungen unter allgemeiner Verwendung des universellen Zahlungsmittels Geld – einer abstrakten Größe – änderte sich die Situation. Mit dem Medium des Geldes ließen sich ganz unterschiedliche Waren und Dienstleistungen vergleichen, beziffern und quantifizieren, nämlich durch ihren Preis. Höchst unterschiedliche Sachverhalte wurden in dem Einheitsmaß von Zahlen ausgedrückt, berechenbar und vergleichbar.³ Die Menschen des Mittelalters lebten in einer Feudalgesellschaft mit klaren und zugeschriebenen Gruppenzugehörigkeiten und einer – im Vergleich zur modernen Gesellschaft – nur geringen Rollendifferenzierung. Ihnen wäre auch der Gedanke der analytischen Kombination immer neuer Gruppen anhand bestimmter Merkmale fremd gewesen, bei der ein immer gleiches Individuum mal als Mann, mal als Maschinenschlosser, mal als 53-Jähriger, mal als Gewerkschaftsmitglied verschiedenen Gruppen zugerechnet wird. Die Egalisierung durch Zahlen und die Typisierung von ganz unterschiedlichen Individuen sind als neuartige Betrachtungsweisen gesellschaftlicher Zusammenhänge ein Produkt des Bürgertums.

Quantifizierung geht also einher mit der Entwicklung und Ausdifferenzierung der modernen Gesellschaft, die sich, um es auf einen verkürzten Nenner zu bringen, selbst als ein immenser Quantifizierungsprozess darstellte. Nicht nur nahm im Zuge des Prozesses des demografischen Übergangs die Zahl der Menschen in den europäischen Gesellschaften ganz beträchtlich zu, es vermehrten sich auch die Rollen und Funktionen. Die Produktion von Gütern wurde im Zuge der industriellen Revolution exponentiell ausgeweitet, das Wissen über die Welt wuchs ebenfalls explosionsartig an und damit auch die Möglichkeiten der Weltanschauung, der Werte, Lebensziele und Erwartungen. Die Kirche, ursprünglich tatsächlich nur im Singular als katholische Kirche existent, hatte bis zur Reformation ein Deutungsmonopol in allen weltanschaulichen Fragen. Dieses Monopol wurde durch alternative Deutungsmuster verschiedener Reformatoren nachhaltig beseitigt. Die mentalen Voraussetzungen für die wirtschaftlichen Entwicklungen der Neuzeit, die sich unter die Überschrift „Industrielle Revolution“ subsumieren lassen, wurden durch die Reformation überhaupt erst geschaffen, wie Weber in seiner berühmten Abhandlung über die protestantische Ethik und den Geist des Kapitalismus ausführlich dargestellt hat.

Parallel dazu wurde das kirchliche Deutungsmonopol auch durch die Entwicklung und Ausdifferenzierung der abendländischen Wissenschaften beginnend in der

³ Geld als Zahlungsmittel ist natürlich sehr viel älter und war bereits in den Stadtstaaten Mesopotamiens in Gebrauch. Ein universell anerkanntes Zahlungsmedium wie Geld stellt aber nur die notwendige Voraussetzung für die Entwicklung überlokaler Märkte und Handelsverbindungen dar. Weitere Bedingungen sind u. a. sichere Handelswege, leistungsfähige Transportmittel und das Vorhandensein von überschüssigen Gütern für den Tausch.

Renaissance angegriffen. Der Beginn der Quantifizierung von Weltanschauungsangeboten ist mithin zeitlich noch vor dem Beginn der Vermehrung von materiellen Ressourcen durch die einsetzende Industrialisierung anzusetzen.

Als Konsequenz dieser vielfältigen Entwicklungen leben wir heute in einer pluralistischen, funktional ausdifferenzierten Massengesellschaft, die in der Tat jeder direkten Beobachtung entzogen ist – anders als die funktional äquivalenten kleinräumig organisierten Sozialverbünde feudaler Gesellschaften. Zugleich ist der Informationsbedarf zur Verwaltung und Steuerung solcher Gesellschaften aber beträchtlich, wobei sich dieses Informationsbedürfnis historisch zunächst auf objektive Bestandsdaten für militärische oder fiskalische Zwecke beschränkte. Die Kriegführung mit stehenden Heeren und ein (auch) dafür benötigtes kalkulierbares Steueraufkommen waren frühe Triebfedern für die Sammlung bestimmter Daten. Die Zentralgewalten der sich etablierenden Nationalstaaten waren deshalb zunächst an bevölkerungs- und wirtschaftsstatistischen Daten interessiert, die durch Zählungen erhoben wurden. Die Geschichte der empirischen Sozialforschung ist zunächst also eine Geschichte des Datensammelns und mit der Entwicklung der modernen Statistik eng verknüpft. Die Entwicklung der Statistik, wie wir sie heute kennen, begann in Form der Sammlung und Auswertung größerer Datenmengen um ca. 1650 in England und Frankreich, den beiden großen europäischen Staaten, in denen sowohl die Zentralisierung und die Entwicklung von Nationalstaaten als auch die Industrialisierung historisch zuerst eingesetzt haben. Statistik war immer schon ein wichtiges Hilfsmittel der merkantilistischen Staatsverwaltung und hat diese Funktion auch in demokratischen Staaten nicht verloren, ihre Entwicklung wurde aber von Anfang an auch von wissenschaftlich interessierten Privatgelehrten und Universitätsprofessoren vorangetrieben. Dabei nutzte die damalige Staatsverwaltung zunächst nur Daten, die wir heute als „prozessproduzierte Daten“ bezeichnen würden – Daten, die automatisch als Nebenprodukt staatlichen Handelns anfallen. Als Beispiel ist etwa das Erheben von Steuern zu nennen, das vor allem dazu dient, den Staatshaushalt zu finanzieren, aber gleichzeitig auch Daten liefert, die zu statistischen Zwecken genutzt werden können (Hald 1990, S. 82; Peters 1987, S. 79). Später wurden dann auch Zählungen mit dem Ziel der Dokumentation von Beständen durchgeführt, zielgerichtete Erhebungen blieben bis Anfang des 19. Jahrhunderts auf solche Zählungen beschränkt, Surveys im heutigen Sinn wurden erst im Gefolge der einsetzenden Industrialisierung entwickelt. Für den Informationsbedarf und die Verwaltungsaufgaben des absolutistischen und ständisch geprägten Staates, in dem die Masse der Bevölkerung nur sehr wenige (legitime) Handlungsalternativen hatte, waren diese Daten völlig ausreichend.

Die Erfassung von Meinungen, Bewertungen, Erwartungen, Lebenszielen, Wünschen usw. wurde erst in dem Maß wichtig, in dem auch die individuellen Dispositionsmöglichkeiten zunahmen und kollektiv-strukturelle Konsequenzen haben konnten. Der erste Schritt auf dem Weg zur modernen Statistik und empirischen Sozialforschung bestand mithin – wenn diese überspitzte Dichotomisierung zulässig ist – in der Entwicklung der Demografie: Die Verwaltung der absolutistischen Staaten benö-

tigte möglichst genaue Angaben über die Zahl gebärfähiger Frauen und wehrfähiger Männer und über Steuerbürger. Der nächste Schritt hin zur Demoskopie verlief parallel zum politischen Wandel von absolutistisch zu demokratisch verfassten Staaten und der damit verbundenen zunehmenden Pluralisierung von Lebenslagen und Lebensstilen.

Dabei muten auch die Probleme hinsichtlich der Datenlage und der Datenbeschaffung auffallend modern oder zeitlos an:

No reliable data were available [...] Two obstacles are mentioned by the authors who have dealt with the work in this period: the unwillingness of the population to give information, because of their fear of increased taxes; and the tendency of governments, whenever statistical information was available, to treat it as highly classified because of its possible military value. (Lazarsfeld 1961, S. 149).⁴

Damals ging es allerdings noch im Wesentlichen um einfache Bevölkerungsstatistiken, heute hingegen handelt es sich bei Surveys und dem Problem des *non response* neben Vermögensfragen, die immer schon eine hohe Non-Response-Rate aufgewiesen haben, um sogenannte heikle oder sensitive Fragen etwa nach der persönlichen Hygiene oder der religiösen Orientierung. Aspekte ökonomischer Verwendbarkeit von Informationen werden häufig von Unternehmen geltend gemacht und entsprechende Angaben als betriebliche Interna behandelt. Wettbewerbsgründe werden z. B. immer wieder von Krankenkassen als Begründung angeführt, Strukturdaten über ihre Versicherten nicht zu publizieren. Schon früh wurden deshalb in der Statistik Schätzverfahren und Indikatoren verwendet, wenn der eigentlich interessierende Sachverhalt einer direkten Beobachtung nicht zugänglich war:

Thus, the ingenuity of early scholars was directed mainly toward obtaining estimates of population size and age and sex distribution from [...] indirect evidence. Multiplying the number of chimneys by an assured average family size or inferring the age structure of the population from registered information regarding age at the time of death were typical procedures in what was then called political arithmetic. (Lazarsfeld 1961, S. 150)

Sehr früh lassen sich aber auch eindeutig wissenschaftlich geprägte Ansätze beobachten, die zentrale Beiträge zur heutigen Statistik und empirischen Sozialforschung geliefert haben.

⁴ So auch Hahn und Braun 1973, S. 43 f.: „Daß die Universitätsstatistiker so gut wie keine Zahlenangaben machen, hängt [...] auch mit der Tatsache [zusammen], daß es zu dieser Zeit nahezu unmöglich ist, quantitatives Datenmaterial über ein Land oder ein Gebiet zu erhalten. Zum einen ist die Bevölkerung aus Furcht vor Steuererhebungen nur schwer dazu zu bewegen, Angaben über ihre Situation zu machen, zum anderen werden die wenigen statistischen Informationen, die in den einzelnen Staaten zusammengetragen werden können, oft wegen ihres wirtschaftlichen und militärischen Wertes geheim gehalten, eine Praxis, die sich bis in das 19. Jahrhundert hinein erstreckt.“

1.3.1 Politische Arithmetik und Moralstatistik

1662 deckte John Graunt (1620–1674), der „Kolumbus der Statistik“, unter Verwendung von Geburts- und Sterbelisten Londons, die von den Kirchengemeinden geführt wurden, bis dahin unbekannte Regelmäßigkeiten auf. Der naheliegende Schluss, dass es in anderen Bereichen des gesellschaftlichen Lebens ebenfalls solche Regelmäßigkeiten geben müsse, etwa bei Selbstmordraten oder Straftaten, wurde in der Folge dann immer wieder untersucht und bestätigt. Diese in England von John Graunt, William Petty (1623–1687) und Edmund Halley (1656–1742) entwickelte Form der tabellari-schen Analyse großer Datensätze wird als „politische Arithmetik“ bezeichnet. Sie wird heute als der eigentliche Vorläufer der amtlichen Statistik sowie generell der quantita-tiven Darstellung und Analyse sozialer Phänomene angesehen (Pearson 1978; Lazars-feld 1961; John 1884).

Die Entdeckung bestimmter Regelmäßigkeiten in großen Datensätzen, etwa die langjährige Konstanz von Heirats- oder Geburtenraten oder bestimmten Straftaten, führte schon relativ früh auch zu der expliziten Forderung nach einer staatlich or-ganisierten systematischen Datenerhebung, also nach einer amtlichen Statistik. So postulierte August Ludwig von Schlözer (1735–1809):

Die wichtigsten statistischen Daten kann nur die Regierung, nicht der Privatmann schaffen. Die Regierungsstatistik ist schon darum notwendig, weil manches Datum sich nur selten zeigt und seine Erheblichkeit erst durch die Summe aller kleinen Zahlen offenbar wird; diese aber kann nur eine Regierung veranstalten. (zit. nach John 1884, S. 112)

Bis zur Umsetzung dieser Forderung dauerte es allerdings noch etwas – insbeson-dere in Deutschland. Erst mit Ernst Engel (1821–1896), von 1860 bis 1882 Leiter des Preußischen Statistischen Büros und Entdecker des „Engelschen Gesetzes“⁵, wurde in Preußen eine amtliche Statistik mit einem kontinuierlichen Erhebungsprogramm in-stitutionalisiert. Engel war ein überzeugter Anhänger des belgischen Mathematikers Adolphe Quetelet (1796–1874), der den Ansatz der englischen politischen Arithmeti-ker konsequent weiterentwickelt hatte (zu Quetelet vgl. John 1884, S. 332 ff.).

Quetelet ging von der These aus, dass das menschliche Leben von der Geburt bis zum Tod von bestimmten Gesetzmäßigkeiten strukturiert wird, die sich nur durch die Analyse aggregierter Zahlen, durch die Abstraktion von den zahllosen individuellen Besonderheiten erkennen lassen. Erst bei der Betrachtung einer großen Zahl von Fäl-len lassen sich Zufälligkeiten von sozialtypischen Regelmäßigkeiten unterscheiden. Quetelet interessierte sich also nicht für individuelle Besonderheiten oder die Lebens-geschichten von einzelnen Persönlichkeiten, sondern für bestimmte Typen von Sozi-

⁵ Das Engelsche Gesetz besagt, dass bei steigendem Einkommen einer Familie ein immer kleinerer Prozentsatz dieses Einkommens für Lebensmittel verwendet wird.

alcharakteren, für „mittlere Menschen“ als idealtypische Aggregationen immer wieder auftretender Eigenschaften und Verhaltensweisen.⁶

Quetelet verwendete für seine Betrachtungen und Analysen vor allem bevölkerungsstatistische Daten wie Eheschließungen, Geburten- und Sterberaten, Selbstmorde oder sonstige Daten der Kriminalstatistik. Subsumiert wurden all diese Statistiken unter den Begriff der „Moralstatistik“. Dieser Begriff wurde zwar nie eindeutig bestimmt, es wurden darunter aber sukzessive alle aggregierten Darstellungen menschlicher Handlungsweisen gefasst. Man kann also sagen, dass „Moralstatistik“ ein Synonym für den heute üblichen Begriff der „Sozialstatistik“ ist, mit der heutigen Bedeutung von „Moral“ dagegen nichts zu tun hat. Der Begriff implizierte nie eine ethische Dimension, sondern sollte nur deutlich machen, dass es hier – anders als in den Naturwissenschaften – um gesellschaftliche Phänomene ging.

Was Quetelet mit seinen Arbeiten in einer vorher nicht dagewesenen Systematik anwendete und deutlich machte, war das sogenannte Gesetz der großen Zahl, wonach bestimmte typische Eigenschaften in einem Aggregat umso deutlicher sichtbar werden, je größer die Zahl der betrachteten Fälle ist.⁷

Der Gedanke der Quantifizierung sozialer Phänomene und der Aufdeckung sozialer Gesetze erhielt im 19. Jahrhundert eine zusätzliche Dynamik, weil die Naturwissenschaften und insbesondere die Physik mit eben diesem Programm und der Entdeckung von Naturgesetzen außerordentlich erfolgreich waren. Deren Methodik, theoretische Begründungen und das zentrale Paradigma der Determiniertheit aller physischen Vorgänge wurden von frühen quantitativ orientierten Sozialforschern wie Quetelet bis in die Begrifflichkeit übernommen. Quetelet etwa nannte seinen Ansatz „Soziale Physik“.

Ganz unzweifelhaft hat Quetelet die empirische Sozialforschung durch sein Lebenswerk ein gutes Stück vorangebracht, allerdings blieb sein Verständnis von sozialen Regelmäßigkeiten letztlich sehr formal und atheoretisch, da er buchstäblich alles als typisch für eine Gesellschaft ansah, was sich quantifizieren ließ.

Quetelet arbeitete letztlich ohne Augenmaß. Er sammelte Material über alle möglichen Faktoren – über körperliche Eigenschaften und demographische Sachverhalte (Fruchtbarkeit/Geburt/Totgeburten, Eheschließungen, Sterbefälle, Bevölkerung, Kraft der Bevölkerung, Einatmung und Pulsschläge, Gewicht des Bluts, Schnelligkeit und Beweglichkeit der Menschen) sowie über die

⁶ Diese Vorgehensweise hat sich bis heute in bestimmten Bereichen erhalten, etwa wenn Warenkörbe des typischen Konsums zusammengestellt werden und man z.B. konstatiert, dass „der deutsche Staatsbürger“ durchschnittlich 2,3 Tassen Kaffee pro Tag trinkt, dreimal in der Woche Fleisch isst, 1,2 Zeitungen abonniert hat und 2,4-mal pro Jahr in Urlaub fährt (um fiktive Zahlen zu nennen).

⁷ Das Gesetz der großen Zahl findet heute vor allem in der Stichprobentheorie Anwendung. Bei Zufallsauswahlen steigt mit wachsendem Stichprobenumfang die Wahrscheinlichkeit, dass der Stichprobenmittelwert (und andere Kennwerte der Stichprobe) mit den Werten der Grundgesamtheit (fast) übereinstimmen. Das Gesetz der großen Zahl ist deshalb ein wichtiges Hilfsmittel zur Festlegung des Stichprobenumfangs, vgl. dazu Abschnitt 3.3 „Grundgesamtheit Grundgesamtheit und Stichprobe“.

sittlichen und geistigen Eigenschaften der Menschen (intellektuelle Fähigkeiten, Geisteskrankheiten, moralische Eigenschaften) – und registrierte eifrig Regelmäßigkeiten, wo immer sie im Material auftraten. Nicht jedes Muster, das in statistischem Material erkennbar ist, sagt nun aber Wesentliches über die Gesellschaft aus, auf die sich die statistischen Daten beziehen. [...] Über die Feststellung statistischer Regelmäßigkeiten allein stößt man nicht auf die typischen Strukturen einer bestimmten Gesellschaftsformation. [...] Welche Regelmäßigkeiten sind letztendlich für die Struktur und Entwicklungsdynamik der bürgerlichen Gesellschaft in der zweiten Hälfte des 19. Jahrhunderts bedeutsam? Auf diese Frage kann Quetelet, von seinem rein statistischen Gesetzesverständnis herkommend, keine Antwort geben. (Kern 1982, S. 43 f.)

Was Quetelet fehlte, war ein theoretischer Zugang zu seinem Material, ein Interpretationsrahmen, der es ihm ermöglicht hätte, sinnvolle Zusammenhänge von bloßen Häufigkeitsverteilungen zu trennen. Auguste Comte (1798–1857), der Schöpfer des Begriffs „Soziologie“ und große Gegenspieler von Quetelet, brachte dies auf die noch heute verwendete und zuvor bereits vorgestellte Formel, dass sinnvolle Empirie ohne Theorie nicht möglich ist (Comte, 1923).

1.3.2 Wahrscheinlichkeitstheorie und das Gesetz der großen Zahl

Ebenso bedeutsam wie das Gesetz der großen Zahl (und damit eng zusammenhängend) ist für die moderne Statistik und empirische Sozialforschung die Wahrscheinlichkeitstheorie. Deren Grundlegung begann 1654, als der Chevalier de Méré dem Philosophen und Mathematiker Blaise Pascal (1623–1662) wegen der Gewinnchancen in einem damals sehr populären Würfelspiel um Rat fragte (Kennedy 1993, S. 88 f.). Pascals Überlegungen dazu haben die Brüder Jakob (1654–1705) und Johann Bernoulli (1667–1748), beide Mathematiker in Basel, bei der Weiterentwicklung der Wahrscheinlichkeitstheorie und der Formulierung des Gesetzes der großen Zahl wesentlich beeinflusst. Die Stichprobentheorie zur Ziehung repräsentativer Zufallsstichproben basiert wesentlich auf diesen Überlegungen (Kennedy 1993, S. 101 ff.).

1.3.3 Deutsche Universitätsstatistik

Demgegenüber ist der deutsche Beitrag zur Entwicklung der modernen Statistik zunächst eher marginal und beschränkte sich im Prinzip auf die Wortschöpfung des Begriffs „Statistik“, wie er in der „deutschen Universitätsstatistik“ zum Ausdruck kommt. Diese Disziplin lässt sich auf Vorlesungen von Herman Conring (1609–1681) an der damals noch existierenden Universität Helmstedt und auf Weiterentwicklungen durch Gottfried Achenwall (1719–1772) zurückführen. Der Begriff „Statistik“ ist ein Lehnwort aus dem Italienischen und leitet sich ab von *statista*, womit ein Mann bezeichnet wurde, der sich mit Angelegenheiten des Staates befasst (Pearson 1978, S. 2).

Die Universitätsstatistik Conringscher und Achenwallscher Prägung ist denn auch ganz wesentlich eine Hilfsdisziplin der Staatswissenschaften, die sich vor allem mit

der Beschreibung von „Staatsmerkwürdigkeiten“ befasst. Die Universitätsstatistik informierte in verbaler Form über die Bevölkerung und die Finanzen von Staaten, über deren militärisches Potenzial, die Staatsverfassung und sonstige rechtliche Fragen, die landwirtschaftliche Produktion, aber auch über klimatische und geografische Besonderheiten. Die deutsche Universitätsstatistik ist damit eher eine Vorläuferin der politischen Wissenschaften als der quantifizierenden Statistik. Sich selbst mit quantifizierenden Darstellungen zu befassen, kam den Universitätsstatistikern nicht in den Sinn. Vertreter und Anhänger der politischen Arithmetik wurden von ihnen als „Tabellenknechte“ diffamiert. Zur Beschreibung von Staaten verwendeten die Universitätsstatistiker in der Regel allgemeine qualitative Ausdrücke wie: dicht bevölkert, groß oder klein, stark oder schwach, tugendsam oder lasterhaft, aber keine Quantifizierungen.

Dass sich heute gleichwohl der Begriff „Statistik“ mit der methodischen Tradition der politischen Arithmetik verbindet, ist darauf zurückzuführen, dass John Sinclair (1754–1835) diesen Begriff 1798 einfach übernahm und erfolgreich für deren Methodik reklamierte (Pearson 1978, S. 2).

1.3.4 Sozialenqueten

Weder die politischen Arithmetiker noch die deutschen Universitätsstatistiker führten allerdings eigene systematische Erhebungen durch. Sie stützten sich bei ihren Arbeiten entweder auf gewissermaßen prozessproduzierte, d. h. durch Verwaltungshandeln entstandene, Daten (im Fall der Verwendung von Moralstatistiken) oder auf Reiseberichte oder sonstige narrative Berichte.

Die stetig zunehmenden und evidenten Probleme der raschen Industrialisierung führten allerdings schon bald zu einem reformerisch orientierten Informationsbedürfnis über die Lage „der Armen“ bzw. der Arbeiterschaft, das nur mit speziellen Erhebungen zu befriedigen war. Mit diesen Erhebungen wurden, beginnend in englischen Industriestädten in den 1930er- und 1940er-Jahren Daten über Arbeits- und Wohnbedingungen und die gesundheitliche Lage der Arbeiterschaft erhoben. Neben diesen objektiven Indikatoren der sozialen Lage wurden erstmals auch subjektive Aspekte erfasst, etwa zur religiösen Einstellung.

So, wie die Industrialisierung in Deutschland später einsetzte als in England und verstärkt erst mit der Reichsgründung 1871 in Gang kam, so wurde auch die „soziale Frage“ im Gefolge dieses Wandlungsprozesses erst später als Problem erkannt, dem man mithilfe sozialer Reformen und staatlicher Sozialpolitik begegnen wollte.

Die dazu benötigten Daten wurden in Orientierung an den englischen Enqueten seit 1873 von dem *Verein für Socialpolitik* erhoben. Wie in England ging es dabei vor allem um die Lage der Arbeiterschaft, und zwar sowohl in den Industriestädten als auch – der spezifischen Wirtschaftsstruktur Deutschlands entsprechend – auf dem Land. Besonders bekannt geworden ist die Enquete von 1892, *Die Verhältnisse der Landarbeiter im ostelbischen Deutschland*, an der Max Weber maßgeblich beteiligt war.

Für die Datenbeschaffung der Enqueten des *Vereins für Socialpolitik* wurden hauptsächlich Fragebögen verwendet. Allerdings blieben die Ergebnisse solcher Untersuchungen vielfach unbefriedigend, was vor allem damit zusammenhing, dass die Auswahl der Zielpersonen oft willkürlich erfolgte und die Teilnahmebereitschaft der Arbeiter eher schwach ausgeprägt war. Zudem steckte die Methodik zur Erfassung und Messung von Einstellungen noch in den Kinderschuhen und konnte auch deshalb nur schwer systematisch vorangetrieben werden, weil es keine geeignete dauerhafte Forschungsinfrastruktur gab.

1.3.5 Institutionalisierung der empirischen Sozialforschung

Entsprechende Bemühungen lassen sich in Deutschland erst in der Phase der Weimarer Republik mit einer Reihe von Gründungen sozialwissenschaftlicher Forschungsinstitute beobachten. 1919 wurde an der Universität zu Köln das *Forschungsinstitut für Sozialwissenschaft* von Lindemann, Scheler, von Wiese und Brauer gegründet, 1924 folgte das *Institut für Sozialforschung* an der Universität Frankfurt a. M., dem u. a. Wittfogel, Horkheimer, Fromm, Marcuse und Adorno angehörten. Beide Institute wurden von den Nationalsozialisten geschlossen, und die meisten ihrer Mitarbeiter emigrierten. Nach dem Zweiten Weltkrieg stellten beide Standorte wichtige Kristallisationskerne für einflussreiche Richtungen der sich neu entwickelnden Soziologie und Sozialforschung dar, die unter den Begriffen *Kölner Schule* und *Frankfurter Schule* bekannt geworden sind (vgl. dazu und zu den folgenden Anmerkungen Kern 1982).

Ein ganz wesentlicher Impuls für die Methodenentwicklung während der Zwischenkriegszeit kam außerdem aus Wien, und zwar von der *Österreichischen Wirtschaftspsychologischen Forschungsstelle* um Paul F. Lazarsfeld. Er und seine Mitarbeiter entwickelten und verwendeten zur Erforschung gesellschaftlicher Probleme eine Vielzahl unterschiedlicher Methoden, die von der quantifizierenden Statistik bis zu qualitativen Interviews und teilnehmender Beobachtung reichten.

Ein besonders eindrucksvolles Beispiel für eine solche Multi-Methoden-Untersuchung ist der inzwischen als Klassiker zu bezeichnende Forschungsbericht *Die Arbeitslosen von Marienthal* von Marie Jahoda, Paul F. Lazarsfeld und Hans Zeisel (1997). Ziel der Untersuchung war eine möglichst umfassende Beschreibung und Analyse des Problems „Arbeitslosigkeit“ und seiner psychischen und sozialen Konsequenzen. Zu diesem Zweck hielt sich ein Forscherteam mehrere Wochen in dem kleinen Ort Marienthal auf. In dieser Zeit wurden – um nur einige Beispiele zu nennen – umfassende Beobachtungen über die Wohnverhältnisse und das Familienleben von 478 Familien erstellt, mit 62 Personen wurden vertiefende biografische Interviews geführt, 80 Personen füllten Zeitverwendungsbögen aus. Schulaufsätze von Kindern zu den Themen „Was ich mir wünsche“ und „Was will ich werden“ wurden inhaltsanalytisch ausgewertet. Die Lehrer wurden zu den Leistungen und der Motivation der Schüler befragt. Statistisch ausgewertet wurden die Umsätze von Gasthäusern und Geschäften. Be-

rühmt worden ist insbesondere die die Beobachtung der Ortsbewohner beim Gehen durch Marienthal, wobei die Schrittgeschwindigkeiten verglichen wurden. Dabei stellte sich heraus, dass die Arbeitslosen signifikant langsamer gingen als Personen, die (noch) Arbeit hatten.

Möglich war diese Art der engagierten und nicht ganz billigen Forschung allerdings nur, weil Lazarsfeld das Institut auch durch die Erstellung von sogenannten Marktstudien finanzierte, also Marktforschung betrieb und im Auftrag verschiedener Produzenten auch die Absatzchancen etwa für Schokolade, Bier oder Konfektionsartikel bestimmter Marken erforschte.

Auch das Wiener Institut wurde nach dem „Anschluss“ Österreichs 1938 von den Nationalsozialisten geschlossen. Lazarsfeld, der schon 1936 in die USA emigriert war, führte seine Arbeit dort allerdings weiter. Bekannt geworden ist er insbesondere durch seine Forschungen im Rahmen des *Radio Research Project* am *Office of Radio Research*, mit dem er 1937 an der Princeton University begann und das er 1939 an der Columbia University in New York weiterführte, woraus 1944 das *Bureau of Applied Social Research* wurde. Nachhaltigen Einfluss hat Lazarsfeld vor allem auf die Massenkommunikationsforschung ausgeübt. Die heute geläufigen Konzepte des *two-step flow of communication* und des *opinion leader*, wie sie gerade auch in deutschen Marktforschungsstudien vielfach Verwendung finden, basieren auf einer Untersuchung über den Präsidentschaftswahlkampf von 1940. Aber auch Datenanalyseverfahren wurden in dieser Untersuchung entscheidend verbessert. Lazarsfeld hat hier die heute noch sehr gebräuchliche mehrdimensionale Tabellenanalyse entwickelt und vorgestellt.

Die Entwicklung der empirischen Sozialforschung in den USA wurde mithin stark von deutschsprachigen Emigranten beeinflusst, die in den USA ihre Arbeit sehr erfolgreich fortsetzen konnten. Allerdings hat sich dort sehr früh auch eine eigenständige empirische Sozialforschung entwickelt, die es überhaupt erst ermöglichte, dass die von den Nationalsozialisten vertriebenen Forscher ihre Arbeit vergleichsweise nahtlos fortsetzen konnten. Anlass zu dieser eigenständigen Entwicklung in den USA waren ebenfalls soziale Probleme im Gefolge der Industrialisierung, der Expansion nach Westen und der massenhaften Immigration.

1.3.6 Social Research in den USA

Zur Untersuchung solcher sozialen Probleme wurden ähnlich wie in Europa Untersuchungen durchgeführt, nur waren diese nicht gruppenspezifisch, sondern regional organisiert. Diese frühen *social surveys* lieferten für spezifische Fragen eine Bestandsaufnahme für bestimmte Regionen – in der Regel für Städte –, etwa zur kommunalen Infrastruktur oder zum Gesundheitswesen. Allerdings waren die Methoden der Datengewinnung dieser Untersuchungen noch sehr heterogen, wobei vor allem bereits verfügbares Material verwendet und analysiert wurde (Kern 1982, S. 182 ff. Ausführlich mit der Entwicklung in den USA von 1890 bis 1960 befasst sich Converse 1987).

Die begriffliche Gleichsetzung von *survey* und Umfrageforschung erfolgte erst später und wurde geprägt durch Arbeiten von Soziologen der sogenannten *Chicago School*, von Gallup und seinen Mitarbeitern, von Soziologen der *Columbia School* um Lazarsfeld sowie von Cantril mit seinem *Office of Public Opinion Research*.

Die *Chicago School* mit ihren prominenten Vertretern Robert E. Park, Louis Wirth, William I. Thomas, William F. Whyte und Ernest W. Burgess arbeitete zunächst insbesondere mit Methoden der Feldforschung, wie sie in der Ethnologie verwendet werden, es wurden also primär Beobachtungsverfahren eingesetzt. Diese wurden aber in der Folge ergänzt durch Befragungen, mit denen man die Einstellungen der Untersuchungspersonen zu bestimmten Sachverhalten erfassen und quantifizieren wollte.

Die Idee, nicht unmittelbar beobachtbare und durchaus auch recht unterschiedliche Phänomene wie subjektive Einstellungen durch bestimmte Messoperationen zu quantifizieren und damit in gewisser Weise auch zu objektivieren, wurde in dieser Form erstmals im Kontext der Attitudenforschung von Soziologen der Chicago-Schule ausformuliert und umgesetzt (Maus 1973, S. 42). Die dazu entwickelten Messinstrumente werden auch heute noch als Skalen bezeichnet. Die erste dieser Skalen wurde 1926 von Emory S. Bogardus entwickelt und sollte die soziale Distanz zwischen Personen aus der Sicht des Befragten erfassen. Ebenfalls Ende der 1920er-Jahre hat Louis Thurstone eine Reihe von Skalen entwickelt, von denen die „Methode der gleicherscheinernden Intervalle“ zumindest in einführenden Lehrbüchern bis heute vorgestellt wird. Auch die heute wohl am häufigsten verwendete Rating-Skala zur Messung von Einstellungen, die sogenannte Likert-Skala, stammt aus dieser Zeit und wurde von Rensis Likert 1932 veröffentlicht.

Neben der Erforschung von Einstellungen gab es in den USA in Vorwahlzeiten immer schon ein großes Interesse daran, wie anstehende Wahlen und insbesondere die Präsidentschaftswahlen ausgehen würden. Zu diesem Zweck wurden sogenannte *pre-election surveys* oder *polls* durchgeführt, und zwar vorrangig von kommerziellen Instituten.

In diesem Zusammenhang bekannt geworden ist vor allem das „Desaster“ der Zeitschrift *Literary Digest*, die 1936 die bislang größte Befragung in der Geschichte der Umfrageforschung durchführen ließ: Es wurden an 10 Millionen Telefonbesitzer Fragebögen verschickt und 2,4 Millionen ausgefüllt zurückgeschickt. Die aus dieser Umfrage abgeleitete Prognose des Wahlsiegs des republikanischen Bewerbers Landon erwies sich als falsch.⁸ Demgegenüber lag George Gallup mit seiner auf einer alters- und geschlechtsquotierten Stichprobe von nur 1.500 Personen richtig, der einen Sieg Roosevelts vorhergesagt hatte (Bryson 1976).

Zwar diskutierte man schon seit den 1920er-Jahren Zufallsstichproben als optimale Verfahren zum Gewinnen von Erkenntnissen über größere Grundgesamtheiten,

⁸ Näheres zu dieser Umfrage und den Gründen ihres Scheiterns findet sich im Abschnitt „Grundgesamtheit Grundgesamtheit und Stichprobe“ in Abschnitt 4.3.

jedoch erwies es sich als Problem, dass es in den USA keine Meldepflicht gibt. Über ein geeignetes Verfahren, wie eine Zufallsauswahl aus einer nicht genau bekannten Population gezogen werden kann, verfügte man noch nicht. Deshalb wurden zunächst andere Verfahren entwickelt, mittels derer man auf der Basis von Stichprobenuntersuchungen zu brauchbaren Ergebnissen zu gelangen glaubte. Die Überlegung, dass man aus einem verkleinerten Abbild der amerikanischen Wahlbevölkerung dann auch präzise das Wahlergebnis ableiten könnte, führte zur Entwicklung von Quotenauswahlverfahren. Zwar war klar, dass man dabei nicht alle Merkmale der Bevölkerung berücksichtigen konnte, man ging aber davon aus, dass die Berücksichtigung der wichtigsten Merkmale hinreichend valide Daten liefern würde. Als „wichtigste“ Merkmale galten Alter und Geschlecht, und dementsprechend wurden die Stichproben quotiert. Die dazu notwendigen Daten entnahm man den Ergebnissen der Volkszählungen. Obwohl die Quotenauswahl Repräsentativität nicht sicherstellen kann (siehe dazu Abschnitt 3.5.2), stellte dieses Verfahren im Vergleich zu den vorher üblichen, völlig willkürlichen Methoden der Rekrutierung von Befragungspersonen einen deutlichen Fortschritt dar, wie sich an dem Erfolg Gallups gegenüber der „Literary-Digest-Stichprobe“ gezeigt hatte. Damit gelangten Polls und Meinungsumfragen auch zu einem gewissen Ansehen in der Bevölkerung.

Dass aber auch die Quotenauswahl nur ein suboptimales Verfahren ist, zeigte sich wiederum bei einer Präsidentschaftswahl, und zwar 1948. Bei dieser Wahl lag das Gallup-Institut falsch und prognostizierte einen Sieg des republikanischen Kandidaten Dewey. Stattdessen wurde aber Truman wiedergewählt.

Allerdings gab es inzwischen auch ein Zufallsauswahlverfahren für nicht genau bekannte Grundgesamtheiten, wie sie etwa die Bevölkerung der USA darstellt. Das *Bureau of the Census* hatte in den 1940er-Jahren ein solches Verfahren auf der Basis einer Flächenstichprobe entwickelt, welches in der Folge zum Standardverfahren der großen amerikanischen Umfrageinstitute avancierte und in modifizierter Form auch von der deutschen Umfrageforschung übernommen wurde (siehe Abschnitt 3.4).

Dass Gallup zunächst an der Quotenauswahl festhielt (und einige deutsche Marktforschungsinstitute wie das Institut für Demoskopie (IfD) Allensbach dieses Verfahren bis heute anwenden), liegt an den niedrigeren Kosten. Quotenauswahlen sind bei gleichem Stichprobenumfang deutlich billiger als Flächenstichproben.⁹

Allerdings waren nicht nur private Meinungsforschungsinstitute an der Entwicklung einer stichprobenbasierten Survey-Forschung beteiligt, sondern auch die Universitäten. Neben den schon erwähnten Arbeiten von Lazarsfeld und seinen Mitarbeitern zuerst an der Princeton- und dann an der Columbia University ist hier nochmals Hadley Cantril zu nennen, der 1940 das *Office of Public Opinion Research* ebenfalls in Princeton gegründet hatte. Cantril wurde kurz darauf zum Berater des Präsidenten

⁹ Für die Bundesrepublik kann als grober Richtwert gelten, dass eine Quotenauswahl – je nach Art der Quotierung – zwischen 20 % und 40 % billiger ist als eine Flächenstichprobe.

berufen. Dies führte dazu, dass der Survey-Forschung eine wichtige politikberatende Funktion zukam und eine Reihe von Umfragestudien über die Akzeptanz bestimmter administrativer Maßnahmen und Einstellungen zur amerikanischen Politik allgemein durchgeführt wurden. Diese Untersuchungen wurden zum Teil vom Präsidenten, zum Teil von den zuständigen Fachministerien in Auftrag gegeben und ab 1941 in zunehmendem Maß von dem neu gegründeten *National Opinion Research Center* (NORC) an der Universität Chicago organisiert und durchgeführt.

1.3.7 Umfrageforschung in Deutschland nach 1945

Nach der Zäsur durch die Barbarei des Nationalsozialismus und den Krieg waren die Bedingungen für empirische Sozialforschung in Deutschland nach 1945 aus mehreren Gründen grundsätzlich recht günstig. Soziologie als die Wissenschaft, die in erster Linie für empirische Sozialforschung (zumindest im größeren Stil) zuständig war, galt als unbelastetes Fach, da nahezu alle deutschen Vertreter dieser Disziplin von den Nationalsozialisten verfolgt und – soweit sie nicht ins Ausland fliehen konnten – ermordet worden sind. Re-emigrierte Soziologen hatten deshalb nach dem Krieg gute Chancen, Fördermittel und Stellen zu bekommen, zumal man vor dem Hintergrund der Nachkriegsprobleme – gänzlich andere Sozialstruktur als vor dem Krieg, Notwendigkeit der Integration von Millionen von Flüchtlingen aus den Ostgebieten, Wiederaufbau – auf soziologische Untersuchungen und Daten dringend angewiesen war (Kern 1982, S. 218). Außerdem sahen insbesondere die Amerikaner Soziologie als wichtiges Hilfsmittel bei der *reeducation* der Deutschen an, was damit zusammenhing, dass die Soziologie in den USA als akademisches Fach wie auch in ihrer praktischen Anwendung seit Beginn des Jahrhunderts einen hohen Stellenwert besaß.

Neben diesem Interesse an konkreten, praktisch verwertbaren Erkenntnissen sollte die Soziologie in gewisser Weise aber auch „Sinnerklärungen“ liefern. Soziologie, verstanden als Gesellschaftsanalyse, bot Erklärungen für die Entwicklung unterschiedlicher Gesellschaftsformen und machte deutlich, dass es keine gleichsam „natürlichen“ Gesellschaften gibt, sondern dass diese stets das Produkt bestimmter, immer nur vorläufiger historischer Konstellationen und sozialer Zusammenhänge sind.

Diese grundsätzlich begrüßenswerte Doppelorientierung des Fachs – praktische empirische Sozialforschung auf der einen Seite und allgemeine Gesellschaftstheorie auf der anderen Seite – ist allerdings bis heute auch ein zentraler Anlass für Konflikte und Auseinandersetzungen innerhalb der deutschen Soziologie. Dabei wurde ein kultur- und gesellschaftskritischer Ansatz bei gleichzeitiger Kritik an einem unmittelbaren Praxisbezug empirischer Sozialforschung vor allem von der schon erwähnten *Frankfurter Schule* um Horkheimer und Adorno verfolgt, die sich mit der Wiedereröffnung des Frankfurter Instituts für Sozialforschung 1950 zu etablieren begann. Aber auch andere bekannte Soziologen der Nachkriegszeit standen der empirischen Sozial-