Vorsicht Umfragedaten!

Rolf Porst[1]

In dreizehn Beiträgen hat sich die Serie „Schüler forschen“ bisher mit der Frage beschäftigt, wie man Methoden der empirischen Sozialforschung in den Sozialkundeunterricht einbringen kann. In jedem dieser Beiträge wurde, bewusst vereinfachend oder sagen wir besser: auf hohem Niveau allgemeinverständlich, ein Schritt oder ein Aspekt der Durchführung von Studien unter Verwendung von Methoden der empirischen Sozialforschung vorgestellt. Die Beiträge sollten LehrerInnen als Anregung und Handreichung dafür dienen, wie empirische Studien von SchülerInnen durchgeführt werden können, die zwar sicher nicht alle Standards der empirischen Sozialforschung erfüllen (können), aber gleichwohl sozialwissenschaftlich (einigermaßen) vertretbar sind. Die rasche Fokussierung der Serie „Schüler forschen“ auf die Vorbereitung, Durchführung und Auswertung von *Umfragen* war der Bedeutsamkeit dieser Form der empirischen Datenerhebung geschuldet.

Die Serie war dann erfolgreich, wenn es uns gelungen ist, Ihnen als LehrerInnen praktikable und effiziente Hilfestellungen für die Durchführung Ihrer schulischen Befragungsprojekte an die Hand zu geben. Wenn wir darüber hinaus ein wenig dazu beigetragen haben, dass Sie sensibler geworden sind gegenüber Umfragen und ihren Ergebnissen – umso besser. Jetzt stellen wir – advocatus diaboli – das Ganze wieder in Frage. Aber nur als Denkmodell.

Der vorliegende Beitrag beschäftigt sich deshalb mit den Problemen und Schwächen von Befragungen, mit vorsätzlichen und fahrlässigen Fehlerquellen und mit der Möglichkeit, Befragungsergebnisse gezielt zu manipulieren oder unbeabsichtigt zu verfälschen.

Rolf Porst
Markt- und Sozialforscher

Betrachten wir zunächst ganz kurz sechs mögliche (beileibe nicht alle) Fehlerquellen bei der Vorbereitung, Durchführung und Auswertung von Umfragen.

Fehlerquellen bei der Vorbereitung, Durchführung und Auswertung von Befragungen

Fehlerquelle 1: Die Ergebnisse, die uns zur Kenntnis kommen, könnten gefälscht sein.

Großes Aufsehen, auch in der nicht-wissenschaftlichen Öffentlichkeit, erregte der südkoreanische Arzt und Wissenschaftler Hwang Woo-suk. Er hatte 2004 eine Studie veröffentlicht, der zufolge es ihm gelungen war, einen menschlichen Embryo zu konstruieren und aus diesem Stammzellen abzuleiten. Am 17. Juni 2005 wurden seine Forschungsergebnisse in der renommierten Zeitschrift „Science" veröffentlicht – im Dezember 2005 stellte sich heraus, dass es sich bei den Ergebnissen seiner Arbeit um eine Totalfälschung handelte. Ähnliche „Meisterwerke" legten z.B. die niederländischen Sozialpsychologen Diederik Stapel und Dirk Smeesters vor. Stapel musste 2011 zugeben, die Daten seiner in renommierten Fachzeitschriften erschienenen Veröffentlichungen frei erfunden zu haben, Smeesters gestand Anfang 2012, dass er Daten gefälscht habe, um in seinen Studien statistische Signifikanz nachzuweisen.

Natürlich sind das nicht alle bekannten (und schon gar nicht alle unentdeckten) Fälle von Betrug in der wissenschaftlichen Forschung. Aber wenn es nicht nur in der Sozialpsychologie, sondern auch in der Medizin (der Biologie, der Chemie oder der Mikroelektronik, um nur einige der „harten" Professionen zu nennen, innerhalb derer Forschern Betrug nachgewiesen werden konnte) möglich ist, Daten zu fälschen oder frei zu erfinden – warum nicht auch in der sozialwissenschaftlichen Umfrageforschung?

Ich konnte im Oktober 2016 bei einer Befragung rheinland-pfälzischer Eltern von schulpflichtigen Kindern ermitteln, dass 56,7% der Betroffenen der Ansicht sind, dass in den Schulen wieder Samstagsunterricht eingeführt werden sollte – aber wissen Sie, ob das stimmt oder ob ich diesen Befund nicht frei erfunden habe?[2]

Fehlerquelle 2: Ergebnisse könnten bewusst und gezielt produziert worden sein.

Befragungen werden häufig durchgeführt, um politische oder gesellschaftliche Entscheidungen oder auch nur eine bestimmte argumentative Position zu untermauern. Wenn ich also z.B. eine rechtspopulistische Partei wäre und zeigen möchte, dass die Mehrheit der Deutschen für die Schließung der Grenzen für Flüchtlinge sei, könnte ich eine bundesweite Befragung durchführen (lassen) und – vorausgesetzt ich finde einen entsprechend unseriösen Anbieter von Befragungen – dafür sorgen, dass bei der Befragung genau das herauskommt, was ich mir wünsche. Allzu schwer wäre das nicht. Ich muss mich z.B. bei der Fragebogenkonstruktion nur an die kognitionspsychologischen Grundlagen der Befragung (vgl. Porst 2015a) oder/und die einschlägigen Experimente zur Manipulation von Befragtenverhalten erinnern. Dann könnte ich durch gezielte Frageformulierungen oder durch Nutzung von Fragensukzessionseffekten bzw. allgemeiner: von Kontexteffekten (vgl. Porst 2015b) Ergebnisse so manipulieren, dass sie meine Position stärken. Etwas überspitzt formuliert: Mit dem Wissen um die Möglichkeiten zur Manipulation von Befragungspersonen könnte ich Ihnen (oder der rechtspopulistischen Partei) jedes Befragungsergebnis liefern, das Sie (oder die rechtspopulistische Partei) haben möchten.

Da wir in den meisten Fällen von den Medien zu veröffentlichten Befragungsergebnissen keine Informationen zur Durchführung der dazugehörigen Befragung geschweige denn den Fragebogen dazu erhalten, ist grundsätzlich erst einmal nicht auszuschließen, dass Ergebnisse bewusst und gezielt produziert worden sind. Die Regel dürfte das allerdings nicht sein. Die meisten – vor allem bundesweiten – Umfragen werden in Deutschland von kommerziellen Umfrageinstituten durchgeführt, die der professionellen Kontrolle durch die Konkurrenz unterliegen. Natürlich gibt es Institute, die z.B. einer der politischen Parteien näher stehen als anderen, aber das Risiko eines Imageverlustes im Falle nachgewiesener Manipulationen ist so groß, dass wir im Normalfall keine bewusst manipulierten Umfrageergebnisse fürchten müssen. Viel größer als das Risiko manipulierter Daten ist das Risiko schlechter Arbeit.

Fehlerquelle 3: Schlechte Arbeit

Wenn wir veröffentlichte Umfrageergebnisse zur Kenntnis nehmen, können wir praktisch nie ersehen, wie „sauber" die Daten und damit auch die Ergebnisse einer Befragung sind. Im gesamten Prozess der Vorbereitung, Durchführung und Auswertung einer Umfrage können die unterschiedlichsten Akteure an den verschiedensten Punkten schlechte Arbeit abliefern. Mögliche Schwachstellen sind zunächst die ForscherInnen selbst, wenn sie sich z.B. für ihre Fragestellung ungeeigneter Methoden bedienen oder wenn Schwächen oder gar Fehler in der Studienanlage und im Studiendesign vorhanden sind – was durchaus vorkommen kann. Mögliche Schwachstellen sind weiterhin die Personen, welche das Studiendesign in die Umfragerealität umsetzen; Fehler hier wären z.B. Stichprobenverfahren, die dem Studienziel nicht angemessen sind, Fehler bei der Fragebogenkonstruktion oder Fehler bei der Datenerfassung und Datenauswertung. Sofern die Umfragen interviewerbasiert stattfinden, stellen die InterviewerInnen die nächste mögliche Fehlerquelle dar: Sie könnten vor Ort die falschen Zielpersonen auswählen, aufgrund der Gegebenheiten vor Ort gar nicht erst versuchen, definierte Zielpersonen zu befragen oder bei der Befragung regelwidrig auf das Antwortverhalten der befragten Personen einwirken. Und schließlich die Befragungspersonen selbst, indem sie z.B. Fragen nicht oder nicht ehrlich beantworten oder, was noch schlimmer ist, Fragen beantworten, die sie nicht verstanden oder zu deren Gegenstand sie eigentlich gar keine Meinung haben. Und ganz am Ende des Forschungsprozesses können Auswertungsverfahren, die den Daten nicht angemessen sind und Interpretationen, die den Auswertungsergebnissen nicht angemessen sind, als mögliche Fehlerquellen dazu führen, dass schlechte Arbeit abgeliefert wird.

Fehlerquelle 4: Mangelnde Transparenz

Wie bereits zu erkennen war, bleiben uns die Hintergründe und das Zustandekommen von Umfrageergebnissen weitestgehend im Dunkeln, so dass wir kaum in der Lage sind, deren Belastbarkeit zu erkennen. Wir lesen z.B.

„Bei einer Umfrage, die das Institut xyz im Januar 2017 für die Zeitschrift zyx durchgeführt hat, waren 47,6% der Befragten der Ansicht, dass Angela Merkel ihre Arbeit als Bundeskanzlerin sehr gut macht."

Aber: Was sagt uns das? Nicht wirklich viel. Wir wissen nichts über die Qualität der Fragen, über die Auswahl der Zielpersonen, über den Befragungsmodus usw. Aber wäre die folgende Aussage besser?

„Bei einer repräsentativen Umfrage, die das Institut xyz im Januar 2017 für die Zeitschrift zyx unter 3.000 Bundesbürgern telefonisch durchgeführt hat, waren 47,6% der Befragten der Ansicht, dass Angela Merkel ihre Arbeit als Bundeskanzlerin sehr gut macht".

Mehr Informationen, aber: Sagt uns das wirklich mehr? Wissen wir jetzt mehr über die Qualität der Umfrage? Ich meine: Nein, nicht wirklich. Auf der Basis der üblicherweise bei der Darstellung von Umfrageergebnissen in den Medien veröffentlichten Informationen zur Umfrage selbst wird es uns in den allermeisten Fällen nicht gelingen, die Seriosität der Ergebnisse zu erkennen. Der einzige Indikator, der mit einer gewissen Sicherheit Rückschlüsse auf die Qualität von Umfrageergebnissen zulässt, ist das durchführende Institut. Aber um das beurteilen zu können, müssten wir schon sehr vertraut mit der „Szene" sein, was die meisten von Ihnen ganz sicher nicht sind. Und selbst wenn das durchführende Institut einen guten Namen hat, kann nicht von vornherein ausgeschlossen werden, dass auch dort einmal das eine oder andere schiefgelaufen sein könnte. Aber ein wenig mehr Hoffnung auf gute und seriöse Arbeit können wir schon haben, wenn bestimmte Institute dafür verantwortlich sind. Sie werden verstehen, dass ich an dieser Stelle keine Institute benennen werde, keine „guten", aber auch keine „schlechten".

Fehlerquelle 5: Bewusste Datenmanipulation

Erinnern Sie sich an die Bundestagswahl 2005? Wie kam Frau Merkel damals am Wahlabend eigentlich dazu, angesichts von 35,2% der Zweitstimmen von einem eindeutigen Auftrag an die CDU/CSU zur Regierungsbildung zu sprechen? Und wie kam Herr Schröder am selben Abend dazu, das Gleiche für die SPD zu reklamieren, die sogar nur 34,3% der Zweitstimmen auf sich vereinigt hatte? Angesichts der Zahlen und angesichts der Professionalität der Akteure kann man hier nur von einer bewussten Fehlinterpretation der Ergebnisse sprechen.

Aber bewusste Fehlinterpretationen gibt es natürlich nicht nur in der Politik, sondern auch im Umgang mit Umfragedaten. Eine bewusste Fehlinterpretation (mindestens!) liegt z.B. dann vor, wenn sich die berichteten Ergebnisse von ihrem Zustandekommen unterscheiden. Ein Beispiel:

Im Fragebogen steht die Frage „Sind Sie für eine Verschärfung des Asylrechts in Deutschland, oder sollte es in dieser Frage so bleiben, wie es heute ist?". In einem Presseartikel dazu lesen wir, dass sich 53% der Befragten für eine Verschärfung ausgesprochen hätten, 41% für die Beibehaltung der derzeitigen Regelung – und 6% plädierten für eine Liberalisierung. Sieht nach sehr wenig aus, ist aber sehr viel, wenn man bedenkt, dass die Kategorie „für eine Liberalisierung" gar nicht im Fragebogen enthalten war, die 6% also spontan zustande gekommen sind. Und wenn wir dann noch wissen, dass Befragungspersonen davon ausgehen, dass nur die Dinge, die gefragt werden, für die Forschenden interessant sind, alles was nicht gefragt wird deshalb von den Befragungspersonen auch nicht berichtet wird (s. dazu Porst 2014: 60), sind 6% spontane Nennungen für eine im Fragebogen nicht vorhandene Antwortmöglichkeit nicht schlecht.

Fehlerquelle 6: Unbeabsichtigte Fehlinterpretationen

Unbeabsichtigte Fehlinterpretationen liegen dann vor, wenn man bei der Interpretation von Umfragedaten zwar ehrlich sein will, es aber nicht schafft. Ein typischer Fall ist die „Scheinkorrelation", bei der zahlenmäßig ein Zusammenhang zwischen zwei Variablen nachgewiesen wird, der aber tatsächlich nicht kausal bedingt ist, sondern

durch eine dritte Variable hergestellt wird. Ein häufig zitiertes Beispiel finden wir schon ganz früh bei Friedrichs (1973: 390), der den Zusammenhang zwischen Geschlecht und Unfällen mit dem Auto anspricht: Frauen verursachen weniger Unfälle als Männer. Fahren Frauen vorsichtiger oder besser? Keineswegs. Der Zusammenhang verschwindet, wenn man die Anzahl der gefahrenen Kilometer kontrolliert. Dann nämlich verursachen Frauen genauso viele Unfälle wie Männer.

Ein sehr beliebtes Beispiel für Scheinkorrelationen findet sich in dem lesenswerten Buch von Andreas Diekmann (1995: 57f.), die „Storchkorrelation": Je mehr Störche in einer Region umherfliegen, umso höher ist dort die Geburtenrate. Also bringt der Storch doch die Kinder? Nein, keineswegs, aber Störche bevorzugen nun mal ländliche Regionen, und in ländlichen Regionen ist die Geburtenrate nun mal höher als in urbanen Gegenden. Oder wie Diekmann (1995: 58) schreibt: „Produziert wird die Scheinkorrelation, die scheinbare Bestätigung der ‚Klapperstorch-Hypothese' durch die wachsende Urbanisierung…, die sowohl mit einem Rückgang der Störche… als auch mit sinkenden Geburtenzahlen verbunden ist".

Bleiben noch der Zusammenhang zwischen Schuhgröße und Einkommen: „Wer auf großem Fuß lebt, verdient mehr" (Diekmann 1995: 57), oder: „Je mehr Feuerwehrleute bei einem Brand eingesetzt werden, umso höher ist der Brandschaden" (Lazarsfeld, zit. nach Diekmann 1995: 603). Überlegen Sie selbst, wie sich diese „Korrelationen" erklären lassen. Und woran liegt es, dass Personen, die hauptsächlich Rockmusik hören, höher springen können als Personen, die hauptsächlich Volksmusik hören" – ganz sicher nicht an der Musik.

Halten wir zunächst einmal zusammenfassend fest, dass der Prozess der Vorbereitung, Durchführung und Auswertung von Befragungen doch recht fehleranfällig sein kann. Die beschriebenen Fehlerquellen müssen nicht wirksam werden, aber sie können es. Aber damit sind noch lange nicht alle Risiken beschrieben, die im Zusammenhang mit Befragungen bestehen. Hervorragende Möglichkeiten, Umfrageergebnisse zu manipulieren oder gezielt zu produzieren, ergeben sich bei der Konstruktion der Fragebogen.

Wie man mit dem Fragebogen Befragungsergebnisse gezielt produzieren kann

Wenn wir alle uns bekannten Möglichkeiten, durch Manipulationen im Fragebogen Ergebnisse bewusst zu produzieren, im Detail vorstellen wollten, möglichst mit Beispielen, könnten wir die Serie „Schüler forschen" damit (fast) unendlich weiterführen. Wir können deshalb im Folgenden nur Andeutungen dazu machen, aber nicht wirklich ausführen, wie wir manipulieren könnten, wenn wir wollten; wer Näheres dazu erfahren möchte, kann das im „Fragebogen-Buch" (Porst 2008, 4. Auflage 2014) nachlesen. Aber: Wir nutzen das Wissen um Manipulationsmöglichkeiten natürlich, um Manipulationen zu verhindern. Schließlich wollen wir „saubere Arbeit" abliefern, wenn wir Befragungen durchführen. Listen wir also im Folgenden einige Manipulationsmöglichkeiten auf:

– Bei Fragebogenfragen kommt es sehr auf die Formulierung an. Durch Verwendung positiv oder negativ geladener Begriffe können Sie Einfluss auf das Antwortverhalten nehmen. Auch sprachliche „Kleinigkeiten" können großen Einfluss auf das Antwortverhalten nehmen.

- Fragen sind in hohem Maße kontextabhängig. Durch Schaffung eines „geeigneten" Kontextes der Frage können Sie die Antworten der Befragungspersonen in die von Ihnen gewünschte Richtung lenken.
- Bei der Verwendung halboffener Fragen, also Fragen mit einem vorgegebenen Set von Antwortkategorien und einer offenen Restkategorie (vgl. Porst 2015: 244), können Sie Ihrem „Favoriten" einen Vorteil verschaffen, wenn Sie ihn bei den vorgegebenen Kategorien platzieren.
- Personen verwenden die vorgegebenen Antwortmöglichkeiten dazu, den Kontext einer Frage zu „erkennen", ihren Sinn zu „interpretieren" und den Spielraum für eine angemessene Antwort ab- und einzugrenzen. Insofern können Sie bei der Vorgabe von Antwortmöglichkeiten dafür sorgen, dass die Antwort in die von Ihnen gewünschte Richtung gelenkt wird.
- Antwortvorgaben beeinflussen die Wahrnehmung einer Frage durch die Befragungspersonen. Personen interpretieren Fragen in Abhängigkeit von den Antwortvorgaben.
- Befragungspersonen neigen dazu, sich positiv darzustellen und „negative" Urteile zu vermeiden. Wenn Sie mit Skalen arbeiten, deren einer Teil z.B. durch Minuszeichen (-5, -4,…) ins Negative verweist, werden Sie positivere Ergebnisse bekommen als bei neutralen Skalen, weil die Befragungspersonen mit negativen Ziffern negative Erfahrungen verbinden und den negativen Teil der Skala meiden.
- Sie können diesen Effekt noch verstärken, wenn Sie innerhalb der Skalen mit fließender Farbgebung arbeiten: Dunkel gefärbte Skalenpunkte werden vermieden, helle Farben locken eher zum Ankreuzen an.
- Die Antwort auf eine bestimmte Frage kann durch eine direkt vorhergestellte Frage und deren Beantwortung beeinflusst werden; wir sprechen von Sukzession- oder Folgeeffekten, ganz allgemein von Kontexteffekten. Durch solche Kontextbildung lenken Sie die Antworten in eine bestimmte Richtung.
- Beim direkten Vergleich zweier Objekte innerhalb einer Frage schneidet in den allermeisten Fällen das zuerst genannte Objekt besser ab; platzieren Sie Ihren Favoriten beim Vergleich vorne, wird er positiver bewertet als das zweite Vergleichsobjekt.
- Befragungspersonen unterstellen optisch präsentierten Befragungshilfen einen Sinn. Wenn Sie mit optischen Vorgaben arbeiten, können Sie durch die grafische Gestaltung der Vorgaben Einfluss nehmen auf das Antwortverhalten.

Wir können es bei dieser Aufzählung belassen; sie ist keineswegs vollständig, aber meines Erachtens beeindruckend. Alles worüber der Fragebogengestalter oder die Fragebogenkonstrukteurin verfügen müssen, sind gute Kenntnisse in den kognitionspsychologischen und kommunikationstheoretischen Grundlagen der Befragung. Dann stehen den Möglichkeiten, durch Fragebogendetails Ergebnisse zu manipulieren, in eine bestimmte Richtung zu lenken oder Ergebnisse gezielt zu produzieren, Tür und Tor offen. Oder noch härter formuliert: Ich produziere Ihnen mit meinen Fragebogen jedes Ergebnis, das Sie haben möchten. Theoretisch. Denn tatsächlich nutzen wir unsere Kenntnisse dazu, gezielte Manipulationen zu erkennen und zu vermeiden.

Kommen wir zu einem weiteren Punkt, der dazu beitragen könnte, unser Unbehagen an Befragungen zu stärken, den sogenannten „fiktiven Fragen".

Warum ist Werner Wilken bekannt, aber nicht beliebt?[3]

Fiktive Fragen sind Fragen, die sich mit Personen, Gegenständen oder Sachverhalten beschäftigen, die es faktisch nicht gibt, aber durchaus geben könnte.

Sie kennen Werner Wilken? Natürlich, Wilken ist ein recht bekannter Politiker, von welcher Partei wollen wir aus Gründen der Neutralität hier verschweigen. Dass Wilken ein bekannter Politiker ist, können wir bei Befragungen regelmäßig und mit einer gewissen Sicherheit erfahren. Aber wir stellen auch genauso regelmäßig fest, dass man mit der Politik von Wilken nicht übereinstimmt. Wie kommt das?

Den Politiker Werner Wilken gibt es nicht (Ähnlichkeiten mit lebenden oder verstorbenen politischen oder sonstigen Personen wären zufällig und völlig unbeabsichtigt), Werner Wilken ist ein „fiktiver Politiker". Fiktive Politiker wiederum sind eine Ausprägung der „fiktiven Fragen". Andere Beispiele:

Auf die Frage...

Do you favor or oppose the Metallic Metal Act?

...antworten 70% der Befragten, dass sie entweder für oder gegen dieses Gesetz seien (Gill 1947); noch immerhin 30,8% der Befragten waren für oder gegen den „Agricultural Trade Act of 1978" (Schuman und Presser 1981) und 26,4% der Befragten für oder gegen das „monetary control bill" (ebenda).

Oder: Das NZZ-Folio der Neuen Züricher Zeitschrift veröffentlichte in Heft 01/2006 die Ergebnisse einer Leserbefragung, bei der 189 von 1883 Teilnehmern (ca. 10 Prozent) angegeben hatten, dass ihnen unter den letzten zwölf Heften das NZZ-Folio mit dem Thema „Katastrophen" am besten gefallen habe.

Nur: Die genannten Gesetze wie auch das NZZ-Folio zum Thema „Katastrophen" gibt es nicht und hat es nie gegeben. Und da möchte ich gerne wieder zu Werner Wilken zurückkommen.

Mit schöner Regelmäßigkeit werden die Namen nicht-existenter, „fiktiver" PolitikerInnen in Itembatterien eingebaut, bei denen es um die Bekanntheit von PolitikerInnen und/oder um die Bewertung derer Politik und/oder die Übereinstimmung mit deren Politik geht.

Regelmäßig lernen wir daraus, dass die fiktiven Politiker durchaus recht bekannt sind (übrigens: Je höher die formale Schulbildung der Befragten, umso bekannter!), aber dass man mit ihnen nicht zufrieden ist bzw. mit ihrer Politik nicht übereinstimmt.

Wie ist das zu erklären? Erwartungsgemäß mit Kognitionspsychologie.

Befragungspersonen gehen grundsätzlich davon aus, dass Befragungen Sinn machen; ansonsten würde man sie ja wohl nicht durchführen. So wie die Befragung als Ganzes machen aus Sicht der Befragungspersonen aber auch die einzelnen Fragen Sinn; ansonsten würde man sie ja wohl nicht stellen.

Nun versetzen Sie sich doch mal in die Lage einer Befragungsperson, der man im Verlaufe einer Befragung (sagen wir im Januar 2017) die folgende Liste vorlegt:

Liste 23: Politiker und Politikerinnen in Deutschland

A) Frauke Petry
B) Winfried Kretschmann
C) Sigmar Gabriel
D) Sahra Wagenknecht
E) Werner Wilken
F) Angela Merkel
G) Horst Seehofer
H) Martin Schulz

Wir fragen zunächst, welche dieser Damen und Herren Politiker man kenne, und sei es auch nur dem Namen nach. Da werden unseren Befragungspersonen doch einige davon bekannt sein, und sie haben natürlich schon mal was von Frau Merkel oder Herrn Gabriel gehört, und von manchen anderen auch. Ausnahme: Werner Wilken. Den kennen die Leute eigentlich nicht. Aber: Wenn der auf einer Liste steht gemeinsam mit Merkel, Gabriel usw. muss er wohl ein „ranghoher" Politiker sein, andernfalls gebe es ja keinen Grund dafür, ihn genau auf diese Liste zu setzen.

Also werden die Befragten – sie wollen ja nicht zeigen, dass sie eigentlich keine Ahnung haben – tendenziell eher bereit sein anzugeben, dass sie – neben den anderen PolitikerInnen – auch Werner Wilken kennen. So weit so gut. Wäre bloß die Nachfrage nicht. Denn die lautet zum Beispiel:

Ich lese Ihnen jetzt noch einmal die Namen der Politiker vor, die Sie mir gerade genannt haben. Sagen Sie mir bitte zu jedem, inwieweit Sie mit der Politik dieser Politikerin bzw. dieses Politikers übereinstimmen.

...(oder so ähnlich) mit einer Skala von 1 = „Stimme überhaupt nicht mit der Politik dieses Politikers überein" bis 7 = „stimme voll und ganz mit der Politik dieses Politikers überein" (oder so ähnlich).

Jetzt haben die Befragungspersonen natürlich ein Problem, wenn sie gerade Werner Wilken als ihnen bekannten Politiker genannt haben.

Also überlegen sie sich Folgendes: An sich habe ich ja gerade (bei der Vorfrage) etwas geschwindelt, indem ich so tat, als wäre mir der Politiker Werner Wilken bekannt. Aber ich weiß natürlich ganz genau, dass ich ihn nicht kenne. Wenn der aber gemeinsam mit Merkel, Gabriel, Schulz usw. auf einer Liste steht, muss – sonst würde die Frage ja keinen Sinn machen – Werner Wilken auch ein „Hochkaräter" sein. Da ich ihn aber trotz seiner mutmaßlichen Wichtigkeit nicht kenne, liegt der Schluss nahe, dass es sich bei Wilken um einen Politiker vom äußeren rechten oder vom äußeren linken Rand des Parteienspektrums handeln muss (die sieht man im Fernsehen halt seltener oder so gut wie gar nicht) – und was liegt bei dieser Deutung der Sachlage näher als zu sagen, mit der Politik dieses Werner Wilken stimme man „überhaupt nicht überein"?

Natürlich werden viele Befragungspersonen auf den Namen Werner Wilken mit einem „kenne ich nicht" oder „nie davon gehört" reagieren. Das ist weniger interessant. Interessanter ist der recht hohe Anteil von Befragungspersonen, die auf fiktive Politiker oder fiktive Themen insgesamt mit inhaltlichen Antworten und damit vorgetäuschtem Wissen reagieren. Aus der Sicht des Fragebogenentwicklers am interessantesten aber ist, dass man dieses Ergebnis schon fast mit an Sicherheit grenzender Wahrscheinlichkeit prognostizieren (und natürlich auch: provozieren) kann.

Aber das ist längst noch nicht alles. Sehen wir uns ein Beispiel dafür an, dass Menschen Fragen beantworten, ohne auch nur ansatzweise zu verstehen, um was es in der Frage geht.

Das Ziel der Nationalsozialisten…

Dazu ein Beispiel zum Frageverständnis: Bei einem Pretest, an dem ich vor einigen Jahren bei – damals noch – GESIS-ZUMA beteiligt war, sollte folgendes Item bewertet werden[4]:

„Das Ziel der Nationalsozialisten war, durch den Krieg gegen die Sowjetunion den jüdischen Bolschewismus zu beseitigen".[5]

Unser Einwand damals, dass diese Aussage nicht geeignet sei, bei einer Allgemeinpopulation abgefragt zu werden, wurde erst akzeptiert, nachdem wir das Item zwanzig Testpersonen vorgelegt und gebeten hatten, mit einer Skala von 1 = „stimme überhaupt nicht zu“ bis 7 = „stimme voll und ganz zu“ zu bewerten. Das Interessante daran war, dass nur zwei der zwanzig Testpersonen angaben, sie könnten diese Bewertung nicht vornehmen; alle anderen 18 Testpersonen haben spontan einen Skalenwert vergeben, und mit Hilfe kognitiver Pretesttechniken konnten wir zeigen, dass alle diese 18 Testpersonen zwar eine Bewertung abgegeben, aber keine klare Vorstellung vom Inhalt der Aussage hatten.

Wir fragten nämlich die 18 Testpersonen, die die Aussage bewertet hatten, warum sie sich gerade für den von ihnen gewählten Skalenwert entschieden hätten.

Eine Auswahl der Begründungen für die Entscheidung für einen bestimmten Skalenwert zeigt uns, – und das ist eines der zentralen Probleme, mit denen sich die Umfrageforschung herumzuschlagen hat – dass Personen Fragen beantworten, ohne überhaupt zu wissen, um was es dabei geht[6] (in Klammern jeweils der gewählte Skalenwert):

- (4) „Das weiß ich nicht. Ich kann damit nichts anfangen.“
- (5) „Es war immer das Ziel der Nationalsozialisten, gegen die Juden zu sein.“
- (2) „Ich weiß es nicht, aber ich glaube es nicht. Habe die „2“ vergeben auf Grund meines Bauchgefühls“
- (6) „Ich vergebe ungern „1“ und „7“. Mit der „6“ möchte ich relativieren und ich glaube, dass es das Ziel der Nationalsozialisten war.“
- (4) „Bin mir relativ unsicher. Ich dachte, ich nehme einen Mittelweg. Die „4“ deshalb, weil ich es nicht weiß.“
- (1)„Weil ich mir einfach nicht vorstellen kann, dass die Nationalsozialisten ein anderes Ziel hatten. Ich vermute es, habe keinen blassen Schimmer.“

Oder ein anderes Beispiel:

Die kognitive Technik des comprehension probing ist bestens geeignet, um das Verständnis von bestimmten Begriffen zu ermitteln, aber auch dazu, zu erkennen, wenn Begriffe falsch verstanden werden. Betrachten wir dazu das folgende (warnende) Beispiel; zu überprüfen war das Verständnis des Begriffes „repräsentative Demokratie“:

„Volksbegehren und Volksentscheide sind eine notwendige Ergänzung der repräsentativen Demokratie.“

Unmittelbar nach der Beantwortung dieses Items wurde nachgefragt:

„Was – würden Sie sagen – ist das, eine „repräsentative Demokratie?“

Dass Sie den Begriff „repräsentative Demokratie“ zumindest in Fragebogen für Allgemeinpopulationen nicht verwenden sollten, zeigen die folgenden Antworten auf diese Frage:

- „Die Demokratie wird hier repräsentiert, sie wird eben repräsentiert, dargestellt.“
- „Eine Demokratie, die nach außen gerne gezeigt wird, weil sie in vielen Dingen gut ist und eine Demokratie mit Menschennähe.“
- „Was sich darstellt. Und Demokratie ist das, was wir haben. Die präsentiert sich halt.“

Sie erkennen selbst das Verständnisproblem, das mit dem Begriff „repräsentative Demokratie“ verbunden ist.

Zum Schluss: Warum – trotz allem – Umfragen notwendig und unersetzbar sind

Stellt sich, angesichts all der beschriebenen Fehlerquellen und Manipulationsmöglichkeiten, der Probleme und Schwierigkeiten, die mit der Vorbereitung, Durchführung und Auswertung von Befragungen verbunden sind, die Frage, ob Befragungen überhaupt Sinn machen oder ob wir nicht besser auf Befragungen verzichten sollten. Die Antwort ist eindeutig: *Ja, sie machen Sinn, und nein, wir sollten nicht auf Befragungen verzichten; sie sind notwendig und unersetzbar. Vor allem aber: Es gibt keine Alternativen dazu.*

Wenn wir wissen möchten, welche Einstellungen Menschen zu bestimmten Personen, Sachverhalten oder Gegenständen haben, müssen wir sie danach fragen. Man sieht Menschen nicht an, ob sie glauben, man brauche eine Familie, um wirklich glücklich zu sein. Man sieht Menschen nicht an, ob sie bei der nächsten Bundestagswahl die CDU wählen möchten. Man sieht Menschen nicht an, ob sie Angela Merkel gut finden oder nicht. Oder ob sie Donald Trump für einen Segen halten. Oder gerne in Sevilla leben möchten, usw., usw. Wenn wir all das wissen wollen, müssen wir die Menschen danach fragen. Und wenn wir herausfinden möchten, wie das bei vielen Menschen aussieht, müssen wir erst recht eine Befragung durchführen, um Antworten zu bekommen.

Aber wir sollten uns bei Befragungen nicht naiv verhalten, weder sie vorschnell ablehnen noch sie bedingungslos annehmen. Befragungen können nur dann taugen, wenn zur Redlichkeit und fachlichen Kompetenz aller Beteiligter ein Höchstmaß an Transparenz ihres Zustandekommens gefordert und realisiert wird. Zumindest bei den Befragungen, die Sie mit Ihren SchülerInnen in schulischen Projekten durchführen, sind Sie selbst dafür verantwortlich, dass alles sachgerecht abläuft. Wir hoffen, Ihnen mit der Serie „Schüler forschen“ ein wenig geholfen zu haben, diese Aufgaben und Anforderungen zu bewältigen.

Anmerkungen

1 Für die kritische Durchsicht des Manuskripts und Vorschläge zu seiner Verbesserung danke ich Ruth Holthof, Studienrätin im Fach Sozialkunde am Eleonoren-Gymnasium in Worms.

2 Ich habe ihn natürlich frei erfunden. Aber publiziert habe ich das ebenso natürlich nicht – nur hier zur Illustration.

3 Das folgende Kapitel ist im Wesentlichen unverändert aus meinem Fragebogenbuch (Porst 2008, 4. Auflage 2014: 126ff) übernommen.

4 Leider ist es uns nicht möglich, diesen Pretest näher zu benennen, weil die Durchführung von Pretests bei ZUMA, bei ZUMA-GESIS und bis vor kurzem auch bei GESIS noch Auftragsarbeit war, deren Ergebnisse dem Auftraggeber vertraulich und zu dessen ausschließlicher Nutzung übereignet worden sind. Ich muss also bitten, dass Sie dieses und die folgenden Beispiele ungesehen als reale Beispiele akzeptieren, auch wenn wir hier keine übliche Zitierung vornehmen können.

5 Wenn Sie beim Lesen dieser Aussage jetzt den Kopf schütteln und der Ansicht sind, niemand würde ernsthaft ein solches Item in einer Befragung vorlegen, muss ich Sie leider enttäuschen: Auch dieses Item wurde nicht von mir zur besseren Illustration erfunden, sondern war in einem von mir beratenen Fragebogenentwurf zur Befragung einer Allgemeinpopulation enthalten; durch die Ergebnisse unseres Pretests konnten wir seine Verwendung in der Hauptbefragung allerdings verhindern.

6 Der Nachweis zu dieser Behauptung lässt sich mit kognitiven Pretests eigentlich einfach leisten; kognitive Pretests helfen aber auch, Lösungen für dieses Problem zu finden, indem sie

Verständnisprobleme überhaupt erst aufzeigen, in Zusammenarbeit mit den Testpersonen zu alternativen Formulierungen oder – wenn es gar nicht anders geht – zu der Erkenntnis führen können, dass man auf eine Frage oder ein Item besser gleich ganz verzichten sollte, bevor man mit unverständlichen Fragen „Rauschen" produziert.

Literatur

Diekmann, A. (1995): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt

Friedrichs, J. (1973): Methoden empirischer Sozialforschung. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag

Gill, S.N. (1947): How do you stand on sin? Tide, (March 14), 17

NZZ-Folio. Die Zeitschrift der Neuen Züricher Zeitung. Heft 1/2006 „Statistik. Zählen und gezählt werden": 14-17

Porst, R. (2014): Fragebogen – Ein Arbeitsbuch. Studienskripten zur Soziologie. Wiesbaden: Verlag für Sozialwissenschaften, 4. Auflage

Porst, R. (2015a): Fragebogenkonstruktion (I) – Grundlagen, Arten von Fragen, Arten von Skalen. S. 239-250 in Gesellschaft. Wirtschaft. Politik 64, Heft 2

Porst, R. (2015b): Fragebogenkonstruktion (II) – „Question Wording" und was sonst noch wichtig ist. S. 381-392 in Gesellschaft. Wirtschaft. Politik 64, Heft 3

Schuman, H. & S. Presser (1981): Questions and Answers in Attitude Surveys. New York: Academic Press

Der Mensch: Randnotiz der Wirtschaft

Sebastian Thieme

Menschengerechtes Wirtschaften?
Subsistenzethische Perspektiven auf die katholische Sozialethik, feministische Ökonomik und Gesellschaftspolitik

2017. 297 S. Kt.
34,90 € (D), 35,90 € (A)
ISBN 978-3-8474-2077-4
eISBN 978-3-8474-1055-3

Wirtschaften, um sich selbst zu erhalten?

Was eigentlich selbstverständlich ist, bleibt in der ökonomischen Theorie und in den sozialpolitischen Debatten oft nur eine Randnotiz. Dem gegenüber diskutiert der Autor auf Basis einer modernen Subsistenzethik die Elemente der Selbsterhaltung in der katholischen Sozialethik und in feministischen Ansätzen der Ökonomik sowie sozialpolitische Fragen (Grundeinkommen, Mindest-Einkommensgarantie, Workfare, Einkommensobergrenzen).

www.shop.budrich-academic.de