Methoden der empirischen Sozialforschung

Soziologie bzw. allgemeiner Sozialwissenschaften sind in ihrem Kern empirische Wissenschaften. D. h. die Entwicklung von Hypothesen und Theorien und ihre Konfrontation mit den realen Fakten, deren Resultat die vorläufige Bestätigung oder vorläufige Zurückweisung der theoretischen Überlegungen zur Folge hat, ist ein ganz zentraler Bestandteil der erfahrungswissenschaftlichen Vorgehensweise. Das methodologische Fundament dieses Postulates wird in der Wissenschaftstheorie seit A. Comte vertreten und ist zugleich in wichtigen Teilen sehr deutlich differenziert und modifiziert worden. Im letzten Jahrhundert haben vor allem zwei Denktraditionen, der Logische Empirismus (Carnap, 1974) und der Kritische Rationalismus, die eine Vielzahl von Gemeinsamkeiten aufweisen, die Weiterentwicklung bestimmt. Der Logische Empirismus hielt an der Idee der Verifikation fest. Die Prüfung wissenschaftlicher Theorien soll zu Ergebnissen führen, die auch zukünftig Geltung haben. Aber auch wenn die die empirischen Sachverhalte beschreibenden Aussagen bestätigten bzw. in wissenschaftlichen Untersuchungen replizierbar sind, rechtfertigt keine (immer endliche) Menge an konsistenten Beobachtungssätzen eine raum-zeitliche Verallgemeinerung. Wenn aber der dazu notwendige (gehaltserweiternde) Schluss von gegenwärtigen Resultaten der Theorieprüfungen auf zukünftige nicht möglich ist, kann die Verifikation (also der Nachweis der Wahrheit) von Theorien keine brauchbare methodologische Regel darstellen. Auf dieses Problem hat vor allem K. R. Popper (1976) hingewiesen. Das Fehlen einer brauchbaren Induktionslogik schwächt die Überzeugungskraft des Logischen Empirismus und seine forschungspraktische Relevanz nachhaltig. Diese kritischen Hinweise belasten nicht nur den Logischen Empirismus, sondern natürlich auch und noch mehr die (alte) positivistische Wissenschaftslehre. Dem Glauben an die Möglichkeit einer Verifikation wird der Boden entzogen. Schon dadurch bricht die moderne erfahrungswissenschaftliche Methodologie eindeutig mit dem Positivismus.

Die zweite zentrale Überzeugung des Positivismus ist in der scheinbar untrüglichen, zweifelsfreien Beweiskraft der empirischen Fakten zu sehen. Der Kritische Rationalismus hat auch diese Annahme nachhaltig in Frage gestellt (Chalmers, 1986). Die direkt oder indirekt beobachteten Fakten sind selbst als theoriegeleitete bzw. interpretierte empirische Zustände und/oder Prozesse zu betrachten. Diese Sicht der Dinge hat gravierende Konsequenzen für das Konzept der Falsifikation. Wenn Fakten, genauer: die sie beschreibenden Beobachtungssätze (Basissätze), keine fraglose Gültigkeit besitzen, und sich Widersprüche zwischen der Theorie und den Basissätzen ergeben, kann eine Theorie nicht eindeutig falsifiziert werden. Da die empirische Basis selbst theoriegeleitet ist bzw. mess- und wahrnehmungstheoretische Implikationen aufweist, die ihrerseits nicht verifiziert werden können, entfällt (auch) eine sichere Falsifikationsgrundlage. Die ‚richtige', ‚wahre' Abbildung von empirischen Fakten als Testkriterium für Hypothesen und Theorien erscheint vor diesem Hintergrund als unmöglich, sie bleibt aber zugleich eine bedeutsame regulative Idee des wissenschaftlichen Bemühens.

In dieser Situation ist die (vorläufige) Akzeptanz von Theorien immer auch ein Entscheidungsproblem. Irrtümer sind möglich, sichere Falsifikation und Verifikation sind nicht

möglich. Stattdessen geht es um die (immer vorläufige) Bestätigung von theoretischen Aussagen (Lakatos, 1974). Diese ist letztlich in die kritische Diskussion der scientific community eingebettet.
Ein zentrales Kriterium für die vorläufige Akzeptanz von Hypothesen und Theorien liegt in der methodischen Qualität der entsprechenden Untersuchungen bzw. in der methodischen Stärke des Hypothesentests. Damit ist die zentrale Aufgabe der empirischen Sozialforschung im Wissenschaftsprozess umrissen. Sie entwickelt sozialwissenschaftliche Methoden, leitet zu ihrer forschungspraktischen Umsetzung an und strebt dabei nach möglichst zuverlässigen und validen empirischen Aussagen.
Die Erfassung empirischer Fakten und ihre Konfrontation mit den theoretischen Aussagen wird gängiger Weise als ein Forschungsprozess betrachtet, der eine Reihe von Arbeitsschritten und eine spezifische Abfolge dieser Schritte impliziert (Schnell, Hill & Esser, 2013; Diekmann, 2010). Die einzelnen Arbeitsschritte sind aber keineswegs unabhängig voneinander, sondern haben starken Verweisungscharakter und sind interdependent. Insbesondere muss der Forschungsprozess auch als ein Entscheidungsprozess betrachtet werden. Die zentralen Arbeitsschritte sind: Theoriebildung, Forschungsdesign, Operationalisierung und Messung, Auswahlverfahren, Datenaufbereitung und Datenanalyse.
Theoriebildung: Von oft unterschätzter Bedeutung ist eine möglichst klare Vorstellung von den Fakten oder Prozessen, die untersucht werden sollen. Jeder denkbare sozialwissenschaftliche Forschungsgegenstand erlaubt verschiedene Perspektiven, weist vielfältige Dimensionen auf und ist differenten Konstruktionen und Interpretationen zugänglich. Seine Komplexität ist begrifflich aufzuarbeiten und erzwingt klare Abgrenzungen und damit Entscheidungen hinsichtlich der Konturierung der Forschungsfrage. Diese Festlegungen wirken sich massiv auf alle folgenden Forschungsschritte aus. Es wird - vor dem Hintergrund der vorhandenen Forschungsressourcen - nicht nur festgelegt was exakt untersucht wird, sondern implizit auch was nicht erforscht wird. Je nach Eigenheiten des interessierenden Phänomens wählt man einen allgemeinen theoretischen Zugang - etwa lern-, handlungs- oder entscheidungstheoretischer Art. Innerhalb dieses Konzeptes sind dann Hypothesen zu entwickeln, die einen erklärenden Charakter und zugleich einen Realitätsbezug aufweisen, der einen empirischen Zugang erlaubt. Gelegentlich erscheint keiner der bekannten theoretischen Ansätze auf den zu erklärenden Forschungsgegenstand anwendbar, und somit sind auch keine Hypothesen ableitbar. In solchen Fällen ist eine explorative Phase unumgänglich. Je nach Forschungsstand ist die Konkretisierung des Problems innerhalb von Voruntersuchungen sinnvoll. Das Fehlen eines angemessen erscheinenden Theorierahmens und konkreter
Hypothesen macht eine theoretische Exploration notwendig. Für diesen Arbeitsschritt der Hypothesengenerierung gibt es keine standardisierten Verfahren oder Limitationen. In diesem Entdeckungskontext wird auf unterschiedlichste Quellen zurückgegriffen: Alltagserfahrungen, Analogienschlüsse, Spekulation, Phantasie u. a. Entscheidend für die nächsten Forschungsschritte sind klare begriffliche Abgrenzungen und explizite Hypothesen, deren empirischer Test avisiert wird.
Forschungsdesign: Das Design einer Untersuchung kennzeichnet die Anordnung der Messungen hinsichtlich der Zeitpunkte und der Anzahl über die in der Untersuchung teilnehmenden Personen (bzw. Objekte). Die klassische wissenschaftliche Untersuchungsform ist ohne Zweifel das Experiment (Saris, 1991). Es sieht in seiner Minimalform eine Experimental- und eine Kontrollgruppe vor. Die Zuordnung der Untersuchungseinheiten zu den Gruppen erfolgt über einen Zufallsprozess (Randomisierung). Zu einem ersten Zeitpunkt erfolgt in beiden Gruppen die Messung der Ausprägung der abhängigen Variable (Wirkung). Danach werden die

Untersuchungseinheiten der Experimentalgruppe dem Stimulus (Ursache) ausgesetzt. Man spricht auch vom Treatment, Treatmentfaktor oder der unabhängigen Variable. Hat der Stimulus die hypothetisch vermutete Auswirkung, dann müssen bei der zweiten Messung der abhängigen Variablen signifikante Unterschiede zwischen den beiden Gruppen nachweisbar sein, während bei der ersten Messung keine signifikanten Differenzen zwischen den Gruppen feststellbar sind. Da idealerweise beide Gruppen bis auf die Setzung des Stimulus gleich behandelt werden und die Gruppenzuordnung nach einer Randomisierung erfolgt, sind die Veränderungen (mit bestimmbarer Wahrscheinlichkeit) auf das Treatment zurückzuführen. Die skizzierte Versuchsanordnung entspricht einem (einfachen) Experiment. Dieses basale Design kann in vielfältiger Weise verfeinert, differenziert und erweitert werden. Sogenannte mehrfaktorielle Versuchspläne integrieren mehrere unabhängige Variablen, wodurch der Einfluss der verschiedenen Faktoren abgeschätzt werden kann. Experimentelle Designs sind im Vergleich zu anderen Forschungsstrategien eher robust, aber keineswegs frei von Störeffekten, die ihre Aussagekraft einschränken. Die interne und externe Validität von Experimenten ist kritisch zu diskutieren bzw. zu untersuchen.

Elaborierte experimentelle Designs stellen die beste Untersuchungsstrategie dar, können aber in der Forschungspraxis häufig nicht angewandt werden. Insbesondere sind Experimente in der Soziologie (leider) sehr selten anzutreffen. Die Ursache dafür liegt in der aus ethischen bzw. praktischen Gründen nicht durchführbaren experimentellen ‚Behandlung' der Untersuchungsgruppe. Stimuli - wie Arbeitslosigkeit, Ehescheidung, ethnische Zugehörigkeit - können nicht zu experimentellen Zwecken gezielt eingesetzt werden. Zur Untersuchung entsprechender Effekte wird das sogenannte Survey-Design (Umfragedesign, Ex-post-facto-Design) ein- gesetzt. Dabei wird der Stimulus nicht kontrolliert vom Forscher gesetzt, sondern entsprechende Faktoren und Prozesse treten (unkontrolliert) in der Alltagswelt auf. Ex post lassen sich dann u. U. verschiedene Gruppen bilden (Verheiratete vs. Geschiedene, Arbeitslose vs. Erwerbstätige), die dann verglichen werden um die Effekte der genannten Ereignisse bzw. Zustände zu erfassen und zu quantifizieren. Alle interessierenden Informationen hinsichtlich der abhängigen und unabhängigen Variablen werden in einem Datenerhebungsprozess erfasst (Querschnittsuntersuchung). Bei dieser Untersuchungsstrategie treten jedoch nicht triviale methodische Probleme auf. Insbesondere können die u. U. beobachteten Gruppenunterschiede nicht mit Sicherheit auf die unabhängige Variable zurückgeführt werden, da diese mit anderen Variablen (z. B. Bildung, Einkommen) konfundiert ist. Zusätzlich sind Survey-Designs in ihrer einfachen Form nur sehr eingeschränkt in der Lage, die kausale Abfolge von Effekten zu analysieren. Die methodische Güte lässt sich dadurch verbessern, dass man mehrere Messzeitpunkte in das Design integriert. Entsprechende Längsschnittuntersuchungen können die Aussagekraft von Survey-Designs deutlich steigern. Unter ‚Längsschnittuntersuchungen' werden in der Praxis der empirischen Sozialforschung sehr verschiedene Strategien subsumiert. Panelstudien, die die Ausprägungen der gleichen Variablen an den gleichen Personen zu verschiedenen Zeitpunkten erfassen, bilden eine Gruppe von Längsschnittstudien. Mit ihrer Hilfe lassen sich intraindividuelle und gesellschaftliche Veränderungen erfassen.

Bei sogenannten Trendstudien wechseln die untersuchten Personen in jeder Erhebungswelle. Mit Daten aus Trendstudien lassen sich keine intraindividuellen Änderungen beschreiben, aber (aggregierte) Prozesse gesellschaftlichen Wandels (z. B. Beschäftigte im tertiären Sektor, Ethnozentrismus, Einstellungswandel) lassen sich rekonstruieren. Die zeitlichen Abstände zwischen den Erhebungszeitpunkten sollten sich bei Panel- und auch Trendstudien an den inhaltlichen Fragen der Untersuchungen orientieren. Kurzfristig mögliche Einstellungswechsel machen geringe

Zeitabstände sinnvoll, eher langfristige Prozesse, etwa des intergenera- tionalen sozialen Auf- oder Abstiegs, erlauben größere Abstände zwischen den Erhebungen bzw. Messzeitpunkten. In der Forschungspraxis trifft man häufig auf Erhebungsabstände von einem oder zwei Jahren, insbesondere wenn sie weniger theorieprüfenden Charakter haben und primär gesellschaftlichen Wandel dokumentieren wollen.

Operationalisierung und Messung: Unter Operationalisierung versteht man in der empirischen Sozialforschung eine Anweisung für die Zuweisung von Objekteigenschaften zu theoretischen Begriffen. Alle empirischen Wissenschaften müssen die in Theorien und Hypothesen verwendeten Begriffe in empirisch direkt oder indirekt beobachtbare Sachverhalte ‚übersetzen'. Interessierende Eigenschaften und ihre jeweiligen Ausprägungen (Merkmale und Merkmalsausprägungen) sind erst dann empirischen Analysen zugänglich, wenn die Forschung empirische Indikatoren benennen kann, die einem theoretischen Begriff zugeordnet werden (Korrespondenzproblem). In der Geschichte der empirischen Sozialforschung wurden verschiedene Korrespondenzregeln diskutiert. Gegenwärtig wird zumeist der ‚kausalanalytische Ansatz' präferiert. Er interpretiert die theoretischen Begriffe als latente Variablen. Die latenten Variablen bedingen bzw. verursachen die Ausprägungen der empirischen Indikatoren, beispielsweise reagiert eine Person mit postmaterialistischer Einstellung auf bestimmte Fragen in einem Fragebogen mit einem spezifischen Antwortmuster, zeigt also beobachtbare Reaktionen. Die Reaktionen (Antwortverhalten) können variieren (z. B. Zustimmung, starke Ablehnung). Als Messung bezeichnet man nun die Zuordnung von Zahlen zu den Antwortvariationen (Merkmalsausprägungen). Nicht jede Zahlenzuordnung ergibt aber eine sinnvolle Messung, sondern eine adäquate Messung beruht auf einer strukturtreuen Abbildung (Stevens, 1946; Steyer, & Eid, 1993). Typischerweise werden die Merkmalsausprägungen durch Zahlen repräsentiert. Diese numerischen Repräsentationen sind dann die Grundlage der angestrebten statistischen Datenanalyse. In der empirischen Sozialforschung sind Messungen nicht fehlerfrei. Die Güte von Messungen wird im Allgemeinen über zwei Kriterien bestimmt: Reliabilität und Validität. Die erste betrifft die Zuverlässigkeit - also die Konstanz der Messergebnisse über die Zeit - die Zweite zielt auf die korrekte inhaltliche Erfassung des zu Messenden durch die Messinstrumente. Geht man von den spezifischen Parametern der klassischen Testtheorie aus, lassen sich Maße für Reliabilität und Gültigkeit schätzen, die Hinweise auf die fehlende Güte der Messung geben. Die klassische Testtheorie selbst geht davon aus, dass sich die empirischen Messwerte aus dem wahren Wert der zu erfassenden Eigenschaft und einem Fehlerterm zusammensetzen. Zudem werden vier Annahmen über die Verteilung der Messfehler postuliert. Diese Annahmen sind aber nicht immer realitätsnah. Diese Schwächen versucht die probabilistische Testtheorie zu umgehen, die jedoch in der Praxis der empirischen Sozialforschung noch wenig Resonanz gefunden hat.

Messungen erfolgen auf unterschiedlichen Mess- bzw. Skalenniveaus. In den Sozialwissenschaften differenziert man im Allgemeinen zwischen Nominal-, Ordinal-, Intervall- und Ratio-Skalen. Je nach Messniveau können die Skalenwerte (Messwerte) unterschiedlichen mathematisch-statistischen Transformationen unterzogen werden. Der Einsatz statistischer Analysemethoden korrespondiert mit den Skalenniveaus der zu analysierenden Daten. Die Berechnung eines arithmetischen Mittels ist beispielsweise für Nominaldaten inadäquat, für Intervalldaten aber angemessen und hilfreich.

Innerhalb der empirischen Sozialforschung verwendet man sehr häufig Messinstrumente, die sich aus mehreren Indikatoren zur Messung der gleichen Eigenschaft zusammensetzen. Indizes bestehen im Allgemeinen aus einer Kombination von Indikatoren, die verschiedene Dimensionen eines theoretischen Begriffs zu einem Messinstrument zusammenfassen. Solche mehrdimensionalen Messinstrumente sind in den Sozialwissenschaften häufig anzutreffen, zum

Beispiel bei der Erfassung der sozialen Schichtzugehörigkeit oder der Intelligenz. Eindimensionale Skalierungsverfahren haben ebenso eine lange sozialwissenschaftliche Tradition. Fast immer bestehen sie aus einer Vielzahl von Items, die alle auf die gleiche Messdimension zielen. Zu den bekanntesten Skalierungsverfahren gehören Likert-Skalen, Guttman-Skalen und Magnitude-Skalen. Solche Skalierungstechniken werden ganz überwiegend zur Erfassung von Einstellungen verwendet. Die Güte von Skalenkonstruktionen wird üblicherweise über spezielle Verfahren der Reliabili- täts- und Validitätsschätzung ermittelt. Es ist offensichtlich, dass die empirische Sozialforschung sich in besonderem Maße um die Entwicklungen zuverlässiger und gültiger Messverfahren bzw. Indizes und Skalen bemüht. Wenn empirische Fakten die Basis der Theorieprüfung darstellen, ist eine hohe Datenqualität zwingend.

Auswahlverfahren: Messungen bzw. Datenerhebungen erfolgen immer an Objekten. Innerhalb des Forschungsprozesses ist zu entscheiden, an welchen Objekten die Messungen erfolgen sollen. Natürlich sind es in den Sozialwissenschaften zumeist Menschen. Hypothesen haben einen Objektbereich, d. h. benennen Objekte (z. B. Europäer, deutsche Kinder, Menschen) für die die Hypothesen gelten sollen. Grundsätzlich können die Messungen entweder an allen interessierenden Objekten oder an einer Teilmenge erfolgen - man spricht von Voll- oder Teilerhebung. Vollerhebungen sind nur sinnvoll, wenn der Objektbereich eher klein ist, wenn also die Hypothese auf eine überschau- und erreichbare Population (z. B. die aktuellen Landtagsabgeordneten im Saarland) zielt. Aus guten Gründen gilt das wissenschaftliche Interesse aber zumeist größeren Populationen bis hin zu ‚allen Menschen'. Vollerhebungen sind dann praktisch nicht realisierbar und u. U. auch fehlerbehafteter als Teilerhebungen. Zur Rekrutierung von Personen aus der Gesamtheit der zu interessierenden Population (Grundgesamtheit) werden Auswahlverfahren eingesetzt. Sehr grob lassen sich Wahrscheinlichkeitsauswahlen und willkürliche (inkl. bewusste) Auswahlen unterscheiden. Verfahren der letzteren Gruppe (z. B. Quotenauswahl, Internet oder E-Mail basierte willkürliche Auswahlen) sind durchaus verbreitet aber genügen den üblichen wissenschaftlichen Standards nicht (Schnell, Hill & Esser, 2013).

Zur Gruppe der Wahrscheinlichkeitsauswahlen (random sampling) gehört eine Vielzahl konkreter Techniken, die allesamt Zufallsstichproben generieren. Zufällig bedeutet, dass jedes Element bzw. jede Person die gleiche bzw. eine angebbare

Chance hat, in die Stichprobe zu gelangen. Auf dieser Grundlage sind dann auch Inferenzschlüsse von der Stichprobenpopulation auf die Grundgesamtheit gut begründet möglich - was aber eben bei willkürlichen Auswahlen nicht der Fall ist. Je nach Untersuchungsziel und Problemlage werden einfache, geschichtete oder Klumpenstichproben zugrunde gelegt. Telefonische Umfragen werden häufig mit Dual-Frame-Stichproben verbunden. Bei großen Bevölkerungsumfragen dominieren mehrstufige Auswahlverfahren (Schnell, Hill & Esser, 2013).

Der Umfang von Zufallsstichproben ist abhängig von der tolerierbaren Fehlermarge hinsichtlich der auf Grundlage der Stichprobenstatistiken geschätzten Parameter der Grundgesamtheit. Für sozialwissenschaftliche Untersuchungen, die Aussagen über die nationale Wohnbevölkerung anstreben, sind Stichprobengrößen zwischen 1.000 und 3.000 Personen nicht unüblich.

Die Qualität einer Stichprobe ist in der Forschungspraxis auch von der Ausschöpfungsrate (als komplimentäre Größe der nonresponse rate) abhängig (Schnell, 1997). Die angestrebte Stichprobengröße kann aufgrund von Ausfällen (z. B. Nichterreichbarkeit oder Verweigerung von Personen) zumeist nicht realisiert werden. Besonders problematisch sind Ausfälle, wenn sie nicht neutral sind, d. h. mit in der Untersuchung erhobenen Informationen zusammenhängen. Dadurch entstehende Verzerrungen sind auch nicht über Gewichtungsverfahren (Redressment) korrigierbar. Da eine Zunahme des Nonresponse (insbesondere durch Verweigerung) in den letzten Jahrzehnten

zu verzeichnen ist, muss die empirische Sozialforschung verstärkt Gegenstrategien entwickeln und für die Teilnahme Anreize entwickeln.
Datenerhebung: Die wichtigsten Datenerhebungsformen sind Befragung, Beobachtung und Inhaltsanalyse. In der sozialwissenschaftlichen Praxis dominiert das Interview als Erhebungsverfahren ganz eindeutig. Beobachtung und Inhaltsanalyse sind derzeit nur gelegentlich anzutreffen, was aber nicht dauerhaft so bleiben muss, da sich durch den verstärkten Einsatz moderner Technik (z. B. Video, GPS-Ortung, computerbasierte Dokumentenanalyse) durchaus Änderungen ergeben könnten. Alle Erhebungsmethoden haben ihre spezifischen Stärken und Schwächen. Als eine besonders wichtige Eigenschaft ist dabei die Reaktivität der Erhebungsverfahren zu sehen. Darunter versteht man das Ausmaß, die Empfindlichkeit, einer Erhebungsmethode für unerwünschte Effekte, die durch die Erhebungsmethodik selbst bedingt sind und damit auf die intendierten Messungen verzerrend Einfluss nehmen. Befragte reagieren u.U. nicht nur auf das Messinstrument (Fragen im Fragebogen), sondern auch auf Eigenschaften der Interviewer. Innerhalb einer Beobachtung, über die die Beobachteten nicht informiert sind, sind auch keine Reaktivitätseffekte erwartbar.
Das Interview wurde in den 1950er und 1960er Jahren als ‚Königsweg' der empirischen Sozialforschung eingeschätzt, da mit überschaubaren Mitteln umfangreiche
Informationen gesammelt werden können, zumindest im Vergleich zu anderen Erhebungsformen. Mit Hilfe von Interviews konnten zu sehr unterschiedlichen Forschungsfragen Daten erhoben werden, die zudem auch relativ schnell verfügbar waren. Befragungen (Groves et al., 2009; Schnell, 2012) können in sehr verschiedenen Formen stattfinden: als persönliches Interview, als postalische oder telefonische Befragung, mit Unterstützung durch einen Computer (CATI bzw. CAPI), als Gruppenbefragung bzw. Gruppeninterview, als Expertengespräch u. a. Zudem lassen sich verschiedene Arten von Fragen differenzieren: Einstellungsfragen, Faktenfragen, Verhaltensfragen etc. Fragen bzw. Befragungen lassen sich auch nach dem Grad der Standardisierung differenzieren. Nichtstandardisierte Leitfadeninterviews geben keine konkreten Frageformulierungen vor, sondern nur Gesprächsthemen, teilstandardisierte Interviews arbeiten mit offenen Fragen, die die Frageformulierung wörtlich festschreiben, aber keine festen Antwortkategorien verwenden und standardisierte Interviews geben Fragestellung und Antwortkategorien vor.
Welche Befragungsform zum Einsatz kommt ist von der theoretischen Fragestellung abhängig und auch vom Stand des Vorwissens. Allgemein gilt: Je geringer das Vorwissen, desto eher können weniger standardisierte Techniken zum Einsatz kommen. Stark standardisierte Befragungsformen schaffen besser vergleichbare Messsituationen und dienen so der Güte der Messung. In der Praxis gibt es sehr häufig eine Entwicklung von wenig standardisierten Formen, die primär der Erkundung dienen, zu standardisierten Messungen, die oft zur Entwicklung von Skalen führen. Diese Instrumentenentwicklung macht die Durchführung von Pretests (‚Vorprüfungen') zwingend notwendig. Neben der Instrumentenentwicklung sind im Vorfeld der Befragung weitere vorbereitende Schritte unumgänglich, insbesondere die Interviewerschulung. Alle diese Bemühungen zielen darauf, die Güte der Messungen zu verbessern und Antwortverzerrungen (Response-Errors) - wie etwa soziale Erwünschtheit, Akquieszenz, Sponsorship- oder Interviewereffekte - zu vermeiden (Sudman & Bradburn, 1974; Schuhman & Presser, 1981).
Aufgrund der Dominanz der Befragungsmethodik ist mittlerweile eine explizite Forschung zu entsprechenden Methodenfragen (Artefakteforschung) und auch eine Theorie des Befragtenverhaltens (Esser, 1986) entstanden. Für die Beobachtung und die Inhaltsanalyse kann das kaum gesagt werden. Beobachtungen (Friedrichs & Lüdtke, 1977; Babbie, 2001) sind mittlerweile in der Soziologie nur noch selten zu finden. Inhaltsanalysen haben ebenfalls in der

soziologischen Forschung weniger Relevanz. Dies sagt aber wenig über die Leistungsfähigkeit dieser Strategien aus. Insbesondere ist für Beobachtungen festzuhalten, dass sie tatsächliches Handeln bzw. Verhalten erfassen und messen. Bei Befragungen werden zumeist nur Handlungsabsichten und -gründe erfasst. Bekannterweise sind diese aber nicht mit faktischen Handlungen oder Verhalten gleichzusetzen. Manche Formen der
Beobachtung können zudem als nicht-reaktive Verfahren gelten. Dieser Vorzug trifft auch auf die Inhaltsanalyse bzw. Dokumentenanalyse zu (Krippendorff, 2004), deren Einsatzgebiet von der Analyse politischer Propaganda, sowie von Zeitschriften und Parlamentsdebattenprotokollen bis zu Comics und Todesanzeigen reicht. Dabei wird auch hier theoriegeleitet gearbeitet. Die Dokumente sind Objekte, die Merkmale und Merkmalsausprägungen aufweisen, die wiederum als Operationalisierung für theoretische Begriffe - wie z. B. gesellschaftlicher Wandel' stehen. Wie angesprochen könnte das Potential dieser Erhebungsmethode durch den verstärkten Einsatz neuer Kommunikationstechniken in den nächsten Jahren stärker genutzt werden. Damit käme man auch der gewünschten Kombination verschiedener Erhebungsmethoden (Triangulation) näher, in der die jeweiligen spezifischen Vorteile genutzt werden.
Datenaufbereitung und Datenanalyse: ist die Datenerhebung erfolgt, steht die Aufbereitung der Daten an (Schnell, Hill & Esser, 2013). Bei Erhebungen, die Informationen in nicht standardisierter Form (etwa auf Tonträgern aufgezeichnete qualitative Interviews) erfassen, ist eine Transkription notwendig. Die entstehenden Textkorpora sind dann der Ausgangspunkt der Auswertung bzw. Interpretation des Materials.
Bei quantitativen Daten gestaltet sich die Datenerfassung je nach der Erhebungsform unterschiedlich. Gelegentlich ist noch eine manuelle Datenerfassung notwendig, die die Informationen des Papierfragebogens mit Hilfe eines EDV-Programmes in elektronischer Form ab-speichert. Erfolgt die Datenerhebung bereits mit Hilfe einer Software (z. B. bei CATI oder bei Internetbefragungen), ist eine explizite Datenerfassung kaum mehr notwendig. In jedem Fall werden die erhobenen Daten in nummerischer Form in einer Datenmatrix zusammengetragen. In dieser Matrix sind die Zeilen über die Untersuchungseinheiten (z. B. Personen) defi-niert, und die Spalten repräsentieren die Variablen und ihre Ausprägungen. Eine Datenmatrix wird nach Regeln erstellt, die im Codeplan festgeschrieben sind. Jeder Eintrag in der Matrix kann mit dessen Hilfe entschlüsselt werden. So stehen die Zahlen 3 und 4 in Spalte 7 und 8 in der 9. Zeile für die Information, dass der Befrage Nr. 9 ein Alter von 34 Jahren hat.
Die Informationen der Datenmatrix sind Grundlage der quantitativen Datenanalyse. In den ersten Auswertungsschritten werden zumeist deskriptive univariate Verteilungen erstellt. Es folgen bivariate und schließlich multivariate statistische Analysen, die sich allesamt natürlich an den theoretischen Ausgangsfragen bzw. den Hypothesen orientieren, d. h. Hypothesen testen. Die statistischen Analysen geben dann Auskunft über die empirische Stärke der theoretisch postulierten Zusammenhänge. Sie sind ein zentrales Kriterium für die vorläufige Akzeptanz der Hypothesen und Theorien. Zugleich sind die empirischen Ergebnisse von
Forschungsprojekten fast immer Ausgangspunkt für Theorierevisionen, Vertiefungen und Spezifikationen, die ihrerseits neue, weiterführende Fragestellungen und Forschungen anregen.
Die Skizze der zentralen Arbeitsschritte verdeutlicht auch nochmal die Interdependenz der vielen notwendigen Entscheidungen: Die Wahl der Forschungsfrage hat direkte Konsequenzen für die angemessene Erhebungsart. Arbeitet man experimentell, dann sind andere Operationalisierungen geboten als innerhalb eines Surveydesigns. Will man Reaktivitätseffekte vermeiden, scheinen Interviews nicht unbedingt angeraten. Ist man an verallgemeinerbaren Daten über eine umfangreiche Grundgesamtheit interessiert, dann sind Beobachtungen zur Datenerhebungen kaum

ratsam. Empirische Forschungsprojekte bedürfen somit einer Gesamtplanung, bei der die methodenbezogenen Entscheidungen zu einer sinnvollen Projektkonzeption zusammengeführt werden. Richtungsweisend ist dabei aber immer die theoretische Fragestellung.
Empirische Sozialforschung verursacht natürlich auch (gelegentlich hohe) finanzielle Kosten. Entsprechende Forschungsunterstützung wird kaum von universitärer Seite geleistet. Sogenannte Drittmittel werden von außen eingebracht bzw. eingeworben. Geldgeber sind Stiftungen, Ministerien und sonstige wissenschaftsfördernde Institutionen. Der weit größte und wichtigste Forschungsförderer ist die Deutsche Forschungsgemeinschaft (DFG). Sie unterstützt empirische Forschungsprojekte, nachdem diese vorgängig ein Begutachtungsverfahren (peer review) durchlaufen haben. Die im Rahmen der Forschung erhobenen Daten werden im Normalfall in wissenschaftlichen Archiven gespeichert und auch anderen Forschern zur Verfügung gestellt.
Die empirische Sozialforschung (zumindest in Deutschland) ist derzeit auch von einer Kontroverse zwischen quantitativer und qualitativer Forschung (Lamneck, 1989) geprägt. Dabei geht es im Kern nicht um den angemessenen Einsatz entsprechender Techniken. Es dürfte wohl Konsens sein, dass nicht-standardisierte, explorative Erhebungstechniken vor allem dann einzusetzen sind, wenn die inhaltliche Fragestellung noch unklar ist bzw. das Vorwissen (etwa aus vorgängigen Studien) noch sehr gering ist. Im Kern ist der sogenannte ‚Methodenstreit' eben kein Methodenstreit, sondern eine methodologische Kontroverse. Wenn man die beiden Standpunkte schlagwortartig zusammenfassen will, ist es nicht unplausibel die quantitativ-orientierten Wissenschaftler, die natürlich auch gar nicht selten qualitative Methoden einsetzen, als Vertreter einer erklärenden Sozialwissenschaft zu charakterisieren; sie suchen nach den kausalen Regeln, denen soziales Handeln folgt, und sie bemühen sich um die Prognosefähigkeit der Sozialwissenschaft. Qualitativ-orientierte Wissenschaftler halten dieses Streben für müßig, oder auch den Sozialwissenschaften für prinzipiell nicht angemessen oder falsch. Diese

Haltung resultiert im Wesentlichen aus der Überzeugung, dass soziales Handeln eines verstehenden Zugangs bedarf, dem eine einheitswissenschaftliche Methodologie nicht gerecht werden kann. Damit schimmert aber wieder eine alte, aus guten Gründen längst überwunden geglaubte Debatte, die zwischen ‚Erklären' und Verstehen' erneut hervor.

Revision #1
Created 12 October 2025 01:03:00 by investigatione
Updated 12 October 2025 01:03:16 by investigatione