

Grundgesamtheit, Stichprobe, Repräsentativität

Rolf Porst¹

Wenn wir im Fernsehen, in der Zeitung, im Internet oder wo auch immer Ergebnisse einer Umfrage zur Kenntnis nehmen, reagieren wir mit einer gewissen Erleichterung, wenn wir erfahren, dass es sich bei der berichteten Studie um eine „repräsentative“ Befragung handelt. Warum aber beruhigt uns das? Und was ist das überhaupt: „Repräsentativität“?

Um uns mit dem Begriff der „Repräsentativität“ angemessen auseinandersetzen zu können, müssen wir Begriffe wie „Population“ und „Stichprobe“ kennenlernen. Wir sollten wissen, welche Arten von Stichproben es gibt und welche Arten von Stichproben welche Arten von Aussagen zulassen. Und wir sollten uns selbstverständlich auch mit der Frage beschäftigen, ob Ihre SchülerInnen „repräsentative“ Befragungen durchführen können – und ob sie das überhaupt sollten.

Wir machen also einen Ausflug in die Welt der Verfahren zur Auswahl von Befragungspersonen. Dazu beginnen wir mit einigen Begriffsbestimmungen.

Grundgesamtheit

Am Anfang unserer Überlegungen zur Auswahl von Befragungspersonen müssen wir zunächst einmal abklären, über wen, also über welche Gruppe(n) von Personen wir mit den Ergebnissen unserer Befragung später Aussagen treffen wollen. Die Menge der Elemente, über die wir Aussagen machen wollen, bezeichnen wir als *Grundgesamtheit*.

Wenn Sie eine Befragung zu den Berufswünschen von zukünftigen Abiturienten in



Rolf Porst, Markt- und Sozialforscher
Römerberg

der Stadt Mannheim planen, dann könnte Ihre Grundgesamtheit definiert werden als „alle SchülerInnen der Oberstufe an Mannheimer Gymnasien und vergleichbaren, die allgemeine Hochschulreife vermittelnden Schulen im Schuljahr 2015/2016“. Wenn Sie wissen möchten, was die SchülerInnen an Ihrer Schule von dem neuen Musiksaal halten, dann ist Ihre Grundgesamtheit „alle SchülerInnen an unserer Schule im Schuljahr 2015/2016“.

Weitere Beispiele für Grundgesamtheiten anderer Befragungen wären etwa alle wahlberechtigten BürgerInnen in Deutschland, alle Mütter mit zwei Kindern in Wattenscheid, alle Studierende an der Universität zu Köln oder alle Personen, die an einem bestimmten Tag zwischen 9 Uhr und 18 Uhr die Prinzhorn-Ausstellung in Heidelberg besucht haben.

Wichtig ist: Grundgesamtheiten werden definiert, können also nicht falsch oder richtig sein. Aber sinnvoll sollten sie schon sein. Wenn Sie sich für die Betreuung von Vorschulkindern in der Stadt Duisburg interessieren, macht es wenig Sinn, Ihre Grundgesamtheit als „alle Männer über 60 Jahre in Duisburg“ zu definieren, auch wenn das formal „richtig“ und „zulässig“ wäre – sinnvoll wäre das nicht.

Elemente

Die Grundgesamtheit besteht aus *Elementen*. Elemente sind Untersuchungseinheiten, in Befragungen zumeist Personen, die für die Befragung ausgewählt und befragt werden sollen. Bleiben wir bei dem Beispiel „Berufswünsche“: Als Grundgesamtheit haben wir „alle SchülerInnen der Oberstufe an Mannheimer Gymnasien und vergleichbaren, die allgemeine Hochschulreife vermittelnden Schulen im Schuljahr 2015/2016“ definiert, die Elemente der Grundgesamtheit wären dann alle SchülerInnen der Oberstufe an Mannheimer Gymnasien und vergleichbaren, die allgemeine Hochschulreife vermittelnden Schulen im genannten Schuljahr.

Vollerhebung

Von einer *Vollerhebung* sprechen wir dann, wenn wir alle Elemente der Grundgesamtheit befragen wollen. Eine Vollerhebung ist dann sinnvoll, wenn der Umfang der Grundgesamtheit nicht allzu groß ist oder wenn alle Elemente auch einer großen Grundgesamtheit leicht zu erreichen sind. Wenn wir eine Befragung von MitarbeiterInnen bei der Firma Schulz & Söhne durchführen wollen und wir alle 850 MitarbeiterInnen über das Intranet erreichen, sollten wir uns für eine Vollerhebung entscheiden.

In den Sozialwissenschaften bilden Vollerhebungen eine wenngleich nicht so ganz seltene Ausnahme, was damit zusammenhängt, dass hier häufig Aussagen über „die Gesellschaft“ getroffen werden sollen, und „die Gesellschaft“ nun einmal ein recht großes Gebilde verkörpert. Eine Vollerhebung wäre hier also nicht nur unbezahlbar, sondern auch aus anderen Gründen kaum zu realisieren.

In der Marktforschung kommen Vollerhebungen häufiger vor, z.B. wenn Investitionsentscheidungen analysiert werden sollen. Dabei gilt es herauszufinden, welche Personen in einem Unternehmen an einer Entscheidung beteiligt sind und diese Personen in ihrer Gesamtheit zu befragen.

Der große Vorteil einer Vollerhebung besteht aus statistischer Sicht darin, dass durch die Befragung die gesamte Verteilung aller interessierender Merkmale in der Grundgesamtheit bekannt ist und nicht mit statistischen Verfahren geschätzt werden muss.

Teilerhebung/Stichprobe

Weil Vollerhebungen besonders bei großen Grundgesamtheiten mit nicht einfach zu erreichenden Elementen zu teuer und zu aufwändig sind, werden in den meisten Befragungen nur Teile der Grundgesamtheit befragt. Werden die Elemente für die Teilerhebung durch festgelegte Regeln bestimmt, spricht man normalerweise von einer *Stichprobe*. Das Prinzip der Stichprobe basiert – vereinfacht gesagt – auf der Möglichkeit, aus den Ergebnissen der Befragung der Stichprobe auf die Ergebnisse der die Stichprobe „repräsentierende“ Grundgesamtheit zu schließen. Wir werden gleich zeigen, dass diese Möglichkeit sehr stark von der *Art* der Stichprobe abhängt, die wir ziehen.

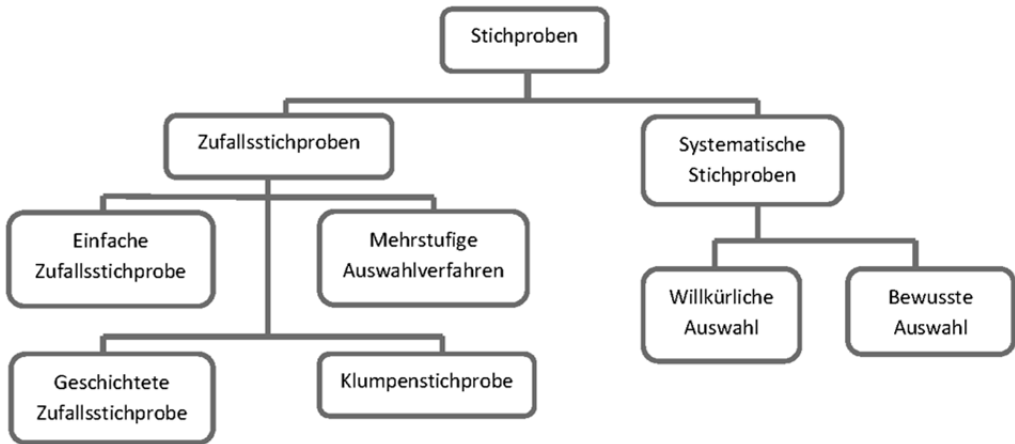
Stichproben können danach klassifiziert werden, ob die Auswahl der Elemente der Grundgesamtheit von einem Zufallsprozess gesteuert wird oder nicht. Wenn ja, sprechen wir von einer *Zufallsstichprobe*, wenn nein, sprechen wir von *systematischen Stichprobenverfahren*.

Allerdings: Der Zufall in der Stichprobentheorie hat eine gänzlich andere Bedeutung als der Zufall in der Alltagssprache. Während wir im Alltag mit Zufall meist ein unvorhersehbares Ereignis meinen, das scheinbar keinen nachvollziehbaren Regeln unterliegt, ist im mathematisch-statistischen Sinne mit Zufall ein Ereignis beschrieben, das Gesetzmäßigkeiten folgt, die angegeben werden können, nämlich dem *Zufallsprozess*. Dieser lässt sich mittels Wahrscheinlichkeiten mathematisch exakt beschreiben. So sind *Zufallsstichproben* (sogenannte *Random Samples*) Stichproben, deren Auswahlregeln es ermöglichen, vor Durchführung einer Auswahl für jedes Element der Grundgesamtheit die Wahrscheinlichkeit zu berechnen, mit der es Teil der Stichprobe werden könnte. Die Auswahlwahrscheinlichkeit jedes Elements der Grundgesamtheit sollte also unbedingt *bekannt* und *größer als Null* sein. Der große Vorteil von Zufallsstichproben ist die Tatsache, dass man mit den auf diese Weise gewonnen Daten mit Hilfe inferenzstatistischer Verfahren in der Lage ist zu bestimmen, inwieweit die Stichprobenergebnisse die Ergebnisse der Grundgesamtheit abbilden, wenn man die Grundgesamtheit befragt hätte.

Der Repräsentativitätsschluss, also der Schluss von den Stichprobenergebnissen auf die Ergebnisse der Grundgesamtheit, ist nur möglich, wenn die Befragungspersonen über eine Zufallsstichprobe ermittelt worden sind.

Auf die Begründung dieser Behauptung und auf die Darstellung der mathematischen Grundlagen wollen wir hier verzichten. Wir verschaffen uns stattdessen einen Überblick über die unterschiedlichen Stichprobenarten. Dabei hilft uns die folgende Abbildung:

Abbildung 1: Arten von Stichproben.



Schauen wir uns die unterschiedlichen Stichprobenarten näher an. Wir fokussieren uns dabei auf Stichprobenarten, die für schulische Befragungsprojekte „realisierbar“ erscheinen. Beginnen wir mit den Zufallsstichproben.

Zufallsstichproben

Wir unterscheiden Zufallsstichproben in *einfache* Zufallsstichproben, *geschichtete* Zufallsstichproben, *Klumpenstichproben* und *mehrstufige Auswahlverfahren*. Allen Zufallsstichproben, wie kompliziert sie auch sein mögen, liegt die einfache Zufallsstichprobe zugrunde.

Einfache Zufallsstichprobe

Bei der einfachen Zufallsstichprobe hat jedes Element der Grundgesamtheit die gleiche Chance, in die Stichprobe zu gelangen. Diese Chance oder Wahrscheinlichkeit ist vorab berechenbar und muss größer sein als Null. Einfache Zufallsstichproben zeichnen sich darüber hinaus durch einen einzigen Auswahlvorgang aus.

Das Paradebeispiel ist die Ziehung der Lottozahlen. Wir wissen von vornherein, wie groß die Population ist (49 Kugeln), wir wissen, wie viele Elemente wir ziehen wollen (6 Kugeln), wir kennen jedes Element, und wir kennen die Auswahlwahrscheinlichkeiten (für die erste Ziehung $1/49$, für die zweite Ziehung $1/48$... und für die sechste Ziehung $1/44$). Wenn wir die Auswahlwahrscheinlichkeiten betrachten, ist die Ziehung der Lottozahlen für jeden einzelnen Ziehungsvorgang tatsächlich eine einfache Zufallsstichprobe. Über den gesamten Ziehungsprozess ist sie das aber nicht, weil mit jedem Ziehungsvorgang die Auswahlwahrscheinlichkeit pro Kugel größer wird. Wir könnten aus der Ziehung der Lottozahlen aber durchaus eine einfache Zufallsstichprobe machen, nämlich dann, wenn wir alle sechs Gewinnkugeln gleichzeitig aus der Trommel fallen ließen.

Nun wollen wir aber in unseren Forschungsprojekten in der Schule nicht Lotto spielen, sondern Befragungen durchführen. Wie könnte das für das Beispiel der einfachen Zufallsstichprobe aussehen?

Kommen wir zurück zu dem Beispiel der Befragung zu den Berufswünschen von Abiturienten in der Stadt Mannheim. Als Grundgesamtheit hatten wir definiert: „alle SchülerInnen der Oberstufe an Mannheimer Gymnasien und vergleichbaren, die allgemeine Hochschulreife vermittelnden Schulen im Schuljahr 2015/2016“. Gehen wir der Übung halber davon aus, dass die Menge aller so definierten SchülerInnen 2.000 beträgt.² Nehmen wir weiterhin an, dass uns eine vollständige Liste aller Elemente vorliegt und dass wir eine Stichprobe von 500 SchülerInnen befragen wollen. Um eine einfache Zufallsstichprobe zu generieren, hätten wir zwei Möglichkeiten:

- a) Wir ordnen jedem Element der Grundgesamtheit eine eindeutige Nummer zwischen 0001 und 2000 zu. Mithilfe eines Zufallszahlengenerators lassen wir uns dann 500 Zufallszahlen aus diesem Zahlenfeld ermitteln. Diejenigen SchülerInnen, deren Nummer mit den Zufallszahlen übereinstimmen, gelangen in die Stichprobe.
- b) Wir schreiben die Namen der 2.000 SchülerInnen auf kleine Zettel, geben die Zettel in eine Lostrommel und ziehen per Zufall 500 Zettel aus der Trommel.

Eine dritte Möglichkeit wäre eine sog. systematische Zufallsstichprobe:

- c) Wir nehmen die Liste der Elemente, sortieren sie alphabetisch, ordnen ihnen wieder eindeutige Ziffern zwischen 0001 und 2000 zu und wählen aus den 2.000 Elementen auf der Liste per Zufall ein Element aus, das uns als Startelement dient. Ausgehend von diesem Startelement arbeiten wir die Liste dann so ab, dass jedes k -te Element gezogen wird. Die Größe des Stichprobenintervalls k berechnet sich aus der Größe der Population und der Größe der Stichprobe. In unserem Beispiel würden wir 2.000 durch 500 teilen und erhielten für k den Wert 4. Das heißt wir beginnen mit einem zufällig ausgewählten Namen auf unserer Liste der Elemente und zählen von dort aus immer 4 Namen weiter. Voraussetzung für das Funktionieren dieses Verfahrens ist allerdings, dass die Liste keiner Ordnung unterliegt, die unsere Auswahlmethodik beeinträchtigt – von daher die alphabetische Sortierung der Elemente über alle Schulen hinweg.

Diese Vorgehensweise erinnert daran, wie wir früher Stichproben aus Telefonbüchern gewonnen haben, um lokale Telefonbefragungen durchzuführen.

Abgesehen davon, dass die Konstruktion einfacher Zufallsstichproben relativ aufwändig ist, ist ihre Verwendung aus inhaltlichen Gründen oft gar nicht erforderlich oder auch nicht wünschenswert. Man greift dann auf komplexere Verfahren zurück.

Geschichtete Zufallsstichprobe

Angenommen wir untersuchen die Hypothese, dass AbsolventInnen mit Migrationshintergrund sich in ihren Berufswünschen von denen ohne Migrationshintergrund unterscheiden. Laut amtlicher Statistik der Stadt Mannheim sind 9,9% der SchülerInnen an Gymnasien „ausländische Herkunft“³; „runden“ wir die einfach mal auf 10% auf, um uns die Arbeit leichter zu machen.

Wenn wir jetzt eine einfache Zufallsstichprobe aus unserer Population ziehen würden, könnte es – wie der Zufall es so will – sein, dass wir nur wenige, im Extremfall gar keine SchülerInnen mit Migrationshintergrund in der Stichprobe hätten – für unsere Fragestellung natürlich fatal. Von daher arbeiten wir mit einer *geschichteten Zufallsstichprobe*.

Wir konstruieren eine geschichtete Zufallsstichprobe (für unser Beispiel der Größe $N = 500$), indem wir die Elemente der Grundgesamtheit bezüglich eines bestimmten Merkmals (hier: Migrationshintergrund/kein Migrationshintergrund) in Schichten teilen. Innerhalb dieser Schichten ziehen wir dann getrennt voneinander einfache Zufallsstichproben (hier: eine aus den SchülerInnen *mit*, eine aus den SchülerInnen *ohne* Migrationshintergrund). Wählen wir die Größe der Stichproben innerhalb der Schichten anhand des Anteils der Schichten an der Grundgesamtheit (der uns dann bekannt sein muss; hier: 10 % oder 50 Personen mit Migrationshintergrund aus insgesamt 500 für unsere gesamte Stichprobe), sprechen wir von einer *proportional geschichteten Stichprobe*.

Wären uns die 50 Befragungsteilnehmer mit Migrationshintergrund aber zu wenig, um unsere Fragestellung zu bearbeiten, könnten wir den Anteil dieser Gruppe auch erhöhen, indem wir z.B. 100 oder 150 SchülerInnen innerhalb der Schicht „mit Migrationshintergrund“ per Zufall ziehen und sie mit den verbleibenden 400 bzw. 350 SchülerInnen ohne Migrationshintergrund vergleichen. Wir sprechen dann von einem *disproportional geschichteten Stichprobendesign*, in dem die gezogenen Anteile der einzelnen Schichten nicht den realen Anteilen in der Grundgesamtheit entsprechen. Die interessierende Schicht, also SchülerInnen mit Migrationshintergrund, wird überrepräsentiert.

Damit können wir nun unsere Fragestellung bearbeiten; darüber hinaus sind wir in der Lage für jede der beiden „Schichten“ unabhängige Schätzungen auf deren (Teil-) Grundgesamtheiten vorzunehmen. Der Nachteil: Da beim disproportionalen Schichten das Kriterium derselben Chance der Auswahl für jedes Element der Grundgesamtheit verletzt ist, müssen bei der Datenauswertung die einzelnen Schichten unterschiedlich gewichtet werden, um unverzerrte Aussagen über die Grundgesamtheit treffen zu können. Wenn man das überhaupt will.

Klumpenstichprobe

Von einer Klumpenstichprobe sprechen wir dann, wenn sich die Auswahlregeln nicht auf die Elemente der Grundgesamtheit, sondern auf zusammengefasste Elemente beziehen und jeweils alle Elemente eines solchen Klumpens befragt werden. Die Klumpenstichprobe wenden wir z. B. dann an, wenn keine Liste der Elemente der Grundgesamtheit vorliegt, wohl aber eine Liste der zusammengefassten Elemente der Grundgesamtheit.

Nehmen wir den nicht unrealistischen Fall an, wir hätten keine Liste der einzelnen Mannheimer SchülerInnen, wie oben definiert. Aber: Wir haben die Liste der Gymnasien bzw. vergleichbaren Schulen in Mannheim. Dann würden wir bei einer Klumpenstichprobe zunächst durch eine einfache Zufallsauswahl einige Gymnasien auswählen und in den ausgewählten Gymnasien alle SchülerInnen befragen, die der Grundgesamtheit angehören. Eine Klumpenstichprobe würde bei unserem Beispiel besonders dann Sinn machen, wenn wir einen Zusammenhang der Studienwünsche der einzelnen SchülerInnen mit denen der MitschülerInnen vermuten würden.

Dieses Design sollte jedoch nicht angewendet werden, wenn die Klumpen in sich sehr homogen sind, sich untereinander aber stark unterscheiden. Hätten wir bei unserem Beispiel zufällig ein reines Wirtschaftsgymnasium gezogen und würden dort alle Elemente der Grundgesamtheit befragen, könnten wir annehmen, dass sich die Studienwünsche dort deutlich zugunsten wirtschaftswissenschaftlicher Studiengänge von

denjenigen der SchülerInnen anderer Gymnasien unterscheiden. Wir hätten also einen für die Grundgesamtheit untypischen Klumpen gezogen, der bei unserer Untersuchung stark ins Gewicht fallen würde, da wir ja alle Schüler dieses Klumpens befragen.

Die bisher geschilderten Auswahlverfahren sind bei vielen Fragestellungen undurchführbar, insbesondere bei Untersuchungen, die die allgemeine Bevölkerung als Grundgesamtheit besitzen. Hier müssen komplexere Verfahren gewählt und in mehreren Stufen einfache Zufallsauswahlen vorgenommen werden. Auch wenn Ihre SchülerInnen Studien dieser Art nicht durchführen werden, wollen wir doch kurz darstellen, wie mehrstufige Auswahlverfahren aussehen und sei es nur, um zu erfahren, wie die Stichproben für große, bundesweite Befragungen erstellt werden.⁴

Mehrstufige Auswahlverfahren

Bei mehrstufigen Auswahlverfahren werden zunächst sog. primary sampling units zufällig ausgewählt; das können z.B. Gemeinden sein. Innerhalb der ausgewählten Gemeinden werden dann die secondary sampling units per Zufall ermittelt, in der Regel Haushalte; und schließlich innerhalb der Haushalte die zu befragende Person.

Wir betrachten dies etwas näher – aber dennoch stark verkürzt – am Beispiel des sog. ADM-Designs.⁵ Beim ADM-Design (ADM 1999) handelt es sich um eine dreistufig geschichtete Klumpenauswahl, mit deren Hilfe bundesweit Personen für persönlich-mündliche Befragungen per Zufall ermittelt werden sollen.

Die Bundesrepublik Deutschland wird dabei auf der Basis von Daten der amtlichen Statistik in mehr als 50.000 *Flächen* eingeteilt. Im *ersten Schritt* der Stichprobenziehung zieht ADM nach dem Zufallsprinzip eine Stichprobe aus diesen Flächen; die gezogenen Flächen werden als *sampling points* bezeichnet. Für 1.000 zu realisierende Interviews im westlichen Teil Deutschlands werden – das ist Konvention – 210, im östlichen Teil 48 *sampling points* gezogen. Die *sampling points* umfassen dabei ca. 600 bis 700 Haushalte. Schichtungskriterien für die Auswahl der *sampling points* sind Bundesländer, Regierungsbezirke und Kreise in Kombination mit einer Gemeindetypisierung.

Im *zweiten Schritt* werden innerhalb der *sampling points* diejenigen Haushalte ermittelt, in denen Befragungen stattfinden sollen. Dazu kommt das sog. random route- bzw. das adress random-Verfahren zum Einsatz. Bei diesen Verfahren erhält der Interviewer pro *sampling point* vom Befragungsinstitut eine Startadresse, von der aus er nach exakt festgelegten Regeln den Bezirk begehen und die Zielhaushalte zu ermitteln hat. Dabei gibt es zwei Varianten: Entweder listet der Interviewer die nach diesen Begehungsregeln ermittelten Haushalte auf und schickt die Liste zurück an sein Institut, welches dann andere Interviewer dorthin schickt, um die Interviews durchzuführen (adress random). Oder der Interviewer führt die Interviews in den von ihm ermittelten Haushalten selbst durch (random route).

Im dritten Schritt schließlich werden innerhalb der ausgewählten Haushalte die Zielpersonen ermittelt; es liegt auf der Hand, dass man nicht einfach diejenige Person befragen kann, die „gerade da“ ist, weil das natürlich zu einer Überrepräsentation älterer Menschen, von Hausfrauen/-männern, Arbeitslosen oder sonstigen Personen führen würde, die eher zuhause anzutreffen sind. Also wird auch innerhalb der Haushalte zufällig ausgewählt. Dazu verwenden wir zumeist die sog. last-birthday-Methode, bei der der Interviewer diejenige Person im Haushalt zu befragen hat, die zuletzt Geburtstag hatte.

Sie merken: Während auf der ersten und zweiten Stufe des ADM-Designs zumindest theoretisch jedes Element, also jede Fläche und jeder Haushalt, die gleiche Chance hat, in die Stichprobe zu gelangen, haben im dritten Schritt die Mitglieder kleinerer Haushalte eine größere Auswahlwahrscheinlichkeit als die Mitglieder größerer Haushalte.

Man braucht keine Fachkenntnisse, um zu erkennen, dass das ADM-Design recht fehleranfällig und seine Umsetzung nur schwer zu kontrollieren ist. Wir wollen darauf aber nicht näher eingehen, weil – wie eingangs gesagt – Ihre SchülerInnen Studien auf dieser Basis nicht durchführen werden.

Wenden wir uns stattdessen von den zufallsgesteuerten Stichproben ab und den nicht-zufallsgesteuerten, den systematischen Stichproben zu.

Systematische Stichproben

Wir unterscheiden die systematischen Stichproben in *willkürliche und bewusste Stichproben*.

Willkürliche Auswahl

Von einer willkürlichen Auswahl sprechen wir dann, wenn es keinerlei Regeln gibt, die der Auswahl der Befragungspersonen zugrunde liegt; ob jemand befragt wird oder nicht, erfolgt unkontrolliert und steht ausschließlich im Belieben des Interviewers bzw. der Interviewerin.

Stellen Sie sich zur Illustration vor, dass Ihre SchülerInnen eine Befragung „auf der Straße“ durchführen und ohne Vorgabe irgendwelcher Regeln einfach die Personen zur Befragung anheuern, die gerade vorbeikommen und fröhlich dreinschauen. Oder die sie attraktiv finden oder die aussehen, also ob man sie leicht rekrutieren könnte.

Auch wenn das sicherlich die einfachste Art der Gewinnung von Befragungspersonen ist, kann ich Ihnen davon nur abraten; dass willkürliche Stichproben für wissenschaftliche Zwecke wertlos sind, versteht sich von selbst. Allemal besser sind die ebenfalls noch leicht zu realisierenden bewussten Auswahlen.

Bewusste Auswahl

Bewusste Auswahlen haben im Vergleich zu willkürlichen Auswahlen den Vorteil, dass sie einem Auswahlplan folgen, die Kriterien der Auswahl also nachvollziehbar und bekannt sind; allerdings können auch hier keine inferenzstatistischen Techniken zum Einsatz gebracht werden. Grob gesagt: Aus den Daten der Stichprobe kann nicht auf die Ergebnisse irgendwie definierter Grundgesamtheiten geschlossen werden.

Zu den bewussten Auswahlen gehören z.B. die Auswahl *typischer Fälle* und das *Konzentrationsprinzip*. Bei der Auswahl *typischer Fälle* wählen wir Personen aus, von denen wir vermuten oder erwarten, dass sie besonders charakteristisch für die Grundgesamtheit sind. Beim *Konzentrationsprinzip* suchen wir nach Fällen, bei denen ein interessierendes Merkmal besonders ausgeprägt ist und damit die gesamte Verteilung in der Grundgesamtheit bestimmt. In der Marktforschung wird dieses Verfahren meist in der Investitionsgüterindustrie angewendet. Wenn z.B. 10% der Kunden 85% des Um-

satzes eines Betriebes ausmachen, kann es – je nach Fragestellung – ratsam sein, bei einer Kundenbefragung nur innerhalb dieser 10% der Kunden zu befragen.

Eine weitere Form der bewussten Auswahl stellt das *Schneeballverfahren* dar, das dann eingesetzt werden kann, wenn in der Bevölkerung selten vorkommende Personen befragt werden sollen. Nehmen wir als Beispiel Personen, die zweimal geschieden und zum dritten Mal verheiratet sind. Weil wir davon ausgehen können, dass diese Kombination eher selten auftritt, rekrutieren wir per Schneeball: Wir suchen und finden eine Person, auf die unsere Definition zutrifft, befragen sie und bitten sie, uns Personen zu nennen, auf die das Gleiche zutrifft. Ausgehend von der Vermutung homogener Soziotope können wir unterstellen, dass Personen mit bestimmten Merkmalen und Eigenschaften andere Personen mit den gleichen Merkmalen und Eigenschaften kennen.

Kommen wir zur letzten der bewussten Auswahlen, dem sogenannten *Quotenverfahren*. Selbst wenn man gelegentlich anderes hört: Auch das Quotenverfahren lässt den wissenschaftlich gesicherten Schluss von der Stichprobe auf die Grundgesamtheit nicht zu (DFG 1999: 47). Es besteht aber ungeachtet dessen kein Zweifel daran, dass Quotenstichproben sich in der Praxis bewähren (ebenda).

Von daher und weil sie relativ leicht zu organisieren sind, könnte die Befragung von Personen, die „per Quote“ ermittelt worden sind, sicherlich für viele oder sogar für die meisten der schulischen Forschungsprojekte als das Mittel der Wahl gelten. Was bedeutet „per Quote“?

Beim Quotenverfahren suchen die InterviewerInnen die Befragungspersonen danach aus, ob sie bestimmte, vorab definierte Merkmalsausprägungen vorzuweisen haben, z.B. einer bestimmten Altersgruppe angehören, eine bestimmte Schulbildung vorzuweisen haben, sich einer bestimmten Partei nahe fühlen.

Wird nur ein einziges Merkmal vorgegeben, sprechen wir von einer „einfachen Quote“; der Interviewer/die Interviewerin erhält die Aufgabe: „Befragen Sie einen Mann“. Oder: „Befragen Sie eine Person mit Abitur“. Von einer „kombinierten Quote“ ist dann die Rede, wenn mehrere Merkmalsausprägungen kombiniert sein sollen: „Befragen Sie eine Frau zwischen 45 und 64 Jahren mit Hauptschulabschluss“. Oder: „Befragen Sie einen Mann zwischen 18 und 45 Jahren, der verheiratet ist und der SPD nahesteht“. Sie können die Merkmale und Merkmalsausprägungen beliebig erweitern, aber mit jeder Erweiterung erschweren Sie die Suche nach „passenden“ Personen.

Wichtig ist: Wenn die InterviewerInnen losziehen, um ihre „Quote“ zu finden, müssen sie sich wirklich nur an den Vorgaben orientieren, die ihnen gemacht wurden. Ob der zu suchende Mann über die Quotenvorgabe hinaus Kinder hat, politisch aktiv ist, gerne Billard spielt, erwerbstätig ist oder gerne ins Kino geht – das alles spielt für die Auswahl der Zielperson keine Rolle.

Ein wichtiger Vorteil der Befragung nach Quote ist, dass sie einfach umzusetzen ist. Entweder Ihre SchülerInnen kennen Menschen, die in die Quote „passen“, dann können sie gezielt befragen und ggfs. von dort aus weitere Personen rekrutieren. Oder: Sie können auch in Fußgängerzonen beliebige Personen befragen, die „so aussehen“, als ob sie die Quote erfüllen und sie – wenn sich das an ihren Daten zeigt – in den realisierten Quotenplan aufnehmen oder sie – wenn es nicht zutrifft – halt weglassen. Dann war das Interview leider umsonst. Sie können also den Quotenplan im Nachhinein füllen.

Die Quotenmerkmale können Sie beliebig definieren; besser ist es aber, wenn Sie mit Merkmalen arbeiten, deren Verteilung in der Population bekannt ist. Wenn Sie z.B. eine Befragung in Mannheim machen wollen, können Sie vorab aus den Daten der kommunalen Statistik herauslesen, wie die Variablen Geschlecht, Alter, Schulbildung oder Einkommen in der Bevölkerung dort verteilt sind und dann Ihre Quoten an dieser

Verteilung ausrichten. Anders ausgedrückt: Ihr Quotenplan könnte eine maßstabsgerechte Miniatur der realen Verhältnisse in Mannheim darstellen – damit erreichen Sie natürlich immer noch keine „repräsentativen“ Daten. Aber die müssen Sie ja meist gar nicht erreichen. Und: Was genau ist „Repräsentativität“ eigentlich?

Repräsentativität – oder die Definition eines Mythos

Dass es einen Zusammenhang gibt zwischen dem Begriff der „Repräsentativität“ und der Art und Weise der Stichprobenziehung dürfte Ihnen beim Lesen des bisherigen Textes deutlich geworden sein. Mehrmals war dort die Rede vom (wissenschaftlich gesicherten) Schluss von der Stichprobe auf die Grundgesamtheit gewesen. Nichts anderes meint der Begriff der „Repräsentativität“.

Dass Umfragedaten „repräsentativ“ sein sollen, ist – aus der Sicht der Rezipienten – vielleicht die wichtigste Anforderung an ihre Qualität. Zugleich ist Repräsentativität auch einer der ganz wenigen „Fachtermini“ (Vorsicht! Gleich mehr dazu.), die Eingang in die öffentliche oder zumindest in die veröffentlichte Diskussion gefunden haben; Repräsentativität wird zumeist gleichgesetzt mit Qualität und Wahrheitsgehalt von Umfragedaten. Mit dem Hinweis auf die Repräsentativität der Ergebnisse wird ihre Bedeutung hervorgehoben.

Diesem Bemühen tut nun die Tatsache keineswegs einen Abbruch, dass es repräsentative Stichproben (und damit „Repräsentativbefragungen“) eigentlich gar nicht, oder sagen wir besser, nur in bestimmten Zusammenhängen gibt; Diekmann (1995: 368) hält denn auch die „Redeweise von der ‘repräsentativen Stichprobe‘“ für „nicht mehr als eine Metapher, eine bildhafte Vergleichung“. Man könne, so Diekmann weiter, „mit einiger Berechtigung auch von einem Mythos der repräsentativen Stichprobe sprechen“ (ebenda: 369).

Qualitätsnachweis und Gütekriterium oder Metapher und Mythos? Was hat es mit der Repräsentativität von Umfragen wirklich auf sich?

Umfrageergebnisse wären dann repräsentativ, wenn die Umfrage, der sie entstammen, repräsentativ wäre. Eine Umfrage wäre repräsentativ, wenn die Stichprobe, die aus der Population gezogen wird, repräsentativ wäre für diese Population.

Wenn wir überhaupt von „Repräsentativität“ und „repräsentativen Stichproben“ sprechen wollen und können, müssen wir dafür Sorge tragen, dass jedes Element der – zumindest behelfsmäßig definierten – Grundgesamtheit (z.B. „alle Personen mit deutscher Staatsbürgerschaft, die in Deutschland in Privathaushalten leben und spätestens am 1. Juni 2015 das 18. Lebensjahr vollendet haben“) die gleiche Chance hat, in unsere Stichprobe zu gelangen. Damit reduziert sich der hohe Anspruch nach „Repräsentativität“ auf die Frage einer exakten Stichprobenziehung – *und es sind alleine Zufallsstichproben, die es uns ermöglichen, aus ihren Ergebnissen in Bezug auf die Verteilung aller Merkmale innerhalb bestimmter statistischer Fehlergrenzen auf die entsprechenden Verteilungen innerhalb der Population zu schließen*. In diesem Verständnis sind „repräsentative Stichproben“ und „Zufallsstichproben“ synonyme Begriffe.

Also: Der Begriff der „repräsentativen Stichprobe“, der im übrigen – wir hatten ihn vorhin als „Fachterminus“ bezeichnet – gar kein Fachbegriff ist (Diekmann 1995: 368), hat nur dann eine gewisse Berechtigung, wenn wir ihn als Synonym für Zufallsstichproben verstehen. Ansonsten ist er wertlos. Wir sollten uns angewöhnen, ihn zu vermeiden und an seiner statt von Zufallsstichproben zu sprechen. „Repräsentativität“ ist

kein Merkmal einer Befragung an sich, sondern etwas, das über das Verfahren einer sauber gezogenen Zufallsstichprobe angezielt werden soll.

Zum Schluss

Nach der Beschäftigung mit den unterschiedlichen Stichprobenverfahren und nach der Lektüre aller in der Reihe „Schüler forschen“ behandelten Themen sind wir nun so weit, dass wir von der Vorbereitung der von uns geplanten Befragung zu ihrer Durchführung kommen oder – wie wir sagen – „in’s Feld“ gehen können.

Abhängig von dem von uns gewählten Stichprobenverfahren müssen wir Personen auswählen, die befragt werden sollen („Zielpersonen“), je nach dem von uns gewählten Befragungsmodus müssen wir die Fragen an den Mann/an die Frau bringen und die Antworten von dort wieder an uns zurück. Wie finden wir jetzt die „richtigen“ Zielpersonen? Welche Probleme können bei der Ansprache potenzieller Zielpersonen auftreten? Dürfen wir als SchülerInnen einfach Leute „auf der Straße“ ansprechen? Ein Telefoninterview durchführen? Oder gar an der Haustür klingeln, um zur Teilnahme an unserer Befragung einzuladen?

Mit Fragen dieser Art beschäftigen wir uns im nächsten Heft.

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 Die Zahl ist fiktiv, aber nicht unrealistisch, wenn man sich die Angaben über die Abgänge an Mannheimer Gymnasien 2013/2014 ansieht: <https://www.mannheim.de/sites/default/files/page/21353/schulstatistik2014.pdf>, Seite 55
- 3 Ebenda, Seite 80
- 4 Bundesweite Stichproben können auch als „Einwohnermeldeamtsstichproben“ realisiert werden oder, als Stichproben für telefonische Befragungen, etwa mit dem Gabler-Häder-Design. Auf die Darstellung dieser Arten von Stichprobenziehung können wir aber gerne verzichten, weil sie für Ihre schulischen Projekte nicht realisierbar sein werden. Einen kurzen Überblick über unterschiedliche Stichprobenverfahren finden Sie bei Interesse z.B. bei Häder & Häder (2014).
- 5 Der Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM) ist eine Landesorganisation, zu der sich die bedeutendsten kommerziellen Umfrageinstitute in Deutschland zusammengeschlossen haben. Siehe <https://www.adm-ev.de/startseite/>

Literatur

- ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., Hrsg. (1999): Stichproben-Verfahren in der Umfrageforschung. Eine Darstellung für die Praxis. Opladen: Leske + Budrich
- Baur, N. & J. Blasius, Hrsg. (2014): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS
- DFG, Deutsche Forschungsgemeinschaft, Hrsg. (1999): Qualitätskriterien der Umfrageforschung. Denkschrift, Hrsg. von Max Kaase. Berlin: Akademie-Verlag

- Diekmann, A. (1995): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt
- Häder, M. & S. Häder (2014): Stichprobenziehung in der quantitativen Sozialforschung. S. 283-297 in: Baur, N. & J. Blasius (2014)