

Notwendige Grundinformationen zur Beurteilung von Umfrageergebnissen

I. Berufsständische Regeln

Um die Aussagekraft von Umfrageergebnissen sowie die Möglichkeiten und Grenzen ihrer Interpretation beurteilen zu können, sind Informationen über die Umfrage selbst und die Modalitäten ihrer Durchführung unumgänglich. Deshalb verpflichten die berufsständischen Regeln des 'IHK/ESOMAR Internationaler Kodex für die Praxis der Markt- und Sozialforschung' die Forscher, „dem Auftraggeber sämtliche angemessenen technischen Einzelheiten eines für diesen Auftraggeber durchgeführten Forschungsprojekts zur Verfügung“ zu stellen.¹

Aus diesem Grund sind auch Journalisten durch den Pressekodex gehalten, „bei der Veröffentlichung von Umfrageergebnissen von Meinungsbefragungsinstituten die Zahl der Befragten, den Zeitpunkt der Befragung, den Auftraggeber sowie die Fragestellung mitzuteilen“.² Dabei handelt es sich allerdings erst um minimale Anforderungen, deren Erfüllung allein noch nicht ausreicht, die Angemessenheit der Interpretation von Umfrageergebnissen vollständig beurteilen zu können.

Der Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM) hat in verschiedenen Pressemitteilungen und - zusammen mit der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und dem Berufsverband Deutscher Markt- und Sozialforscher e.V. (BVM) - in einer „Richtlinie für die Veröffentlichung von Ergebnissen der Wahlforschung“³ sowie zuletzt und am ausführlichsten in seinen 'Standards zur Qualitätssicherung in der

¹ ESOMAR 1995, Regel 25

² Deutscher Presserat 1997, S. 243

³ ADM, ASI, BVM 1997

Markt- und Sozialforschung⁴ die methodisch-technischen Grundinformationen aufgelistet, die notwendig sind, um die Aussagefähigkeit und Tragweite der Ergebnisse von Umfragen beurteilen zu können.

Auch die Deutsche Forschungsgemeinschaft (DFG) hat in einer Denkschrift „Qualitätskriterien in der Umfrageforschung“⁵ einen - nahezu identischen - Katalog notwendiger Grundinformationen bei der Veröffentlichung von Umfrageergebnissen vorgelegt. In welcher Form und in welchem Umfang diese Forderungen in der journalistischen Praxis Berücksichtigung finden, ist neben dem entsprechenden Problembewußtsein der Journalisten im Einzelfall auch von der Art der Publikation abhängig.

II. Notwendige Grundinformationen

Dafür Sorge zu tragen, daß sowohl dem Auftraggeber als auch den Medien und der Öffentlichkeit die Möglichkeiten und Grenzen von Umfrageergebnissen bewußt werden, ist in allen Bereichen der Markt- und Sozialforschung wichtig und notwendig. Besonders deutlich wird diese Notwendigkeit aber in Forschungsbereichen, deren Ergebnisse in der Regel auf großes öffentliches Interesse stoßen und deshalb häufig veröffentlicht und intensiv diskutiert werden. Idealtypisch hierfür ist die empirische Wahlforschung, deren Ergebnisse außerdem zumeist kurze Zeit später auf dem Prüfstand des tatsächlichen Wahlergebnisses stehen. Im einzelnen sind zur vollständigen Beurteilung der Aussagefähigkeit und Tragweite von Umfrageergebnissen die folgenden Grundinformationen notwendig, die deshalb im Forschungsbericht für den Auftraggeber der Untersuchung, aber möglichst auch bei der Veröffentlichung der Untersuchungsergebnisse vollständig anzugeben sind:

1) Name des Instituts

Durch die namentliche Nennung des Forschungsinstituts, das die Umfrage durchgeführt hat, soll keine vordergründige Bewertung der Umfrageergebnisse durch die Reputation des Instituts erfolgen. Man kann davon ausgehen, daß die Durchführung von Umfragen nach anerkannten wissenschaftlich-methodischen Regeln zum beruflichen Selbstverständnis der Markt- und Sozialforscher gehört. Außerdem verpflichten die berufsständischen Regeln des 'IHK/ ESOMAR

⁴ ADM 1998, S. 33

⁵ DFG 1998, S. 99

Internationalen Kodex für die Praxis der Markt- und Sozialforschung' die Institute „Forschung so anzulegen, daß sie kostenwirksam und von angemessener Qualität ist, und sie dann entsprechend den mit dem Auftraggeber im einzelnen vereinbarten Vorgaben auszuführen“.⁶ Die im ADM zusammengeschlossenen Markt- und Sozialforschungsinstitute sind zudem durch die Satzung des Vereins zur Einhaltung der Qualitätskriterien des ADM verpflichtet.

Die namentliche Nennung des Forschungsinstituts soll vielmehr zur Realisierung des Wissenschaftskriteriums der intersubjektiven Überprüfbarkeit von Forschungsergebnissen beitragen. Sie macht es interessierten Personen grundsätzlich leichter, weitergehende Informationen über die den Forschungsergebnissen zugrundeliegende Umfrage zu erhalten. Allerdings sind Auskünfte dieser Art den Instituten nur insoweit rechtlich möglich als dies mit dem Auftraggeber der Umfrage bei der Auftragserteilung geregelt oder von diesem nachträglich gestattet wurde.

2) Aufgabenstellung der Untersuchung

Bei der Interpretation von Umfrageergebnissen ist die Aufgabenstellung und die Zielsetzung der Untersuchung zu beachten, um Fehl- oder Überinterpretationen zu vermeiden. Vor allem bei der Veröffentlichung von Ergebnissen der empirischen Wahlforschung wird dies allzu häufig nicht genügend berücksichtigt. Beispielsweise werden durch die Messung der aktuellen politischen Stimmung vier Wochen vor einer Wahl keine Stimmen gezählt. Es ist deshalb unzulässig, die Ergebnisse einer solchen Umfrage als Wahlprognose zu interpretieren.

Im Bereich der empirischen Wahlforschung lassen sich drei Typen von Untersuchungen unterscheiden. Sinngemäß kann diese Typologie auch auf andere Bereiche der Umfrageforschung übertragen werden: Es gibt erstens aktuelle Messungen der (Partei-)Sympathien oder (Partei-) Präferenzen. Diese Umfragen liefern ein Abbild der (politischen) Stimmungen, die jedoch nicht unbedingt verhaltensrelevant sind; vor allem dann nicht, wenn zwischen der Stimmungsmessung und dem (Wahl-)Termin eine größere zeitliche Distanz liegt.

Zweitens werden aktuelle (Wähler-)Potentiale ermittelt. Hierbei werden neben den Umfrageergebnissen auch externe Daten und Erkenntnisse berücksichtigt und beide Informationen in einer Modellrechnung zusammengeführt. Aber auch aktuelle (Wähler-)Potentiale stellen noch keine (Wahl-)Prognose dar, da

⁶ ESOMAR 1995, Regel 12

sie nur für den jeweiligen Zeitpunkt der Messung gelten, nicht aber für den Termin der (Wahl-)Entscheidung.

Drittens schließlich gibt es echte (Wahl-)Prognosen, die zu einem wesentlichen Teil auf Umfragedaten basieren. Dabei wird mittels einer Modellrechnung, in die die Ergebnisse mehrerer Umfragen sowie externe Informationen einfließen, eine Vorhersage zum Ausgang eines zukünftigen Ereignisses (Wahlergebnis) gemacht.

3) Zielgruppe der Untersuchung (Grundgesamt)

In der quantitativen Markt- und Sozialforschung werden mittels repräsentativer Stichproben Aussagen über die Gesamtheit oder bestimmte Teile der Bevölkerung (bzw. Unternehmen) gemacht. Um im Rahmen einer Umfrage überhaupt eine angemessene Auswahl der zu befragenden Personen (bzw. Unternehmen) treffen zu können, ist eine eindeutige Definition der Zielgruppe der Untersuchung in sachlicher, regionaler und zeitlicher Hinsicht notwendig. Ohne diese Definition der Grundgesamtheit ist weder eine Beurteilung der Stichprobenqualität noch der Aussagefähigkeit und Tragweite der Forschungsergebnisse möglich. Deshalb muß sowohl im Forschungsbericht für den Auftraggeber als auch bei einer Veröffentlichung der Umfrageergebnisse die Zielgruppe der Untersuchung genannt werden.

4) Zahl der befragten Personen

Die Angabe der Zahl der befragten Personen ist notwendig, um die Stichprobenqualität beurteilen zu können; allerdings nicht unter dem Aspekt der repräsentativen Abbildung der Zielgruppe der Untersuchung, sondern hinsichtlich der Meßgenauigkeit der Umfrageergebnisse. Die der Berechnung der statistischen Fehlertoleranz der Ergebnisse (vgl. dazu auch den folgenden Abschnitt) zugrundeliegenden Annahmen können erst ab einer bestimmten Größe der Stichprobe bzw. der einzelnen Teilstichproben als gegeben unterstellt werden. Ähnliches gilt für die Anwendung multivariater Verfahren bei der mathematisch-statistischen Analyse der erhobenen Daten.

Um die Meßgenauigkeit und die Angemessenheit der multivariaten Analysen von Umfrageergebnissen beurteilen zu können, ist die Zahl der in einer Umfrage tatsächlich befragten Personen eine notwendige Information, die deshalb weder im Forschungsbericht noch bei der Veröffentlichung der Untersuchungsergebnisse fehlen darf. Dies gilt umso mehr, als es für bestimmte Bereiche der

Umfrageforschung - beispielsweise für Werbeträger-Analysen - auch wissenschaftliche Konventionen zur Mindestgröße von Stichproben gibt.⁷

5) Statistische Fehlertoleranz der Ergebnisse

Die in der Markt- und Sozialforschung durchgeführten Umfragen sind in der Regel Stichproben aus einer Grundgesamtheit. Deshalb sind ihre Ergebnisse zwangsläufig mit einer statistischen Fehlertoleranz behaftet. Die Ergebnisse von Stichproben schwanken mit einer angebbaren Genauigkeit und Wahrscheinlichkeit, d.h. dem sogenannten Konfidenzintervall, um den „wahren“ Wert in der Grundgesamtheit. Die Größe der statistischen Fehlertoleranz von Umfrageergebnissen ist in erster Linie abhängig vom Stichprobenverfahren (geschichtete und/oder mehrstufige Stichprobe) von der Stichprobengröße, d.h. der Zahl der tatsächlich befragten Personen, und dem ermittelten Anteilswert des Merkmals in der Stichprobe.

Wenn beispielsweise durch eine Umfrage mit 2.000 befragten Personen kurz vor einer Wahl für eine bestimmte Partei ein Anteilswert von fünf Prozent ermittelt wird und diese Partei bei der folgenden Wahl tatsächlich einen Stimmenanteil zwischen 4,2 und 5,8 Prozent erzielt, dann handelt es sich dabei nach mathematisch-statistischen Kriterien um eine gute Prognose des Wahlergebnisses. Berücksichtigt man allerdings das deutsche Wahlsystem, dann wird deutlich, daß sich das politische Schicksal dieser Partei unabhängig von der statistischen Fehlertoleranz von Umfrageergebnissen vollzieht.

6) Untersuchungszeitraum

Soweit die in Umfragen ermittelten Einstellungen und Meinungen auf den grundlegenden Wertorientierungen der befragten Personen basieren, verändern sie sich im Zeitablauf nur langsam. Soweit sie aber (auch) durch aktuelle Ereignisse beeinflusst werden, sind auch kurzfristig erhebliche Schwankungen möglich. Beispielsweise kann die nur wenige Wochen zuvor ermittelte politische Stimmung sich bis zum Zeitpunkt der Wahl noch deutlich verändern. Umfrageergebnisse können also nur für einen bestimmten Zeitraum aktuelle Gültigkeit beanspruchen. Dieses „Verfallsdatum“ ist unter anderem abhängig vom konkreten Untersuchungsgegenstand. Zum Beispiel weisen die in einer Umfrage ermittelten grundlegenden Einstellungen zu Beruf, Ehe, Familie und Religi-

⁷ vgl. ZAW 1994, S. 28f

on gewöhnlich einen längeren Gültigkeitszeitraum auf als Parteipräferenzen und Wahlabsichten. Das „Verfallsdatum“ von Umfrageergebnissen macht es zur Beurteilung ihrer Aussagekraft unumgänglich, den Zeitraum anzugeben, in dem die Umfrage durchgeführt wurde.

7) Stichproben-Methode

Durch die Angabe der Stichproben-Methode soll nicht in jedem Forschungsbericht und jeder Veröffentlichung von Umfrageergebnissen der alte wissenschaftliche Streit „Random versus Quota“ wieder neu aufgelegt werden. Die Berechnung der statistischen Fehlertoleranz ist aber strenggenommen nur bei Zufallsstichproben möglich. Die Meßgenauigkeit der Umfrageergebnisse wird unter anderem auch durch das konkrete Verfahren der Zufallsauswahl beeinflusst. Eine Befragung spezieller, eng abgegrenzter Gruppen der Bevölkerung ist mit einem vertretbaren finanziellen Aufwand häufig nur auf der Grundlage einer Quotenauswahl zu realisieren. Informationen zur Stichproben-Methode sind deshalb zur Beurteilung der Aussagekraft und Tragweite von Umfrageergebnissen unerlässlich.

8) Erhebungsverfahren

Umfragen der Markt- und Sozialforschung können mittels persönlich-mündlicher, schriftlicher oder telefonischer Interviews durchgeführt werden. Diese drei grundlegenden Verfahren der Datenerhebung weisen jeweils spezifische Vorteile auf, sind aber auch mit jeweils spezifischen Problemen behaftet. Die Entscheidung für ein bestimmtes Erhebungsverfahren sollte deshalb in erster Linie an seiner wissenschaftlich-methodischen Angemessenheit im Hinblick auf den Forschungsgegenstand und die Untersuchungskonzeption erfolgen.

Den Anwendungsmöglichkeiten jedes Erhebungsverfahrens sind bestimmte Grenzen gesetzt. Beispielsweise war nach der politischen Wiedervereinigung Deutschlands in den fünf neuen Bundesländern die Durchführung telefonischer Umfragen wegen der geringen Telefondichte in der Bevölkerung und der damit zusammenhängenden methodischen Probleme der Repräsentativität zunächst nicht möglich. Das hat sich erst geändert nachdem sich die Telefondichte dem hohen Niveau in den alten Bundesländern weitgehend angenähert hat. Die Angabe des Erhebungsverfahrens im Forschungsbericht und bei der Veröffentlichung von Umfrageergebnissen ist deshalb als Bestandteil des „Steckbriefs“ einer Umfrage unverzichtbar.

9) Stichproben-Ausschöpfung

Nicht bei allen für eine Umfrage ausgewählten Adressen bzw. Telefonnummern kann ein Interview realisiert werden, weil bestimmte Adressen oder Telefonnummern nicht zur Grundgesamtheit der Untersuchung gehören. Das ist beispielsweise der Fall, wenn es sich um die Adresse oder Telefonnummer eines Privathaushalts handelt, in dem ausschließlich ausländische Personen leben, die Zielgruppe der Untersuchung aber die in Privathaushalten lebende deutsche Bevölkerung ist. Man bezeichnet diese und ähnliche Fälle als qualitätsneutrale Ausfälle, weil sie die Repräsentativität der Stichprobe nicht beeinflussen.

Aber auch bei der sogenannten Netto-Stichprobe, d.h. der um die qualitätsneutralen Ausfälle bereinigten Stichprobe, kann nicht bei allen Adressen bzw. Telefonnummern ein Interview realisiert werden. In einigen Fällen sind die zu befragenden Personen trotz mehrmaliger Kontaktversuche nicht anzutreffen; in anderen Fällen ist die Zielperson im Haushalt nicht bereit, ein Interview zu geben, oder der Kontakt zur Zielperson wird durch eine andere Person im Haushalt verhindert. In all diesen und ähnlichen Fällen spricht man von relevanten oder systematischen Ausfällen, weil sie die Repräsentativität der Stichprobe beeinflussen können.

Deshalb ist es erforderlich, daß im Forschungsbericht die verschiedenen Ausfallgründe in einem detaillierten Ausschöpfungsprotokoll quantifiziert und dargestellt werden. Bei der Veröffentlichung von Umfrageergebnissen sollte zumindest die sogenannte Ausschöpfungsrate als Anteil der relevanten Ausfälle an der Netto-Stichprobe angegeben werden.

Die differenzierte Darstellung der verschiedenen qualitätsneutralen und relevanten Ausfallgründe gewinnt noch durch die Tatsache an Bedeutung, daß im Bereich der Werbeträger-Analysen die Höhe der zu erreichenden Ausschöpfungsrate auf siebzig Prozent festgelegt wurde.⁸ Dabei handelt es sich allerdings um eine Konvention, die nicht durch wissenschaftliche Kriterien begründet und deshalb auch nicht in allen Bereichen der Umfrageforschung gleichermaßen akzeptiert wird. Beispielsweise steht man dieser Festlegung in der empirischen Sozialforschung durchaus skeptisch gegenüber.

⁸ ZAW 1994, S. 20 ff.

10) Zahl der eingesetzten Interviewer

Bei persönlich-mündlichen und telefonischen Interviews, die in Deutschland gegenwärtig zusammen rund neunzig Prozent der in der quantitativen Umfrageforschung durchgeführten Interviews ausmachen, werden die Daten im Rahmen einer Kommunikation zwischen dem Interviewer und dem Befragten erhoben. In der Methodendiskussion der Markt- und Sozialforschung wird unter anderem auch diskutiert, ob und gegebenenfalls in welchem Umfang die persönlichen Merkmale und das situative Verhalten des Interviewers das Antwortverhalten des Befragten beeinflussen (können). Ein weiteres in diesem Zusammenhang diskutiertes Problem ist die Möglichkeit, daß die Auswahl der zu befragenden Personen trotz klarer Auswahlvorschriften daneben auch interviewerspezifischen Einflüssen unterliegt.

Diese möglichen Interviewereinflüsse sollen - soweit sie nicht durch eine entsprechende Auswahl, Schulung und Kontrolle der Interviewer sowie exakte Vorschriften zur Auswahl der Befragungspersonen und zur Durchführung des Interviews ausgeschaltet werden können - durch den Einsatz einer ausreichend großen Zahl von Interviewern minimiert werden. Allerdings stehen hier wissenschaftliche Qualitätsansprüche und eventuelle Zeit- und Budgetrestriktionen des Auftraggebers in einem grundsätzlichen Spannungsverhältnis. Die Zahl der zur Durchführung einer Umfrage eingesetzten Interviewer sowie die Angaben zur Häufigkeit und Verteilung der durchgeführten Interviews auf die verschiedenen Interviewer geben Aufschlüsse über das Ausmaß möglicher Interviewereinflüsse. Deshalb sollten diese Informationen im Forschungsbericht für den Auftraggeber und möglichst auch bei der Veröffentlichung von Umfrageergebnissen nicht fehlen.

11) Methode und Ergebnis der Interviewerkontrollen

In den 'Standards zur Qualitätssicherung in der Markt- und Sozialforschung' werden geeignete Verfahren zur Kontrolle der Interviewer gefordert, um sicherzustellen, „daß durch - bewußt oder unbewußt - nicht korrekt durchgeführte Interviews keine signifikante Verfälschung der Untersuchungsergebnisse erfolgt. ... Methode und Häufigkeit der Kontrollen sind zu dokumentieren, damit ein Auftraggeber sich von den institutsspezifischen Interviewerkontrollen ein Bild machen kann“.⁹

⁹ ADM 1998, S. 38

Dieser qualitätsrelevanten Verpflichtung der Institute entspricht die Forderung, Methode und Ergebnis der Interviewerkontrollen im Forschungsbericht für den Auftraggeber anzugeben. Auch bei der Veröffentlichung von Umfrageergebnissen im Rahmen von wissenschaftlichen Publikationen sollten Informationen zu den durchgeführten Interviewerkontrollen selbstverständlich sein. Dagegen dürfte diese Angabe bei sonstigen Veröffentlichungen von Umfrageergebnissen den möglichen Umfang des „Steckbriefs“ der zugrundeliegenden Untersuchung häufig überschreiten.

12) Fragebogen bzw. Fragestellung

Die Beurteilung, ob die Umsetzung des der Untersuchung zugrundeliegenden Forschungsproblems in die entsprechenden Forschungsfragen gelungen ist, und die Bewertung, ob die Interpretation der Umfrageergebnisse mit den tatsächlich ermittelten Sachverhalten übereinstimmt, ist ohne Kenntnisse der genauen Fragestellungen nicht möglich. Zudem können in vielen Fällen die Forschungsergebnisse auch durch kleine Änderungen im Wortlaut der gestellten Fragen sowie deren Reihenfolge beeinflusst werden. Bei der Veröffentlichung von Umfrageergebnissen sind deshalb die zugrundeliegenden Fragestellungen genau anzugeben. Außerdem sind Angaben darüber zu machen, welche Antwortmöglichkeiten den Befragten vorgegeben wurden, oder ob es sich um offene Fragen ohne feste Antwortvorgaben handelte. Der Forschungsbericht für den Auftraggeber der Untersuchung sollte darüber hinaus den für die Umfrage verwendeten Fragebogen enthalten.

13) Gewichtungsverfahren

Gewichtungen der in Umfragen erhobenen Daten können aus verschiedenen Gründen notwendig sein, um zwangsläufig entstehende oder bewußt herbeigeführte Disproportionalitäten der Stichprobe nachträglich wieder auszugleichen, d.h. den entsprechenden Strukturen in der Grundgesamtheit anzupassen. Das ist beispielsweise der Fall, um die zumeist haushaltsbasierten Stichproben bei allgemeinen Bevölkerungsumfragen in Personen-Stichproben zu transformieren und damit die in Abhängigkeit von der Haushaltsgröße unterschiedlichen Auswahlwahrscheinlichkeiten von Personen auszugleichen.

Eine Gewichtung der erhobenen Daten ist auch notwendig, wenn im Rahmen eines disproportionalen Stichprobendesigns die Stichprobenanteile einzelner Bevölkerungsgruppen bewußt überhöht wurden, beispielsweise um auch kleinen Bevölkerungsgruppen angehörende Personen in ausreichender Zahl befra-

gen zu können. Darüber hinaus wird häufig noch eine sogenannte Ausfallgewichtung durchgeführt, um die Abweichungen der Stichprobe von den Strukturen der Grundgesamtheit auszugleichen, die sich aus der Tatsache ergeben, daß die durch fehlende Teilnahmebereitschaft und Erreichbarkeit bedingten Ausfälle in der Netto-Stichprobe sich nicht proportional über die verschiedenen Bevölkerungsgruppen verteilen.

Vor allem in der empirischen Sozialforschung wird die Ausfallgewichtung nicht nur zustimmend sondern auch mit Skepsis gesehen, weil man dadurch unkontrollierbare Auswirkungen auf die Analyse der erhobenen Daten befürchtet. Um einerseits die Aussagefähigkeit der Umfrageergebnisse beurteilen und andererseits die Gefahr unkontrollierbarer Einflüsse besser abschätzen zu können, sind Informationen über die Art der eingesetzten Gewichtungsverfahren und der berechneten Gewichtungsfaktoren notwendig.

III. Interpretation von Umfrageergebnissen

Forschungsberichte für den Auftraggeber und Veröffentlichungen von Umfrageergebnissen enthalten neben der Darstellung häufig auch eine Interpretation der Ergebnisse und eventuell darauf begründete Empfehlungen. Die Durchführung von Umfragen der Markt- und Sozialforschung ist eine Tätigkeit, die nach anerkannten wissenschaftlichen Regeln durchgeführt wird und deshalb die durch Artikel 5 des Grundgesetzes für die Bundesrepublik Deutschland garantierte Forschungsfreiheit genießt. Sofern sich Interpretationen und Empfehlungen aus dem der Umfrage zugrundeliegenden Erkenntnisinteresse bzw. dem Problem des Auftraggebers ableiten und von den Forschungsergebnissen gestützt werden, sind sie ein integraler Bestandteil angewandter empirischer (Auftrags-)Forschung.

Nicht durch Umfrageergebnisse gestützte Interpretationen und Empfehlungen sind dagegen kein Bestandteil empirischer Forschung und deshalb auch kein Bestandteil des Forschungsberichts oder wissenschaftlicher Publikationen. Soweit Veröffentlichungen von Umfrageergebnissen Interpretationen und Empfehlungen enthalten, die nicht durch empirische Forschung gestützt werden, stellen sie persönliche Ansichten und Überzeugungen dar, die als solche kenntlich gemacht und von der Darstellung und Interpretation der Umfrageergebnisse getrennt werden müssen.