

ZUMA-Arbeitsbericht 98/02

**Im Vorfeld der Befragung:
Planung, Fragebogenentwicklung, Pretesting**

Rolf Porst
Mai 1998

ZUMA
Quadrat B2,1
Postfach 12 21 55
D-68072 Mannheim

Telefon: (0621) 12 46 - 228
Telefax: (0621) 12 46 - 100
E-mail: Porst@zuma-mannheim.de

Inhalt

	Seite
1. Vorbemerkung: Warum dieser Bericht geschrieben wurde und was er soll	3
2. Projektplanung und Projektdurchführung	4
2.1 Theoretische Vorarbeiten.....	5
2.2 Planung und Durchführung der Datenerhebung	7
2.3 Datenerfassung und Datenaufbereitung.....	10
2.4 Auswertung	11
3. Befragungen als Verfahren zur Erhebung sozialwissenschaftlicher Daten	12
4. Fragebogenerstellung	17
4.1 Zielsetzungen eines Fragebogens	21
4.2 Kognitionspsychologische Grundlagen der Befragung.....	22
4.3 Arten von Fragen.....	23
4.4 Fragen- und Antwortformulierungen.....	27
4.5 Skalen.....	28
4.6 Aufbau des Fragebogens und Fragensukzession.....	30
4.7 Befragungshilfen, Filter, Layout	33
5. Pretesting	34
5.1 Stellenwert und Ziel eines Pretests.....	34
5.2 Pretestverfahren.....	36
6. Literatur.....	41

1. Warum dieser Bericht geschrieben wurde und was er soll

Am 30. und 31. November 1995 führte der Autor dieses Berichts bei der „Kooperationsstelle Arbeitswelt und Wissenschaft“ der „Siegerland Consult Gesellschaft zur Förderung eines sozial- und umweltverträglichen Strukturwandels“ in Siegen einen Workshop zum Thema „Entwicklung und Design von Fragebogen“ durch. Aufgabe des Workshops sollte es sein, Basisinformationen zur Konstruktion von Fragebogen sowie neuere Erkenntnisse und Erfahrungen in diesem Bereich vorzustellen. Der Schwerpunkt des Workshops sollte auf der Vorbereitung eines Fragebogens für eine sozialwissenschaftliche Umfrage liegen, sich also mit Fragebogenerstellung und Pretesting beschäftigen. Es sollte aber auch Wert gelegt werden auf die Einordnung des Prozesses der Fragebogenentwicklung in die gesamte Projektplanung und Projektdurchführung.

Diesem Anspruch zufolge gliederte sich der Workshop in einzelne Lektionen zu den Themen

- Projektplanung und Projektdurchführung
- Datenerhebungsverfahren
- Fragebogenerstellung
- Pretesting

und fand selbstverständlich seine Abrundung in einem Abschlußgespräch der TeilnehmerInnen.

Bei diesem Abschlußgespräch wurde der für Veranstaltungen dieser Art nicht ungewöhnliche Wunsch geäußert, eine Kopie der Folien erhalten zu dürfen, die die einzelnen Lektionen illustriert hätten. Die Frage, ob man nicht vielleicht gleich auch noch die entsprechenden Vortragsmanuskripte erhalten könne, führte dann zu der Idee, man sollte am besten doch alles einmal zusammenstellen und zusammenkopieren, so daß man es später auch verwenden könnte für eigene Lehrveranstaltungen insbesondere in der Methodenausbildung des Grundstudiums, oder auch bei anderen Gelegenheiten, bei denen man sich mit dem Thema zu beschäftigen habe.

In dem vorliegenden Bericht wird diese Idee aufgenommen und materialisiert. Er integriert die überarbeiteten (und um neuere Entwicklungen seit 1995 erweiterten) Fassungen der damals verwandten Vortragsskizzen mit den Folien, die zum Einsatz gekommen sind. Die Folien können als WinWORD-Dateien auf Diskette beim Verfasser angefordert und für Zwecke der Methodenausbildung oder für welche Zwecke auch immer verwendet werden (sofern die Nutzerin/der Nutzer ihre Herkunft preisgibt).

2. **Projektplanung und Projektdurchführung**

Wäre es nur darum gegangen, einen Überblick zu geben über die Planung und Durchführung eines sozialwissenschaftlichen Forschungsprojekts, hätte man sich die Mühe mit diesem Bericht sparen können; es gibt eine ganze Reihe guter und sehr guter Darstellungen der Phasen, die ein empirisches Forschungsprojekt in den Sozialwissenschaften (idealerweise) zu durchlaufen hätte (von Alemann 1984; Porst 1985; Friedrichs 1990; Schnell, Hill & Esser 1992; Diekmann 1995). Da es uns hier aber um eine - sogar auf Folie gebannte - überblickartige Kurzdarstellung geht, haben wir uns zunächst auch hier mit Projektplanung und Projektdurchführung zu beschäftigen.

Ein sozialwissenschaftliches empirisches Forschungsprojekt (im folgenden nur noch mit „Projekt“ bezeichnet), durchläuft idealerweise vier Phasen:

Folie Pro_1: Phasen der Projektdurchführung

1. Theoretische Vorarbeiten

- **Definition des Erkenntnisinteresses**
- **Bearbeitung der Literatur**
- **Formulierung des theoretischen Bezugsrahmens**
- **Generierung der Hypothesen**

2. Planung und Durchführung der Datenerhebung (Umfrage)

- **Operationalisierung der Forschungsfragen**
- **Planung der Durchführung der Umfrage**
- **Population, Stichprobe für den Pretest**
- **Pretesting**
- **Stichprobe für die Hauptstudie**
- **Durchführung der Hauptstudie**

3. Datenerfassung und Datenaufbereitung

- **Datenerfassung und Verkodung offener Fragen**
- **Datenbereinigung**

4. Auswertung

- **Analysen**
- **Publikation der Ergebnisse**

Merke:

Nicht alle empirischen Studien laufen nach dem gleichen Schema ab!

Die vier Phasen

- theoretische Vorarbeiten
- Planung und Durchführung der Datenerhebung¹
- Datenerfassung und Datenaufbereitung
- Auswertung

lassen sich jeweils (auch wieder idealtypisch) in einzelne Arbeitsschritte untergliedern. Am Anfang steht das Erkenntnisinteresse, am Ende die Publikation der Ergebnisse in geeigneter Form.

Soweit der erste Überblick über die Planung und Durchführung eines empirischen sozialwissenschaftlichen Forschungsprojekts. Wir werden uns mit den einzelnen Phasen gleich näher beschäftigen.

Bei dem Überblick nicht berücksichtigt werden Aspekte, die für ein Forschungsprojekt zwar von existentieller Bedeutung sind, aber eigentlich nicht unbedingt aufgrund methodischer Überlegungen. Gemeint sind Fragen der Forschungsförderung, der Forschungsfinanzierung, der Beantragung von Mitteln, usw.

Abschließend soll darauf hingewiesen werden, daß der vorgestellte Überblick über die Planung und Durchführung eines empirischen sozialwissenschaftlichen Forschungsprojekts als idealtypische Darstellung durchaus verallgemeinert werden kann, daß aber in der Praxis selbstverständlich nicht alle Studien nach dem gleichen Schema ablaufen.

2.1 Theoretische Vorarbeiten

Neugierde ist aller Forschung Anfang. Ganz zu Beginn jedes Forschungsvorhabens steht deshalb (s. Folie Pro_2 auf Seite 6) die Definition des Erkenntnisinteresses.

Es ist zu fragen, was denn eigentlich das Ziel der geplanten Forschungsarbeit sein soll, was den Forscher oder die Forscherin interessiert, was er oder sie „rauskriegen“ will. Dabei ist es vollkommen beliebig, wodurch das Erkenntnisinteresse des Forschers/der Forscherin ausgelöst werden mag. Auslöser können z.B. sein:

- wissenschaftliches Interesse an einer Thematik
- eigene frühere oder rezipierte Forschungsarbeiten
- Vorantreiben der eigenen Karriere
- die Wahrnehmung eines sozialpolitischen Problems in den Medien
- die Aufforderung des Professors, sich mal mit dem Thema xxx zu beschäftigen
- ein entsprechender Auftrag eines beliebigen Geldgebers.

So unterschiedlich diese Motive auch sein können, eines ist ihnen gemeinsam: Sie stellen legitime Auslöser für die Definition eines Erkenntnisinteresses dar.

¹ Wobei wir im folgenden Datenerhebung mit (persönlich-mündlicher) Befragung gleichsetzen wollen.

Folie Pro_2: Theoretische Vorarbeiten**Theoretische Vorarbeiten****Definition des Erkenntnisinteresses**

- Was ist das Ziel meiner geplanten Forschungsarbeit?
- Was interessiert mich?
- Was will ich „rauskriegen“?
- Auslöser für das Erkenntnisinteresse: beliebig.

Bearbeitung der Literatur

- Einstieg in die Problematik
- Überblick über den Stand der Forschung
- Angebote theoretischer Bezugsrahmen
- Hilfestellung bei der Formulierung der Hypothesen
- Anregungen

Formulierung des theoretischen Bezugsrahmens

- Auswahl und Darstellung des allgemeinen theoretischen Rahmens für die Erklärung und/oder Prognose
- Definition der zentralen (abhängigen und unabhängigen) Variablen

Generierung der Hypothesen

- Explikation der aus dem theoretischen Rahmen ableitbaren Zusammenhänge

Nachdem wir - auf welchem Wege auch immer - festgelegt haben, was unser Erkenntnisinteresse sei, haben wir damit zu beginnen, uns mit dem Thema näher zu beschäftigen. Da es nur wenig wirklich Neues in der Welt gibt, können wir grundsätzlich davon ausgehen, daß zu jedem Thema, das wir bearbeiten wollen, bereits gedrucktes - neuerdings digitalisiertes und via Internet zu empfangendes - Material vorliegt. Die Bearbeitung der vorhandenen Literatur (der konventionellen wie der digitalisierten) stellt grundsätzlich einen guten Einstieg in die von uns zu behandelnde Problematik dar. Vermittels der wissenschaftlichen Literatur zum Thema verschaffen wir uns einen Überblick über den Stand der Forschung, wir finden Angebote für die Auswahl eines theoretischen Bezugsrahmens und Hilfestellung bei der Formulierung unserer Hypothesen. Ganz allgemein erhalten wir aus der Literatur vielfältige Anregungen, die für unsere weitere Forschungsarbeit von großer Bedeutung sein können.

Der nächste Schritt (der zeitlich auch parallel zum Literaturstudium stehen kann)² besteht in der Formulierung eines theoretischen Bezugsrahmens für unsere Fragestellung. Wir haben - oft unter konkurrierenden Ansätzen - einen allgemeinen theoretischen Rahmen auszuwählen, der als Grundlage für Erklärung oder Prognose dienen kann (bei einer Erklärung liegt ein bestimmter sozialer Tatbestand vor, und man sucht nach den Anfangsbedingungen und Ursachen dieses Tatbestandes sowie nach einer Theorie bzw. nach Gesetzesaussagen zu seiner Erklärung; bei der Prognose sind Anfangsbedingungen und Theorie gegeben, und man sucht nach einem zukünftigen Tatbestand). Neben der Entscheidung zwischen Erklärung und Prognose steht die inhaltliche Definition der zentralen unabhängigen und abhängigen Variablen an.

Am Ende der theoretischen Vorarbeiten ist schließlich die Generierung der Hypothesen zu leisten, also die Explikation der aus dem theoretischen Rahmen und der Literatur ableitbaren Zusammenhänge. Da wir diese dann später auch in empirischen Studien ermitteln und überprüfen wollen, müssen wir uns nun mit der Planung und Durchführung der Datenerhebung beschäftigen, und wir tun dies am Beispiel der persönlich-mündlichen Befragung.

2.2 Planung und Durchführung der Datenerhebung

Die Entscheidung für das Beispiel der persönlich-mündlichen Befragung ist dabei im Prinzip willkürlich gefallen; viele der einzelnen Schritte, die wir im folgenden erwähnen werden (z. B. die Notwendigkeit, die Zielpopulation zu bestimmen), treffen auch auf andere Befragungsformen zu.

Folie Pro_3: Planung und Durchführung der Datenerhebung (Umfrage)

Planung und Durchführung der Datenerhebung (Umfrage)

- **Auswahl der Erhebungstechnik und Entwicklung des Erhebungsinstruments**
- **Planung der Durchführung der Datenerhebung (Umfrage)**
- **Bestimmung der Population**
- **Art und Realisierung der Stichprobe für den Pretest**
- **Pretesting**
- **Art und Realisierung der Stichprobe für die Hauptstudie**
- **Durchführung der Hauptstudie**

² Die meisten der hier beschriebenen Schritte können zeitlich hintereinander, aber ebenso gut auch zeitlich parallel ablaufen.

Zunächst einmal aber muß diese Entscheidung im Forschungsprozeß getroffen werden. Wir haben also unter einer größeren Zahl anerkannter Datenerhebungsverfahren dasjenige auszuwählen, das zum Beantworten unserer Forschungsfrage am geeignetsten erscheint: Entscheiden wir uns für ein Beobachtungsverfahren oder für eine Befragung? Führen wir die Befragung persönlich-mündlich durch oder am Telefon? Verwenden wir einen weitgehend offenen Leitfaden, oder basiert unsere Datenerhebung auf einem standardisierten Fragebogen?

Egal, wie wir uns auch entscheiden, der nächste Schritt besteht in der Entwicklung eines Erhebungsinstrumentes, sei es nun ein Beobachtungsbogen, ein Leitfaden oder ein standardisierter Fragebogen.

In gewisser Abhängigkeit von der Auswahl der Erhebungstechnik und der damit verbundenen Entwicklung eines Erhebungsinstrumentes stellt sich die Frage, wer denn die Datenerhebung letztendlich durchführen soll. Die Antwort auf diese Frage ist sicherlich nicht unabhängig vom gesamten Forschungsdesign: Will ich nur wenige Personen mit einem Leitfaden befragen, kann ich das durchaus selbst tun. Plane ich eine telefonische Befragung von 500 Personen, kann ich das - vielleicht mit Unterstützung von Studenten - ebenfalls noch selbst in Angriff nehmen. Will ich aber z.B. 5.000 erwachsene Deutsche in Privathaushalten in einer bundesweiten Studie persönlich-mündlich befragen, werde ich die Grenzen meiner Kapazitäten sicherlich nicht zu diskutieren haben.

Grundsätzlich stellt sich also die Frage: Kann und will ich meine Daten selbst erheben (eventuell mit Unterstützung weiterer Personen), oder vergebe ich die Datenerhebung an ein kommerzielles Datenerhebungsinstitut?

Entscheide ich mich für ein Datenerhebungsinstitut, sollte ich zwischen mehreren wählen können. Dazu schreibe ich mein Projekt bzw. meine Datenerhebung an mehrere Institute aus und bitte um deren Angebote. Ich prüfe die eingehenden Angebote und wäge deren jeweilige Vor- und Nachteile und vor allem die Kosten gegeneinander ab (das Billigste ist nicht immer das Günstigste!). Auf der Basis der so zustande gekommenen Informationen (und ggfs. weiterer direkter Kontakte mit den Instituten) wähle ich das Institut aus, das die Datenerhebung meines Projektes durchführen soll.

Entscheide ich mich dafür, die Datenerhebung selbst zu organisieren, muß ich Vorbereitung und Durchführung des Feldes planen; dies beinhaltet so banal erscheinende Fragen wie den Druck des Fragebogens, aber auch weniger banale Fragen wie Auswahl und Schulung von Interviewern, Möglichkeiten der Feldkontrolle usw. Vor allem beim persönlich-mündlichen Interview stellt sich dabei oft schnell das Gefühl ein, überfordert zu sein, wenn man sich mit Fragen beschäftigen muß wie: woher bekomme ich dreißig Interviewer, und wer schult sie mir so, daß sie einigermaßen erfolgreich arbeiten können? oder: habe ich überhaupt die Kapazität, 400 persönlich-mündliche Interviews in Baden-Württemberg in Eigenregie durchzuführen?

Eine weitere Frage, die ich bei der Planung meiner Datenerhebung zu beantworten habe, ist die nach der Population, die ich befragen und über die ich Aussagen treffen will: Interessieren mich alle erwachsenen deutschen Staatsangehörigen, die zum Zeit-

punkt xyz in Deutschland einen festen Wohnsitz haben, alle erwachsenen Einwohner der Stadt Mannheim, alle weiblichen Mitarbeiter ZUMAs?

Und: wie viele davon gehen in meine Befragung ein und auf welche Art und Weise? Plane ich die Befragung aller Einheiten der Population (Vollerhebung), oder beschränke ich mich auf eine Stichprobe aus dieser Population? Beim Stichprobenansatz hätten wir uns zu entscheiden zwischen einer einfachen Zufallsstichprobe (address random, random route) oder einer geschichteten Zufallsstichprobe. Oder vielleicht eine Quotenstichprobe? Oder gar eine Vollerhebung? Was sind die Vor-, was sind die Nachteile?

Während die Entscheidung zwischen einer Vollerhebung und einer Stichprobenerhebung vor allem durch die Größe der Population determiniert wird, hängt die Art der Stichprobe eher von stichprobentheoretischen Argumenten und der Intention der zu treffenden Aussagen ab. Alleine Zufallsstichproben, aber nicht Quotenstichproben lassen den Schluß von den Stichprobenparametern auf die Populationsparameter, damit Verallgemeinerung der gewonnenen Ergebnisse zu. Die Größe der Stichprobe, damit letztendlich auch die Anzahl der Fälle, die man auswerten kann, ist zum Teil abhängig von den verfügbaren finanziellen Ressourcen, sie wird vor allem aber bestimmt durch die gewünschte Tiefengliederung der vorgesehenen Auswertungen: will man die Antwort auf eine bestimmte Frage nur nach dem Geschlecht der Antworter auswerten, braucht man natürlich nicht annähernd so viele Fälle wie bei einer Auswertung nach Kombinationen von Geschlecht, Alter, Schulbildung und Körpergröße.

Ein wichtiger Schritt vor Beginn der Datenerhebung ist schließlich die Durchführung eines Pretests, der von allen Experten als unabdingbare Voraussetzung einer erfolgreichen Hauptbefragung angesehen wird. Im Pretest werden die Art und die Qualität der Meßinstrumente überprüft, ihre Eignung für die Fragestellung, ihre Verständlichkeit und Handhabbarkeit.

Wird die Hauptbefragung von einem kommerziellen Datenerhebungsinstitut durchgeführt, kann ich mich nicht unbedingt zurücklehnen und der Dinge harren, die da kommen werden; es ist zu empfehlen, in der Feldphase einen intensiven Kontakt zum Institut zu pflegen, vielleicht auch direkt vor Ort vorbeizuschauen (und sei es auch nur, um hinreichend glaubhaft und überzeugend Kontrolle zu signalisieren). Führe ich meine Umfrage selbst durch, muß ich mich mit Feldproblemen beschäftigen (ein Interviewer wird krank, Zielpersonen beschweren sich), die Feldsteuerung durchführen, die Interviewer kontrollieren etc. Dabei ist nicht immer eindeutig, welche der beiden Varianten - Institut oder Eigenregie - dem Forscher weniger Sorgen und Mühen bereitet.

Aber selbst wenn es manchmal etwas länger dauert: Irgendwann ist das Feld abgeschlossen und die Bauchschmerzen verklungen, und man freut sich auf die ersten Ergebnisse. Bevor man allerdings dorthin gelangt, muß man erst noch die Phase der Datenerfassung und Datenaufbereitung absolvieren.

2.3 Datenerfassung und Datenaufbereitung

Wenn man nicht ohnehin computergestützt befragt, hat man nämlich am Ende seines Feldes „nur“ einen mehr oder weniger großen Stapel ausgefüllter Papierfragebogen vor sich liegen. Eine effiziente Auswertung der Fragebogen setzt aber voraus, daß die Daten maschinenlesbar vorliegen, so daß gängige Auswertungsverfahren (SPSS, SAS, NSDstat usw.) zum Einsatz kommen können. Bevor es aber soweit ist, müssen die Daten erfaßt werden, die offenen Fragen verkodet, die Daten bereinigt und in einen mit dem vorhandenen Auswertungsprogramm zu bearbeitenden Systemfile übersetzt werden.

Der erste Schritt der Datenaufnahme besteht in der EDV-Umsetzung („Erfassung“) der geschlossenen Fragen, also der Fragen, zu denen es im Fragebogen nur Kreuze zu machen oder Ziffern einzutragen gilt.

Folie Pro_4: Datenerfassung und Datenaufbereitung

Datenerfassung und Datenaufbereitung

Erfassung der geschlossenen Daten

- Per Dateneingabemaske
- Per CAI-Programm

Verkodung und Erfassung der offenen Fragen

- Erstellen des Codeplans
- Verkodung
- Erfassung der verkodeten Daten

Datenbereinigung

- Fehlende Werte
- Wild codes
- Filterprüfungen
- Formale Inkonsistenzen
- Inhaltliche Inkonsistenzen

Die Datenerfassung kann erfolgen durch direkte Eingabe der Codeziffern in Dateneingabemasken, durch Simulation einer computergestützten Befragung, bei der jeder einzelne Fragebogen über das CAI-Programm erfaßt wird, oder auch automatisch per Belegleser.

Die offenen Fragen, zu denen es freie Antworttexte im Fragebogen gibt, müssen zunächst verkodet werden, bevor sie - wie die geschlossenen Fragen - erfaßt werden können. Dazu muß zu jeder offenen Frage ein Codeplan entwickelt werden, der die

wichtigsten Dimensionen enthält, die aus den Antworten auf die Frage abgeleitet werden können. Die Texte werden dann den durchnummerierten Dimensionen (dem Codeplan) zugeordnet, dadurch mit Ziffern versehen und danach wie geschlossene Fragen in den Datensatz aufgenommen.

Sind alle Daten aufgenommen, müssen sie bereinigt werden. Fehlende Werte müssen gekennzeichnet, wild codes (unzulässige Werte) ebenso korrigiert werden wie Filterfehler oder formale Inkonsistenzen (der Vierzehnjährige mit Ehepartner).

Der so bereinigte Datensatz wird dann - sofern Datenaufnahme und Datenauswertung nicht ohnehin mit dem gleichen Programm geleistet werden - in die Sprache des vorhandenen Auswertungsprogramms „übersetzt“ (z. B. vom ASCII- zum SPSS-Systemfile), so daß er analysiert werden kann. Jetzt gibt es Grund, sich auf die ersten Auswertungen zu freuen.

2.4 Auswertung

Je nach Kenntnissen und Fähigkeiten, je nach eigenen Ansprüchen und der Zielgruppe, der man seine Ergebnisse kommunizieren möchte, kann man Analysen unterschiedlicher Komplexität durchführen:

Folie Pro_5: Auswertung

Auswertung und Interpretation

Analysen

- **Univariate Analysen: Randverteilungen, Häufigkeitsauszählungen von Fragen über alle Personen**
- **Bivariate Analysen (z. B. Tabellenanalyse, bivariate Kovarianzen, bivariate Korrelationen) = Betrachtung von Zusammenhängen zwischen zwei meßbaren Variablen**
- **Multivariate Analysen (z. B. Faktorenanalyse, Clusteranalyse, multiple Regressionen) = Erklärung der Zusammenhänge zwischen zwei und mehr meßbaren oder latenten Variablen**
- **Interpretation der Ergebnisse.**

Publikation der Ergebnisse

- **Forschungsbericht**
- **Arbeitsberichte**
- **Vorträge und Kongresse**
- **Beiträge in einschlägigen Fachzeitschriften**
- **Beiträge in Sammlungen**
- **Monographien**

Beginnen sollte man die Auswertung auf alle Fälle mit univariaten Analysen, also mit der Betrachtung von Randverteilungen oder Häufigkeitsauszählungen. Im nächsten Schritt können dann bivariate Analysen folgen, bei denen nach den Zusammenhängen zwischen zwei meßbaren Variablen gefragt wird (z.B. Tabellenanalyse, bivariate Korrelationen). Bei multivariaten Analysen (z.B. Faktorenanalysen, Clusteranalysen, multiple Regressionen) schließlich wird nach Zusammenhängen zwischen mehr als zwei meßbaren oder latenten Variablen gefragt.

Da Wissenschaft als System absolut undenkbar ist ohne Kommunikation und Diskussion von Forschungsergebnissen - aber auch aus anderen, eher „materiellen“ Gründen - muß der letzte Schritt des Forschungsprozesses in der Publikation der Analysen und Ergebnisse bestehen.

Neben den auf Folie Pro_5 aufgeführten traditionellen Publikationsforen gewinnt neuerdings auch das Internet als Publikationsforum an Bedeutung, wobei hier mehr als sonstwo daran zu denken ist, daß Beiträge im Internet nicht unbedingt irgendeine Form der wissenschaftlichen, nicht einmal der redaktionellen Begutachtung durchlaufen haben müssen.

3. Befragungen als Verfahren zur Erhebung sozialwissenschaftlicher Daten

Um Informationen zu erhalten, die man dazu verwenden kann, Aussagen zur Beschreibung und Erklärung sozialer Sachverhalte zu treffen, kann man auf ein relativ breites Spektrum bewährter Verfahren zurückgreifen. Für welches man sich letztendlich entscheidet, sollte davon abhängig gemacht werden, welches der vorhandenen Verfahren am besten geeignet und angemessen erscheint, die jeweils mit dem spezifischen Forschungsinteresse verbundenen spezifischen Forschungsfragen zu beantworten.

Man kann die vorhandenen Verfahren nach unterschiedlichen Kriterien systematisch verorten, wobei allerdings darauf zu achten ist, daß man die eigentlichen Verfahren selbst von den zugrundeliegenden Forschungsdesigns unterscheidet.³

Ein wichtiges Differenzierungskriterium ist die Art der Einbeziehung des Untersuchungsobjektes in den Ablauf der Datenerhebung. Wir unterscheiden zwischen

- Verfahren, in denen das Untersuchungsobjekt aktiv und bewußt mit dem Untersucher (oder dessen Repräsentanten) interagiert
- Verfahren, in denen das Untersuchungsobjekt bewußt oder unbewußt Objekt im eigentlichen Sinne ist, also passiver Gegenstand von Handlungen des Untersuchers und
- Verfahren, in denen nicht das Untersuchungsobjekt selbst, sondern von ihm oder über ihn erstellte Daten das Interesse des Untersuchers sind.

³ Eine Panelbefragung z. B., also die wiederholte Befragung der gleichen Person in gewissen Abständen, ist vom Forschungsdesign her wesentlich von einer Einmal-Befragung verschieden, vom Verfahren her aber nun mal nichts anderes als eine Befragung.

Anders ausgedrückt sprechen wir (vgl. Schnell u.a. 1992:325) von Befragung, Beobachtung und Inhaltsanalyse.

Folie Dat_1: Datenerhebungsverfahren

Datenerhebungsverfahren

Befragungen

**Umfragen allgemein
Gruppendiskussionen
Soziometrische Verfahren**

Beobachtung

**strukturiert/unstrukturiert
teilnehmend/nichtteilnehmend
registrierend/kategorisierend usw.**

Inhaltsanalyse

**Frequenzanalyse
Valenzanalyse
Intensitätsanalyse
Kontingenzanalyse**

Neben Befragung, Beobachtung und Inhaltsanalyse gibt es einige Verfahren, die im Prozeß der Gewinnung sozialwissenschaftlicher Daten eine gewisse Sonderrolle einnehmen, z. B. nicht-reaktive Datenerhebungsverfahren (vgl. z. B. Milgram u.a. 1965; Webb u.a. 1966; Bungard & Lück 1982) oder Experimente (vgl. z. B. Zimmermann 1972; Holzkamp 1981; Atteslander 1993).

Unter der Überschrift „Entwicklung und Design von Fragebogen“ können wir alle Datenerhebungsverfahren, die nicht unter die Befragungen fallen, ignorieren. Nur bei Befragungen braucht man, so trivial das klingt, Fragebogen im eigentlichen Sinne. Überspringen wir deshalb die Bereiche Beobachtung und Inhaltsanalyse, und wenden wir uns direkt den Befragungen zu.

Wir können Befragungen unterscheiden nach dem Grad ihrer Strukturiertheit und nach ihrer Kommunikationsform.

Folie Dat_2: Arten von Befragungen

Befragungen nach dem Grad ihrer Strukturiertheit

- **Narrative, explorative oder Tiefeninterviews**
- **Leitfadeninterviews**
- **Standardisierte Interviews**

Befragungen nach der Kommunikationsform

- **Persönlich-mündliche Befragung**
- **Telefonische Befragung**
- **Schriftliche Befragung**
- **Computergestützte Befragung**

Nach dem Grad der Strukturiertheit unterscheiden wir Befragungen in

- narrative, explorative oder Tiefeninterviews
- Leitfadeninterviews
- standardisierte Interviews

Narrative, explorative oder Tiefeninterviews (vgl. z. B. Karakalos 1979; Schütze 1983; Wiedemann 1986; Lamnek 1989; Spöhring 1989) zeichnen sich durch geringe Standardisierung des angesonnenen Interviewer- wie auch des Befragtenverhaltens aus. Sie ähneln am meisten einem „normalen“ Gespräch, bei dem ein neugieriger Frager einem auskunftswilligen Antworter gegenüber sitzt und bei dem es zwar um ein bestimmtes Thema, aber ansonsten „um Gott und die Welt“ geht. Die Vorgehensweise bei der Ermittlung der benötigten Informationen obliegt dem Befrager mehr oder weniger alleine, der Antworter kann alle Informationen in der ihm genehmen Form geben.

Leitfadeninterviews (vgl. z. B. Hopf 1978; Atteslander 1993) schränken die Freiheit des Befragers insofern ein, als er nicht nur ein bestimmtes Thema, sondern vorgegebene und das Thema differenzierende Themenbereiche abfragen muß. Dabei hilft ihm ein Katalog von vorformulierten Stichworten oder Globalfragen, der sogenannte Leitfaden, den er im Prinzip in jeder beliebigen Form und Reihenfolge abarbeiten kann; er muß aber sicherstellen, daß alle Fragenbereiche behandelt worden sind. In manchen Fällen wird auch Wert darauf gelegt, daß die Fragen des Leitfadens wörtlich vorgelesen werden.

Das standardisierte Interview (vgl. z. B. Payne 1951; Erbslöh 1972; Scheuch 1973; Schuman & Presser 1981; Sudman & Bradburn 1983; Converse & Presser 1986; Fowler & Mangione 1990; Converse & Presser 1986) unterscheidet sich von den vorher genannten Formen dadurch, daß sowohl dem Interviewer als auch der befragten Person außerordentlich eng vorgegeben ist, wie das Interview-Gespräch ablaufen soll. Der Interviewer muß alle Vorgaben des Fragebogens nach Anweisung erfüllen, muß Fragen und Antwortkategorien wörtlich vorlesen, Listen oder Kartenspiele vorlegen usw., die befragte Person muß ihr Antwortverhalten an den Erwartungen des Interviewers ausrichten, das heißt in der Regel ihre Antwort in vorgegebene Antwortkategorien einordnen. Das standardisierte Interview fordert für alle Befragte gleiche Befragungsbedingungen; sofern ein Interviewer auftritt, beschränkt sich seine Aufgabe idealerweise auf das reine Vorlesen der Fragen und das kommentarlose Registrieren der Antworten.

Nach der Kommunikationsform, in der eine Befragung durchgeführt wird, kann man Befragungen unterteilen in

- persönlich-mündliche Befragungen
- schriftliche Befragungen und
- telefonische Befragungen.

Das computergestützte Interview stellt, sieht man von allerneuesten Entwicklungen ab, nur eine technologisch verbesserte Form der genannten Befragungsverfahren dar, aber keine neue Datenerhebungsmethode.

Die persönlich-mündliche Befragung galt lange Zeit als der „Königsweg“ der sozialwissenschaftlichen Datenerhebung. Man glaubte - und gelegentlich wird das heute noch geglaubt - daß durch persönlich-mündliche Befragungen gültigere und verlässlichere Daten gewonnen werden könnten, als dies bei schriftlichen oder telefonischen Umfragen der Fall sei. Diese Bewertung verschaffte der persönlich-mündlichen Befragung in den empirischen Sozialwissenschaften eine lange Zeit uneingeschränkte Sonderstellung, die bis heute nachwirkt.

Die schriftliche Befragung (vgl. z. B. Hafermalz 1976; Dillman 1978; Berdie, Anderson & Niebuhr 1986), die in Form der postalischen Befragung zur Zeit eine gewisse Renaissance erlebt, profitierte, wie die telefonische, von den Problemen der persönlich-mündlichen Befragung. Postalische Interviews sind in der Regel billiger als persönlich-mündliche; sie erfordern weniger organisatorischen Aufwand als jene, es gibt keine mühsame Interviewer-Verwaltung, und die Anforderungen an die Feldsteuerung sind begrenzt. Die Zielperson kann selbst bestimmen, wann sie den Fragebogen ausfüllen will. Der Druck des Interviewers, der gerade zur denkbar schlechtesten Zeit oder unter denkbar ungünstigen Umständen vor der Tür steht, entfällt. Die Zielperson kann die Bearbeitung des Fragebogens beliebig unterbrechen, nachdenken, ggfs. in Unterlagen nachsehen. Sowohl die Selbstbestimmtheit als auch die Anonymität der Befragungssituation führen dazu, daß die Antworten bei der postalischen Befragung als „ehrlicher“ gelten, als „überlegter“ und „durchdachter“. Auf der anderen Seite besteht die Gefahr einer Verzerrung der Stichprobe durch „Selbstselektion“: Personen, die gewohnt sind, schriftlich zu arbeiten oder Formulare auszufüllen, nehmen eher teil als solche, die mit dieser Art der Kommunikation weniger vertraut sind. Die Datenerhe-

bungssituation ist, weil kein Interviewer involviert ist, nicht kontrollierbar und man kann deshalb nicht sicher sein, ob diejenige Person, die den Fragebogen ausgefüllt hat, auch wirklich die Zielperson war und unter welchen Bedingungen der Fragebogen ausgefüllt worden ist. Schließlich können im postalischen Interview keine „spontanen“ Reaktionen der Zielpersonen erfaßt werden. Ein letzter Nachteil der postalischen Befragung schließlich ist die außerordentlich große Schwankungsbreite der Rücksenderate (Hippler 1988:244).

Die telefonische Befragung (vgl. z. B. Lucas & Adams 1977; Lavrakas 1987; Groves u. a., Hrsg, 1988; Frey 1989; Frey u. a. 1990) hat nach langer Anlaufzeit in den Sozialwissenschaften ihr „quick-and-dirty“-Image ablegen können. Gründe dafür sind vor allem die Vollversorgung mit Telefonen zumindest in den westlichen Ländern und die Fortschritte in der Computertechnologie, dadurch die Möglichkeit computergestützter Telefonbefragungen. Telefonbefragungen können unter bestimmten Umständen, müssen aber nicht billiger sein als persönlich-mündliche Befragungen. Eindeutig für die Telefonbefragungen sprechen kürzere Feldzeiten; dadurch liegen die Daten schneller vor als bei allen anderen Datenerhebungsverfahren. Bei der Stichprobenrealisierung ist wichtig, daß beim Telefoninterview regionale Klumpungen vermieden werden können. In der Befragungssituation selbst zeichnet sich das Telefoninterview durch leichtere Kontaktmöglichkeiten zu den Zielpersonen, die Möglichkeiten zu einer besseren Kontrolle der Interviewer und eine größere Anonymität der Befragungssituation aus.

Computergestützte Verfahren der Datenerhebung haben einen überaus schnellen und eindrucksvollen Siegeszug in der Umfrage- und empirischen Sozialforschung genommen. Sie stellen aber, sieht man von allerneuesten Entwicklungen ab, „nur“ eine technologisch verbesserte Form der genannten Befragungsverfahren dar, aber keine neue Datenerhebungsmethode. Mit dem Einsatz fortgeschrittener Programme zum computer assisted interviewing (CAI) ergeben sich eine Reihe methodisch wichtiger und interessanter Möglichkeiten, wie z.B. die Abarbeitung auch extrem komplexer Befragungsmuster durch differenzierte Filterführungen und weitgehend individualisiertem, d.h. der jeweiligen Befragungsperson und ihren Antworten angepaßtem Befragungsverlauf. Darüber hinaus ermöglicht der Einsatz entsprechender CAI-Programme, daß man in jeder Phase der Befragung schnell Randverteilungen oder Tabellen zur Beobachtung des bisherigen Feldverlaufes erstellen kann. Ebenso können zu jedem Zeitpunkt während der Befragung auch inhaltliche Auswertungen vorgenommen werden. Der fertige Datensatz kann direkt und unmittelbar nach Feldende erstellt und bearbeitet werden. Die Daten selbst weisen eine größere Genauigkeit und eine bessere formale Qualität auf als Daten, die nicht computergestützt erhoben worden sind (keine Filterfehler, keine ungültigen Werte, usw.).

Ein Vergleich der Befragungsverfahren aus technischer und methodischer Sicht (Hippler & Schwarz 1990) zeigt die Vor- und Nachteile der unterschiedlichen Verfahren (s. Seite 17):

Folie Dat_3: Vergleich der Datenerhebungsverfahren unter technisch-methodischen Aspekten

Vergleich der Verfahren unter technisch-methodische Aspekten

	Telefonisch	Persönl.-mündlich	schriftlich
Kosten	mittel bis hoch	hoch	niedrig
Stichprobengröße	groß	mittel	sehr groß
Ablaufsteuerung	sehr hoch (CATI)	mittel	keine
Datengenauigkeit	hoch	mittel	mittel
Interviewereinfl.	mittel	hoch	kein
Anonymität	hoch?	mittel bis gering	hoch
Stichprobe	teilw. beschränkt	ohne Beschränkung	beschränkt
Fragenkomplexität	gering?	hoch	gering
Ausschöpfung	hoch?	hoch	mittel
Länge	mittel/lang?	lang	kurz/mittel

Zu ähnlichen Ergebnissen kommen auch andere Autoren, z. B. de Leeuw und Van der Zouwen (1988).

4. Fragebogenerstellung

Kommen wir nun zum eigentlichen Schwerpunkt dieses Berichts, nämlich der Fragebogenerstellung. Im einzelnen wollen wir uns dabei mit folgenden Fragen beschäftigen:

Folie Fb_1: Fragebogenerstellung

Fragebogenerstellung

1. Zielsetzungen eines Fragebogens
2. Kognitionspsychologische Grundlagen der Befragung
3. Arten von Fragen
4. Fragen- und Antwortformulierungen
5. Skalen
6. Aufbau des Fragebogens und Fragensukzession
7. Befragungshilfen, Filter und Layout

Trotz dieser umfangreichen Gliederung, die uns auf die Vielzahl der Aspekte verweist, die für die Fragebogenerstellung wichtig sind, könnte alles ganz einfach sein, nämlich so:

Folie Fb_2: Nicht der Interviewer...

**Nicht der Interviewer,
der Fragebogen muß
schlau sein☺**

**Gerhard Schmidtchen,
Der Anwendungsbereich betriebssoziologischer Umfragen,
Bern 1962, S. 9**

Damit wäre alles gesagt, und wir könnten zum nächsten Thema übergehen. Dennoch sollten wir uns doch etwas detaillierter mit der Fragebogenerstellung beschäftigen. Beginnen wir mit einem Beispiel (Schwarz 1991):

Bei einer Befragung sollte unter anderem auch gemessen werden, was die Befragten so ganz allgemein von der CDU hielten. Gefragt wurde: „Alles in allem, was halten Sie ganz allgemein von der CDU?“. Vorgelegt wurde eine Skala von 1=„überhaupt nichts“ bis 11=„sehr viel“. Der Mittelwert (arithmetisches Mittel) lag bei - für die CDU wenig schmeichelhaften - 3,4.

Damit war die CDU natürlich ganz und gar nicht zufrieden, und natürlich war der CDU zu helfen. Man befragte eine vergleichbare Stichprobe mit der selben Frage, und der Mittelwert lag - schon besser - bei 5,2.

Auf die Frage, ob es nicht auch etwas oberhalb des Skalenmittelwertes ginge, wurde eine dritte Befragung durchgeführt, gleiche Frage, vergleichbare Stichprobe, nur daß der Wert jetzt mit 6,5 doch deutlich Positiveres verhiess. Also, man war's zufrieden.

Selbstverständlich ist die Geschichte fiktiv und frei erfunden.⁴ Nicht fiktiv und frei erfunden sind allerdings die in dieser Geschichte berichteten Werte; sie sind Ergebnis einer Studie, die bereits vor längerem bei ZUMA durchgeführt worden ist (Schwarz 1991). Sehen wir uns das noch einmal in der Übersicht an:

⁴ Die CDU möge mir diesen Einstieg in das Thema „Fragebogenentwicklung“ bitte nachsehen; an späterer Stelle wird deutlich werden, warum hier die CDU erhalten mußte und keine andere bestehende oder fiktive Partei.

Folie Fb_3: Bewertung der CDU

Alles in allem: Was halten Sie ganz allgemein von der CDU?

11	= sehr viel		
10			
9		erster Mittelwert:	3,4 (A)
8			
7			
6		zweiter Mittelwert:	5,2 (B)
5			
4			
3		dritter Mittelwert:	6,5 (C)
2			
1	= überhaupt nichts		

Wohlgermerkt: immer die gleiche Fragestellung, immer vergleichbare Stichproben, die befragt worden sind, aber: deutlich unterschiedliche Skalenwerte. Des Rätsels Lösung - ich bitte noch um etwas Geduld - wird später preisgegeben. Hier ist nur wichtig, daß trotz offensichtlich gleicher Voraussetzungen deutlich unterschiedliche Ergebnisse zustande gekommen sind.

Selbst wenn man sich Umfrageergebnisse wie dieses genüßlich vor Augen hält - die grundsätzliche Bedeutung der Befragung als Methode zur Erfassung sozialwissenschaftlich relevanter Informationen steht außer Zweifel (Phillips 1971; Kaase, Ott und Scheuch 1983).

In einem gewissen Widerspruch dazu befindet sich die relativ geringe Aufmerksamkeit, die man in der methodischen Forschung dem Fragebogen als dem zentralen Instrumentarium der Befragung gewidmet hat.

Wer sich mit Fragebogen und der Konstruktion von Fragebogen beschäftigt, findet denn auch in der Literatur häufig zwar teilweise recht bemühte (vgl. z. B. Karmasin & Karmasin 1977), aber oft sehr empiristische, sehr punktuelle, nicht verallgemeinerbare und schließlich auch von Fall zu Fall schlicht banale "Ratschläge": Fragen sollten kurz sein, konkret, eindeutig, sie sollten den Befragten nicht überfordern, usw. usw.

Mit solcherart wertvollen Hinweisen ausgestattet ist demzufolge ganz einfach auch die Aufgabe des Fragebogengestalters. Nach einer ähnlich allgemeinen Auflistung von Ratschlägen zur Frageentwicklung schreiben Karmasin & Karmasin (1977:174):

„Dieser eher weite und sich nur auf die zentralen Prinzipien beschränkende Raster, der naturgemäß eine Fülle von Variationen zuläßt, muß durch das Sprachgefühl des Forschers ergänzt werden, der intuitiv und aus seiner Erfahrung als kompetenter Sprecher der zugrunde liegenden Sprache zu entscheiden hat, welche Formulierung adäquat ist, um für einfache Sprachstile verständlich zu sein, für alle Befragten eindeutig interpretierbar und geeignet, genau den fraglichen Sachverhalt zu erfassen.“ (Sic!)

Ratschläge wie diese tragen zur Aufrechterhaltung der irrigen Meinung bei, daß die Fragebogengestaltung eine auf individueller Erfahrung basierende „Kunstlehre“ sei - der Begriff ist entnommen aus dem Buch von Payne (1951) „The Art of Asking Questions“. Daß ältere Publikationen wie Noelle (1963), Mayntz u.a. (1971) oder Scheuch (1973) den Begriff „Kunstlehre“ dem Wort oder dem Sinn nach verwenden, kann noch entschuldigt werden; die Entwicklung von Fragebogen bis in die jüngste Zeit hinein als Kunstlehre zu verstehen, halte ich allerdings für problematisch. Denn:

Anfang der Achtziger Jahre begannen systematische Versuche, Fragebogengestaltung als integrierten Bestandteil innerhalb eines theoretischen Konzepts der Befragung zu verstehen (z. B. Schuman & Presser 1981; Dijkstra & Van der Zouwen 1982; Sudman & Bradburn 1983). Sozialpsychologen und Psychologen arbeiten gemeinsam mit Umfrageforschern verstärkt an der Erforschung der kognitiven und kommunikativen Prozesse, die der Befragungssituation zugrunde liegen (z. B. Jabine u.a., Hrsg, 1984; Hippler, Schwarz & Sudman, Hrsg, 1987; Jobe & Loftus, Hrsg, 1991; Schwarz & Sudman, Hrsg, 1992; Schwarz & Sudman, Hrsg, 1996), und sie leiten aus ihrer Forschung auch konkrete Empfehlungen für die Gestaltung von Fragebogen ab (Schwarz & Strack 1988; Schwarz 1990; Schwarz 1991).

Auf solcherart Bemühungen werde ich nachher noch kurz zurückkommen, wenn ich die kognitionspsychologischen Grundlagen der Befragung beschreibe. Zunächst möchte ich mich mit den Zielsetzungen eines Fragebogens beschäftigen, danach - als Hauptteil - mit praktischen Aspekten der Fragebogenerstellung.

4.1 Zielsetzungen eines Fragebogens

Was ist eigentlich ein Fragebogen? Versuchen wir eine Definition (Porst 1996:738):

Folie Fb_4: Definition „Fragebogen“

Definition „Fragebogen“:

Ein Fragebogen ist eine mehr oder weniger standardisierte Zusammenstellung von Fragen, die Personen zur Beantwortung vorgelegt werden mit dem Ziel, deren Antworten zur Überprüfung der den Fragen zugrunde liegenden theoretischen Konzepte und Zusammenhänge zu verwenden. Somit stellt ein Fragebogen das zentrale Verbindungsstück zwischen Theorie und Analyse dar.

Bei der Erstellung eines Fragebogens ist zunächst auf die qualitative und quantitative Übereinstimmung des Instrumentariums mit dem Forschungsziel zu achten. Unter quantitativer Übereinstimmung des Fragebogens mit dem Forschungsziel verstehe ich die vollständige, unter qualitativer Übereinstimmung die inhaltlich angemessene Operationalisierung aller Hypothesen bzw. Variablen des zugrunde liegenden theoretischen Konzepts: alle theoretischen Begriffe müssen im Fragebogen enthalten sein; die Frageformulierungen, die Antwortkategorien und die Art der Frage müssen geeignet sein, die angezielten Informationen valide (d. h. gültig) und reliabel (d. h. zuverlässig) zu erfassen. Bereits bei der Entwicklung des Fragebogens ist darauf zu achten, daß zu erwartende (auch unerwünschte) Verhaltensweisen und Reaktionen der Zielpersonen (man nennt das response set) wenn schon nicht ausgeschaltet, so doch möglichst kontrolliert werden können. Dazu ist es wichtig, die kognitiven und kommunikativen Prozesse zu kennen, die der Befragungssituation zugrunde liegen (vgl. z. B. Tourangeau 1987; Strack & Martin 1987; Strack & Schwarz 1992) und sie bereits bei der Fragebogenentwicklung zu berücksichtigen. Ich will dazu nur einen kurze Übersicht geben, sozusagen ein Anstoß, sich intensiver mit der Literatur zu beschäftigen.

4.2 Kognitionspsychologische Grundlagen der Befragung

Die Teilnehmer an einer Befragung müssen mehrere Aufgaben lösen:

Folie Fb_5: Aufgaben des Befragten

Die Teilnehmer an einer Befragung müssen ...

- 1) die gestellte Frage verstehen
- 2) relevante Informationen aus dem Gedächtnis abrufen
 - bei Einstellungsfragen: eine bereits gebildete Meinung aus der Erinnerung abrufen oder relevante Informationen abrufen, die es erlauben, ein Urteil zum Befragungsgegenstand zu bilden
 - bei Verhaltensfragen: relevante Ereignisse erinnern, sie gegebenenfalls zu datieren, eventuell die Zahl der relevanten Ereignisse zu bestimmen oder sie zu schätzen.
- 3) auf der Basis dieser Informationen ein Urteil bilden
- 4) dieses Urteil gegebenenfalls in ein Antwortformat einpassen
- 5) gegebenenfalls ihr "privates" Urteil vor Weitergabe an den Interviewer "editieren"

Vgl. Strack und Martin (1987)

Die Suche nach Information ist natürlich nicht grenzenlos. Personen hören auf zu suchen, wenn sie genug Informationen erinnert haben, um sich mit hinreichender subjektiver Sicherheit ein Urteil bilden zu können. Dieses Urteil beruht primär auf der Information, die dem Befragten zuerst in den Sinn kommt.

Information kann chronisch oder situativ verfügbar sein. Chronisch bedeutet, die Information ist im Gedächtnis leicht verfügbar, weil man zum Beispiel bereits öfters über den gefragten Sachverhalt nachgedacht hat; situativ bedeutet, die Information kommt nur unter bestimmten Bedingungen in Erinnerung, etwa durch das Interview selbst (z. B. aufgrund von Vorfragen).

Befragungen gelten als eine spezielle Form der Konversation. Warum? Sie sind "conversation at random" und weichen in vielen Fällen von der normalen Gesprächsführung ab. Einige dieser Abweichungen werden normalerweise problemlos umgesetzt, z. B. das konzentrierte Fragen und Antworten. Andere Regeln der Alltagskonversation bleiben nicht ohne Einfluß auf das Interview und sein Ergebnis. So glauben Befragungspersonen im Einklang mit den Grundregeln der kooperativen Kommunikation (s. Seite 23; vgl. Grice 1975), der Interviewer bzw. der Forscher tue alles, um informativ, der Wahrheit folgend, bedeutungsvoll und zielgerichtet sowie eindeutig, oder besser: unzweideutig zu sein.

Folie Fb_6: Grundregeln der kooperativen Kommunikation**Grundregeln der (kooperativen) Kommunikation:**

Maxim of Quantity: Make your contribution as informative as is required, but not more informative than is required.

Maxim of Quality: Try to make your contribution one that is true. That is, do not say anything you believe to be false or lack adequate evidence for.

Maxim of Relation: Make your contribution relevant to the aims of the ongoing conversation.

Maxim of Manner: Be clear. Try to avoid obscurity, ambiguity, wordiness, and disorderliness in your use of language.

Nach H. P. Grice (1975)

Weicht der Interviewer von diesen Regeln ab, versucht die Person, der gestellten Frage einen subjektiven Sinn zu geben, indem sie den Kontext der Frage auszuloten und ihre Bedeutung festzulegen versucht. Dies hat natürlich Einfluß auf das Antwortverhalten.

Damit genug zu den kognitionspsychologischen Grundlagen der Befragung. Ich komme bei Gelegenheit darauf zurück, will mich aber jetzt mit den „praktischeren“ Problemen der Fragebogengenerierung beschäftigen. Dabei will ich auf folgende Aspekte kurz eingehen: Arten von Fragen, Fragen- und Antwortformulierungen, Skalen, Fragebogenaufbau und Fragenabfolge, Befragungshilfen, Filter und Layout.

4.3 Arten von Fragen

Beginnen wir mit den Arten von Fragen. Fragen können nach ihrem Inhalt bzw. ihrer Zielrichtung und nach ihrer Form unterschieden werden. Die Aufteilung nach inhaltlichen Gesichtspunkten ist (relativ) beliebig; eine einfache Unterteilung ergibt sich z. B. in Fragen nach Einstellungen oder Meinungen, Fragen nach Überzeugungen und Wertorientierungen, Fragen nach Wissen und Verhalten und Fragen nach Eigenschaften der Befragungsperson (andere Einteilungen finden sich z. B. bei Dillman 1978; Davis 1980; Porst 1985).

Nach ihrer Form können Fragen ebenfalls differenziert werden; wichtig ist vor allem die Unterscheidung von Fragen aufgrund des Grades ihrer Strukturiertheit in offene, halboffene und geschlossene Fragen (s. Seite 24).

Folie Fb_7: Arten von Fragen

38. Wie stark interessieren Sie sich für Politik: sehr stark, stark, mittel, wenig oder überhaupt nicht?

- Sehr stark.....
- Stark
- Mittel
- Wenig
- Überhaupt nicht

16. Welche Staatsbürgerschaft haben Sie? Wenn Sie die Staatsbürgerschaft mehrerer Länder besitzen, nennen Sie mir bitte alle.

Interviewer: Mehrfachnennungen möglich

- Staatsbürgerschaft von:
- Deutschland
 - Dänemark.....
 - Frankreich.....
 - Italien.....
 - Anderes Land

S9 Welche berufliche Tätigkeit üben Sie in Ihrem Hauptberuf aus? Bitte beschreiben Sie mir Ihre berufliche Tätigkeit genau.

Interviewer: Bitte genau nachfragen:

Hat dieser Beruf, diese Tätigkeit noch einen besonderen Namen?

Interviewer: Bitte Liste S2 vorlegen!

S2 Als nächstes kommen jetzt Fragen zu Ihrer Ausbildung, Ihrem Beruf und Ihrer Familie. Beginnen wir mit Ihrer Ausbildung: Welchen allgemeinbildenden Schulabschluß haben Sie?

Interviewer: Nur **e i n e** Nennung möglich. Nur höchsten Schulabschluß angeben lassen!

- A Noch Schüler..... ⇨ S4
 - B Schule beendet ohne Abschluß
 - C Volks-/Hauptschulabschluß bzw. Polytechnische Oberschule mit Abschluß 8. oder 9. Klasse
 - D Mittlere Reife, Realschulabschluß bzw. Polytechnische Oberschule mit Abschluß 10. Klasse.....
 - E Fachhochschulreife (Abschluß einer Fachoberschule etc.).....
 - F Abitur bzw. Erweiterte Oberschule mit Abschluß 12. Klasse (Hochschulreife)
 - G Anderen Schulabschluß, und zwar:
- _____

Bei der offenen Frage (auf Seite 24 Frage S9) wird nur die Frage vorgelesen; es gibt keine Antwortkategorien, die Befragungsperson antwortet in ihren eigenen Worten, und der Interviewer protokolliert ihre Aussage möglichst wörtlich.

Bei geschlossenen Fragen gibt es dagegen eine begrenzte und definierte Anzahl von möglichen Antwortkategorien, in welche die Befragungsperson ihre Antwort einpassen muß; dabei ist zu unterscheiden zwischen Fragen mit nur einer zulässigen Antwort (Einfachnennung), bei der sich die Befragungsperson für eine der vorgegebenen Alternativen entscheiden muß (auf Seite 24 Frage 38) und Fragen mit mehr als einer zulässigen Antwort (Mehrfachnennung), bei denen sie mehrere der vorgegebenen Kategorien auswählen kann (auf Seite 24 Frage 16).

Offene Fragen haben den Vorteil, daß sie den Befragungspersonen die Möglichkeit bieten, so zu sprechen, wie sie es gewohnt sind; der Nachteil liegt u. a. darin, daß die Ergebnisse sehr stark von der Verbalisierungsfähigkeit des Befragten abhängen und nicht unbedingt von seiner tatsächlichen Einstellung zu dem gefragten Sachverhalt. Ein weiterer zentraler Nachteil der offenen Frage ist der immense Verkodungsaufwand, der mit ihrer Auswertung verbunden ist.

Geschlossene Fragen sind schneller abzuarbeiten, haben aber u. a. den Nachteil, daß sich Befragungspersonen häufig nicht in den vorgegebenen Kategorien wiederfinden können.

Halboffene Fragen (auf Seite 24 Frage S 2) schließlich sind eher ein Ergebnis von Entscheidungsschwierigkeiten des Fragebogenentwicklers, kommen aber in der Praxis sehr häufig vor: einer an sich geschlossenen Frage wird eine zusätzliche Kategorie (z. B.: "Sonstiges, bitte nennen") angehängt, die wie eine offene Frage beantwortet werden kann. Eine halboffene Frage bietet sich immer dann an, wenn das tatsächliche Universum möglicher Antworten auf eine Frage zwar gut abgeschätzt (geschlossene Frage), aber nicht definitiv bestimmt werden kann (offene Frage).

Die Diskussion um offene oder geschlossene Fragen ist wohl fast so alt wie die moderne Umfrageforschung selbst (Lazarsfeld 1944; Krech und Crutchfield 1948); allerdings ist eine Entscheidung zugunsten der einen oder der anderen Fragenform lange Zeit wohl eher durch Erfahrung, common sense, Institutsroutinen und forschungspraktische Überlegungen bestimmt worden als durch systematische Forschung (Converse 1984).

Aufgrund kognitionspsychologischer Arbeiten weiß man heute allerdings, daß die Entscheidung für eine offene oder für eine geschlossene Frage nicht alleine das Registrieren der Antwort und den Aufwand bei der Auswertung beeinflusst, sondern bedeutende inhaltliche Auswirkungen auf das Ergebnis zu dieser Frage haben kann, weil offene und geschlossene Fragen unterschiedliche kognitive Anforderungen an die Befragungsperson stellen (Schuman & Presser 1977; Bishop u.a. 1988).

Kognitionspsychologische Forschungen haben weiterhin gezeigt, daß auch die bei einer geschlossenen Frage vorgegebenen Antwortkategorien großen Einfluß auf das Ergebnis einer Frage haben können. Antwortkategorien haben nämlich nicht nur die vom

Wie ist das Ergebnis zu erklären? Aufgrund der Regeln der kooperativen Kommunikation unterstellen Befragte, daß die Ihnen vorgelegte Skala einen Sinn macht, und der Sinn kann nur darin bestehen, daß die Skala real vorhandene oder als real vorhanden angenommene Verteilungen repräsentiert. Bei einer sehr starken Differenzierung im unteren Bereich der Skala (linke Seite der Folie) gehen die Befragten demzufolge davon aus, daß die meisten Leute wohl weniger als 2 Stunden fernsehen - warum sonst hätte man gerade diese starke Differenzierung im unteren Bereich gewählt? Sie stufen sich also da ein, wo sie sich mit vielen anderen in guter Gesellschaft wähnen. Das gleiche machen sie aber auch bei der anderen Skala (rechte Seite der Folie), nur daß die Differenzierung im oberen Bereich ihnen ein Volk von Vielsehern suggeriert - auch hier stuft man sich eher da ein, wo mutmaßlich auch die meisten anderen sein werden.

Das heißt verallgemeinert: Befragungspersonen gehen davon aus, daß die ihnen vorgelegte Skala sinnhaft konstruiert ist, und sie nehmen an, daß sie in ihren Ausprägungen tatsächliche Verteilungen der Population widerspiegeln. Die Informationen, die sie aus dem Wertebereich einer Skala erschließen, nutzen sie gleich mehrfach: zum einen ziehen sie den Wertebereich als Bezugsrahmen für ihre eigene Verhaltenshäufigkeit heran (und geben zum Beispiel höhere Frequenzen für ihr eigenes Verhalten an, wenn die Skala höhere Häufigkeiten vorgibt), zum zweiten entnehmen sie ihrer eigenen Platzierung auf der Skala Informationen über die relative Häufigkeit ihres Verhaltens verglichen mit dem Verhalten anderer und berücksichtigen dies, wenn sie komparative Urteile bilden; und schließlich ziehen sie bei nicht eindeutigen Fragen die Skalen als Interpretationshilfen für den vermeintlichen Sinn der Frage heran.

Auswirkungen auf das Antwortverhalten können sich auch aus der Reihenfolge ergeben, in der Antwortkategorien präsentiert werden, insbesondere wenn eine größere Anzahl möglicher Antwortkategorien vorgegeben wird: Befragte tendieren entweder dazu, zuerst genannte Kategorien auszuwählen ("primacy effect"), oder sie entscheiden sich eher für zuletzt genannte Kategorien ("recency effect"). Welcher der Effekte zum Wirken kommt, hängt von der Art der Präsentation ab: werden Listen mit einer größeren Anzahl von Antwortkategorien optisch präsentiert, kommt es eher zum primacy effect, werden die Antwortkategorien vorgelesen, ist eher mit einem recency effect zu rechnen.

4.4 Fragen- und Antwortformulierungen

Kommen wir zu den Frage- und Antwortformulierungen. Es gibt eine Reihe von "Faustregeln" zur Verbalisierung von Fragen und Antwortkategorien in einem Fragebogen; sie finden sich in jedem Lehrbuch der empirischen Sozialforschung und gehen explizit oder implizit zumeist auf Payne (1951) zurück. Wir lesen dort (und später z. B. bei Karmasin & Karmasin 1977; Dillman 1978; Converse & Presser 1986; Schnell u.a. 1992), daß Fragen und Antworten einfach, kurz und konkret formuliert sein, keine Fremdworte und keine unverständlichen Begriffe enthalten sollen; sie sollen nicht suggestiv, semantisch nicht positiv oder negativ befrachtet sein, nicht hypothetisch; sie sollen die Befragungsperson nicht überfordern, aber auch nicht trivial klingen; Fragen sollten eindeutig sein, nicht mehrere Stimuli oder doppelte Verneinungen enthalten.

Diese und ähnliche Regeln haben natürlich eine gewisse Berechtigung und machen Sinn, allerdings wird man bei näherer Beschäftigung damit schnell feststellen, daß sie wirklich nur als sehr grobe Faustregeln zu verwenden sind (während die eine Zielperson durch eine Frage überlastet ist, erscheint sie der nächsten schon als trivial).

4.5 Skalen

Wenden wir uns den Skalen zu. Der Beantwortung einer Frage liegt technisch betrachtet grundsätzlich der Prozeß des Messens zugrunde, worunter wir jegliche regelhafte und kodifizierte Benennung bestimmter Aspekte oder Ausprägungen einer manifesten (z. B. Verhalten) oder latenten (z. B. Einstellungen) Variablen mit Hilfe von Symbolen oder Zahlen verstehen. Das dem Meßvorgang zugrunde gelegte Bezugssystem bezeichnen wir als Skala. Einige Skalen finden sich auf dieser Folie:

Folie Fb_9: Skalentypen

Skalentypen											
	trifft voll und ganz zu		trifft eher zu		trifft eher nicht zu		trifft überhaupt nicht zu				
	<input type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>				

unwichtig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	sehr wichtig		

unwichtig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	sehr wichtig		

links	1	2	3	4	5	6	7	8	9	10	rechts
rechts	10	9	8	7	6	5	4	3	2	1	links

sehr unsympathisch	-3	-2	-1	0	+1	+2	+3	sehr sympathisch			

Zunächst einmal unterscheidet man verbalisierte von numerischen Skalen, wobei erstere sich dadurch kennzeichnen, daß jedem Skalenpunkt eine verbale Benennung eindeutig zugeordnet ist, während bei der numerischen Skala nur die Endpunkte verbalisiert, die Skalenwerte ansonsten durch Ziffern repräsentiert oder leer sind.

Skalen mit einer ungeraden Anzahl von Skalenpunkten und einer formalen Mittelkategorie (ungerade Skalen) stehen Skalen mit einer geraden Anzahl von Skalenpunkten ohne formale Mittelkategorie (gerade Skalen) gegenüber; die optimale Anzahl der Skalenpunkte ist ebenso von Bedeutung (und strittig) wie die Frage, ob man Skalenwerte von rechts nach links oder von links nach rechts auf- oder absteigen läßt, ob man eindimensionale (z. B. von 1 nach 7) oder zweidimensionale (z. B. von -3 bis +3) Skalen einsetzt, und anderes mehr. Die Diskussion um die „richtige“ Skala kann hier nicht wiedergegeben werden; sie reicht nachweislich zurück bis ins Jahr 1915.

In der Praxis haben sich unter dem Gesichtspunkt der Diskriminierungsfähigkeit numerische Skalen mit sieben plus/minus zwei Skalenpunkten bewährt (Cox 1980; empirisch überzeugend Stadtler 1983). Unter Auswertungsgesichtspunkten ist dabei wichtig, daß solche Skalen den Charakter intervallskalierter Variablen besitzen oder zumindest zu besitzen glaubhaft machen.

Die Frage, ob man eine Mittelkategorie zulassen will oder nicht, ist kein rein technisches Problem. Bietet man eine Mittelkategorie an, so läuft man Gefahr, daß Personen sie als Ausweichmöglichkeit nutzen, weil sie sich nicht auf die eine oder andere Seite der Skala einstufen wollen (mit Einführung der Mittelkategorie reduziert sich der Anteil der "weiß nicht"-Antworten - sofern eine solche Kategorie vorgegeben war - deutlich; vgl. Molenaar 1982). Zwingt man sie zu einer Einordnung, indem man die Mittelkategorie wegläßt, nimmt man ihnen die Chance, sich bewußt in der Mitte einzustufen, wobei diese bewußte Einstufung kein Ausweichen darzustellen braucht, sondern die tatsächliche und inhaltliche korrekte Platzierung in der Mitte der Skala.

Unabhängig von der Art der Skala ist es grundsätzlich wichtig zu wissen, daß auch Skalen (selbst die relativ neutral erscheinenden numerischen Skalen) keineswegs den Charakter eines formalen, sachlichen Meßinstruments aufweisen. In Experimenten konnte vielmehr nachgewiesen werden, daß Skalen, die formal äquivalent, aber unterschiedlich beziffert oder benannt sind, das Antwortverhalten von Personen deutlich unterschiedlich beeinflussen können (s. Seite 30):

Folie FB_10: Beispiel Erfolg im Leben

Wie erfolgreich waren Sie bisher in Ihrem Leben? Sagen Sie es bitte nach dieser Leiter hier <es folgt die formale Skalenerklärung>.

außerordentlich	10	+ 5
	9	+ 4
	8	+ 3
	7	+ 2
	6	+ 1
	5	0

	4	- 1
	3	- 2
	2	- 3
	1	- 4
überhaupt nicht	0	- 5

Nennungen von 4 bis 0:

34 %

Mittelwert: 6,4

Nennungen von - 1 bis - 5

13 %

Mittelwert: 7,3

Die Differenz der Mittelwerte ist hoch signifikant.

Institut für Demoskopie Allensbach, IfD-Studie 5.007, Juli 1988, bundesweite Repräsentativbefragung, N = 1.032

Bei diesem Beispiel sind es schlicht die Minuszeichen vor den Ziffern, die den unteren Skalenwerten einen negativen Anstrich vermitteln und von daher nicht gerne gewählt werden. Der Wert 0 auf der linken Seite wird als „nicht erfolgreich“ interpretiert, nicht als „negativ erfolgreich“ und ist von daher lange nicht so deprimierend wie der Wert -5 auf der rechten Seite, der als „negativ erfolgreich“ interpretiert wird.

4.6 Aufbau des Fragebogens und Fragensukzession

Neben Skalen und Frageformulierungen spielt die Fragensukzession eine wichtige Rolle bei der Entwicklung von Fragebogen. Unter Fragensukzession versteht man die Reihenfolge, mit der Fragen im Fragebogen angeordnet und der Befragungsperson im Verlaufe der Befragung vorgelegt werden. Sie spielt vorrangig bei der persönlichen und bei der telefonischen Befragung eine Rolle; bei der schriftlichen Be-

fragung dagegen könnte die Sukzession irrelevant werden, weil nicht ausgeschlossen werden kann, daß die Befragungsperson den Fragebogen entgegen der Intention des Forschers nicht sukzessive, sondern in beliebiger Abfolge bearbeitet, also beim Ausfüllen vor- und zurückblättert.

Auch für die Fragensukzession gibt es eine Vielzahl technischer Regeln, die mehr oder weniger "intuitiv" definiert sind (vgl. z. B. Karmasin & Karmasin 1977:197ff); auf der anderen Seite haben kognitionspsychologische Forschungen auch hier eine Reihe empirischer Hinweise für die Fragebogengestaltung geliefert.

Zu den eher intuitiven Regeln zählt, daß eine Befragung mit spannenden, themenbezogenen und die Befragungsperson persönlich betreffenden, aber technisch einfach zu bearbeitenden Fragen beginnen sollte, um die Motivation der Befragungsperson zur weiteren Teilnahme aufrechtzuerhalten oder sogar noch zu erhöhen („Eisbrecher-“ oder „Aufwärmfragen“). Einstiegsfragen sollten so konstruiert sein, daß sie von allen Befragten zu beantworten sind, damit bei einer Befragungsperson nicht von vornherein der Eindruck entsteht, sie sei für den Interviewer und die Befragung uninteressant.

Die Logik des Befragungsablaufes sollte für die Befragungsperson gut nachvollziehbar sein; Fragen zum gleichen Thema sollten zu Fragenblocks zusammengefaßt werden. Schwierige oder heikle Fragen sollten eher am Ende der Befragung gestellt werden, damit der Schaden im Falle eines durch die Frage bedingten Interview-Abbruches begrenzt bleibt.

Besondere Aufmerksamkeit ist in der methodischen Forschung der Frage gewidmet worden, welchen Einfluß die Anordnung von Fragen im Fragebogen auf das Antwortverhalten habe. Daß vorausgehende Fragen die Antworten auf nachfolgende beeinflussen können, gehört zu den Gemeinplätzen der Fragebogenkonstruktion. Solche „Sukzessions-“, „Reihenfolge-“ oder „Plazierungseffekte“ - allgemein als „Kontexteffekte“ bezeichnet (vgl. z. B. Schuman & Presser 1981; Bradburn 1983; Hippler & Schwarz 1987; Schwarz & Sudman 1992) - stellen auch für die kognitionspsychologische Forschung einen Schwerpunkt dar. Wir erinnern uns an das Beispiel mit Weizsäcker zu Beginn. Ich will erklären, was passiert ist (s. Seite 32).

Folie Fb_11: Bewertung der CDU die Zweite

Alles in allem: Was halten Sie ganz allgemein von der CDU?

11	= sehr viel		
10			
9		erster Mittelwert:	3,4 (A)
8			
7			
6		zweiter Mittelwert:	5,2 (B)
5			
4			
3		dritter Mittelwert:	6,5 (C)
2			
1	= überhaupt nichts		

- (A) Wissen Sie zufällig, welches Amt Richard von Weizsäcker ausübt, das ihn außerhalb des Parteigeschehens stellt?
- (B) Ohne Vorfrage im gleichen Kontext
- (C) Vorfrage: Wissen Sie zufällig, welcher Partei Richard von Weizsäcker seit mehr als 20 Jahren angehört?

Norbert Schwarz und Herbert Bless, Assimilation and Contrast Effects in Attitude Measurement. *Advances in Consumer Research* 19, 1992, S. 72ff

Grundsätzlich variiert das Ergebnis der CDU-Bewertung in Abhängigkeit von der vorangehenden Frage, die auf der Folie jeweils nachzulesen ist. In Variante A mit dem Mittelwert von 3,4 wurde von Weizsäcker in der Vorfrage explizit („außerhalb des Parteiengeschehens“) aus der CDU herausdefiniert; da von Weizsäcker zweifelsohne in allen politischen Lagern als äußerst integre Persönlichkeit mit hoher Reputation angesehen wird, geht mit seiner Exklusion die Bewertung der CDU nach unten. In Variante C mit dem Mittelwert von 6,5 passiert genau das Gegenteil: durch die explizite Inklusion von Weizsäcker's in die CDU geht dessen Reputation und Integrität in die

Bewertung der Partei mit ein und wirkt sich deutlich positiv aus. Folgerichtig liegt der Mittelwert der Variante B ohne einschlägige, kontextgebundene Vorfrage zwischen den Werten für Inklusion und Exklusion.

4.7 Befragungshilfen, Filter, Layout

Kommen wir zum Ende des Abschnittes über Fragebogenerstellung, und schließen wir mit einigen wenigen Überlegungen zu Befragungshilfen, Filtern und Layout:

Befragungshilfen unterstützen die Arbeit des Interviewers, indem sie der Befragungsperson Informationen optisch präsentieren, die ohne entsprechende Präsentation nicht oder nur schlecht zu verarbeiten wären: lange Itemlisten, Skalen, Bilder, Skizzen, etc. Geläufige Formen der Befragungshilfen sind Listen(hefte) und Kärtchen.

Befragungshilfen sind insbesondere im persönlich-mündlichen Interview von Bedeutung. Beim schriftlichen Fragebogen sind sie ohnehin Bestandteil des Fragebogens, beim telefonischen Interview sind sie praktisch nicht einsetzbar bzw. müssen grundsätzlich neu an die Methode angepaßt werden. Der Wegfall visueller Befragungshilfen wird bei der telefonischen Befragung viel eher durch andere Erhebungstechniken wettzumachen gesucht, als durch Befragungshilfen im engeren Sinne: z. B. unfolding tactic (Groves 1979), bei der komplexere Fragen in Stufen abgefragt werden: „Sind Sie eher für oder eher gegen...?“. Falls „eher für“: „Sind Sie voll und ganz für..., einigermaßen oder gerade noch für...?“. Falls „eher dagegen“: „Sind Sie nur wenig dagegen, einigermaßen oder voll und ganz dagegen?“.

Bestimmte Fragen des Fragebogens sind nicht oder nicht sinnvoll von allen Befragten zu beantworten (z. B. von Unverheirateten Fragen nach Merkmalen des Ehepartners). Um zu vermeiden, daß Befragte mit sie nicht betreffenden Fragen konfrontiert werden, werden im Fragebogen an den entsprechenden Stellen sogenannte "Filter" eingebaut. Durch einen Filter erhält der Interviewer (beim Selbstausfüller: die Befragungsperson) die Anweisung, die folgenden Fragen zu überspringen und die Befragung an einer späteren, durch den Filter exakt definierten Stelle des Fragebogens weiterzuführen.

Unter Layout eines Fragebogens schließlich versteht man alle Maßnahmen, die seine formale und äußere Gestaltung betreffen. Generell ist dazu zu sagen, daß die Gestaltung eines Fragebogens so sein muß, daß sie dem Interviewer die Arbeit möglichst leicht und ihm keine formalen Schwierigkeiten macht. Als hilfreich und wichtig gelten ein einheitliches Präsentationsbild für Frageformulierungen, für Antwortkategorien, für Interviewerhinweise etc., wobei zumindest bei persönlich-mündlichen und telefonischen Interviews weniger die ästhetische Qualität als die leichte Handhabbarkeit des Fragebogens durch den Interviewer (Institutsroutinen) im Vordergrund steht.

Bei schriftlichen Befragungen kommt dem Layout über die leichte Handhabbarkeit des Instruments durch die ausfüllende Person hinausgehende Funktion zu: es soll helfen, Kooperationsbereitschaft herzustellen und den Eindruck von Wichtigkeit und Seriosität der Befragung ebenso vermitteln wie die leichte Lösbarkeit der Aufgabe; deshalb muß ein schriftlicher Fragebogen attraktiv gestaltet, übersichtlich gedruckt und gut lesbar

sein, eingängig und leicht zu bearbeiten. Attraktive Deckblätter, die einen Bezug zum Thema haben, haben sich in der Praxis ebenso bewährt wie ausformulierte Hinweise zum Ausfüllen des Fragebogens, eine Dankesfloskel und freier Platz am Ende des Fragebogens, den die ausfüllende Person nutzen kann, um eigene Anmerkungen zu dem Fragebogen zu machen.

Einen allgemein anerkannten Ansatz zur Durchführung einer postalischen Befragung, der neben einer Anzahl zentraler Aspekte ihrer Abwicklung auch eine Vielzahl von Hinweisen zur Layout-Gestaltung eines schriftlichen Fragebogens liefert, stellt die grundlegende Arbeit von Dillman (1978) zur Total Design Method dar, in der selbst Anregungen für die Größe und Farbe des zu verwendenden Papiers nicht fehlen.

5. Pretesting

Die Durchführung eines oder auch mehrerer Pretests gilt gemeinhin als unabdingbare Voraussetzung für die erfolgreiche Entwicklung eines Fragebogens. Wir wollen uns zunächst mit Stellenwert und Ziel eines Pretests beschäftigen und dann unterschiedliche Pretestverfahren vorstellen (vgl. Mohler & Porst 1996).

5.1 Stellenwert und Ziel eines Pretests

Wo immer man auch hinschaut und nachliest: In der gesamten einschlägigen Literatur werden Pretests in allen Projekten der empirischen Sozialforschung als unabdingbare Voraussetzung zur Vorbereitung einer Hauptstudie bezeichnet (Porst 1985:50; Prüfer & Rexroth 1996), zumindest in all jenen Projekten, die zur Datengewinnung Fragebogen einsetzen wollen.

An dieser Stelle immer wieder zitiert: Sudman und Bradburn (1983):

„If you don't have the resources to pilot test your questionnaire, don't do the study.“

Dieses Zitat, so eindringlich es wirkt, hat allerdings eine ganz zentrale Schwäche, nämlich: es setzt Pretesting sehr einschränkend gleich mit Überprüfung eines Fragebogens.

Eigentlich sollte es aber anders sein, nämlich: Pretests sollten die Aufgabe haben, Hinweise zu liefern über die Funktionsfähigkeit des gesamten Studiendesigns sowie einzelner Bestandteile dieses Designs, und dazu gehört, neben Stichprobenziehung und Stichprobenrealisierung, neben Fragen des Feldes und sogar der Auswertung, eben auch das Befragungsinstrument, der Fragebogen.

Praktisch ist dies aber nicht der Fall: Meistens werden alle anderen Aspekte vernachlässigt, und allein der Fragebogen steht im Mittelpunkt des Pretestinteresses.

Das heißt also: Praktisch handelt es sich beim Pretest um einen Testlauf eines „Fragebogen-Prototyps“, d.h. eines mutmaßlich noch nicht vollständig ausgereiften

Fragebogens. Damit wird auch deutlich, in welchen übergeordneten Zusammenhang Pretests einzuordnen sind: Sie sind ein wesentliches Element im Prozeß der Fragebogen-Entwicklung. Warum wird gerade das Pretesten des Fragebogens immer wieder in den Vordergrund gestellt? Zum einen sicherlich deshalb, weil der Fragebogen bekanntermaßen eine der zentralen potentiellen Fehlerquellen in Umfragen ist. Und zum zweiten: weil wir alle wissen, daß ein perfekter Fragebogen - wenn es denn einen perfekten Fragebogen überhaupt gibt - nicht am Schreibtisch konstruiert werden kann. Wir stimmen Sudman und Bradburn (1983) erneut zu: „*Even after years of experience, no expert can write a perfect questionnaire.*“

Wir stimmen auch Klaus Allerbeck und Wendy Hoag (1985) zu, die im Zusammenhang mit der Konstruktion der Fragebogen ihrer Jugendstudie sagen: „*Erfahrene Umfrageforscher kennen die hier gültige Variante des Murphysischen Gesetzes, daß alles, was schiefgehen kann, auch schiefgehen wird, schon lange*“.

Nun, da wir alle wissen, wie wichtig ein Pretest ist, suchen wir nach einer Definition und - oh Wunder - wir finden keine. Trotz des hohen Stellenwerts, der Pretests gemeinhin zugeschrieben wird, gibt es zumindest in der uns bekannten Literatur weder eine allgemein anerkannte Definition noch konkrete und präzise Regeln zur praktischen Durchführung eines Pretests (eine Feststellung, die sich im übrigen mühelos auf andere zentrale Begriffe der empirischen Umfrageforschung übertragen ließe). Trauen wir uns, eine eigene Definition vorzustellen, so sagen wir:

Pretests sind nichts anderes als die Miniaturausgabe einer beliebigen Form sozialwissenschaftlicher Datenerhebung, wobei sich in der Regel die Konzentration auf die Qualität des Erhebungsinstrumentes richtet. Das heißt:

Folie Pret_1: Aufgaben eines Pretests

**Im allgemeinen sollte ein Pretest Auskunft geben über...
(in Anlehnung an Converse und Presser 1986):**

- **Verständlichkeit der Fragen**
- **Probleme des Befragten mit seiner Aufgabe**
- **Interesse und Aufmerksamkeit des Befragten bei einzelnen Fragen**
- **Interesse und Aufmerksamkeit des Befragten während des gesamten Interviews**
- **Wohlbefinden des Befragten (respondent well-being)**
- **Häufigkeitsverteilung der Antworten**
- **Reihenfolge der Fragen**
- **Kontexteffekte**
- **Probleme des Interviewers**
- **Technische Probleme mit Fragebogen/Befragungshilfen**
- **Zeitdauer der Befragung**

Spätestens nach dieser Aufzählung merkt man, daß es ab jetzt vorrangig um Pretests für mehr oder weniger standardisierte Fragebogen geht.

5.2 Pretestverfahren

Wenden wir uns der Frage zu, welche unterschiedlichen Pretestverfahren zur Verfügung stehen. Auf der einen Seite haben wir den sog. „klassischen“ Pretest, der sich vor allem dadurch kennzeichnet - ich verkürze das - daß er im Feld durchgeführt wird und der Interviewer die eher passive Funktion eines „Beobachters“ innehat. Auf der anderen Seite haben wir Pretestverfahren, die eher und vorrangig in Labors oder inhouse durchgeführt werden, bei denen sowohl Befragte wie auch Interviewer aktive Rollen beim Pretesten übernehmen (wir fassen sie hier unter kognitive und Labor-Methoden zusammen). Daneben gibt es Verfahren, die theoretisch oder praktisch ganz ohne Mitwirkung einer Zielperson ablaufen (z. B. Expertenratings, Fragebogenkonferenzen).

Irgendwo zwischen dem klassischen Pretest und dem kognitiven Laborpretest ist schließlich das Behavior Coding angesiedelt, das bei ZUMA zum Einsatz kommt.

Wenden wir uns zunächst dem klassischen Pretest zu, der auch bezeichnet wird als

- konventioneller Pretest („conventional pretest“: Presser & Blair 1994)
- Standard-Pretest („standard pretest“: Oksenberg, Cannell & Kalton 1991)
- Beobachtungspretest (ZUMA), oder, schon eher etwas respektlos als
- „old-style-pretest“ (Fowler 1992)

Die Bezeichnungen „klassisch“, „konventionell“ und „old-style“ lassen vermuten, daß diese Attribute dem traditionellen Beobachtungspretest erst in jüngster Zeit - genauer seit der Entwicklung alternativer Verfahren - verliehen worden sind.

Auch für die Durchführung eines „klassischen Pretests“ existieren keine verbindlichen bzw. allgemein akzeptierten Regeln. Zwar findet sich in jedem sozialwissenschaftlichen Methodenlehrbuch ein Kapitel zum Thema „Pretest“, im Vergleich miteinander werden jedoch höchst unterschiedliche und zum Teil auch widersprüchliche Empfehlungen bezüglich der Durchführung eines Pretests gegeben. Blättert man die Methodenliteratur durch, so stößt man auf Uneinigkeit zum Beispiel bei

- **der notwendigen Fallzahl** (variiert von $N = 10$ bis 200),
- **dem Einsatz der Interviewer** (erfahrene, speziell ausgebildete Pretestinterviewer oder Querschnitt aller Interviewer),
- **der Informiertheit der Befragten** („participating pretest = Befragter ist über Testcharakter informiert im Gegensatz dazu „undeclared pretest = Befragter nicht informiert, reagiert unter realistischen Bedingungen), ja sogar bei der
- **der Terminologie** an sich (Pretest, Vortest, Testbefragung, pilot study, question testing, trial run etc.)

Obwohl es nirgendwo explizit hingeschrieben steht, gibt es doch zumindest aber eine Art Übereinstimmung darüber, wie das „Grundgerüst“ eines „klassischen“ Pretests beschaffen ist bzw. sein sollte:

Folie Pret_2: Der „klassische“ Pretest

Der „klassische“ Pretest

- **Einmaliges Testen des Fragebogens unter möglichst realistischen Hauptstudie-Bedingungen**
- **Durchführung von 20 bis 70 Interviews (Quota oder Random)**
- **Interviewer haben die Aufgabe, Probleme und Auffälligkeiten bei der Durchführung der Interviews zu beobachten und zu berichten.**
- **In der Regel passives Verfahren, d. h. der Interviewer beobachtet nur („Beobachtungspretest“), ohne aktiv zu hinterfragen.**
- **Zugrundeliegendes Prinzip: Man versucht, aus der Reaktion/Antwort des Befragten Rückschlüsse auf sein Fragenverständnis zu ziehen.**

Im Laufe der 80er Jahre wurde von Kognitionspsychologen in den USA eine Reihe von Techniken zur Untersuchung des Frage-Antwort-Prozesses entweder neu entwickelt, oder bereits vorhandene Techniken wurden zu neuem Leben erweckt (vgl. Akkerboom 1996; Wänke 1996). Dabei ging es zunächst einmal nicht um Pretests. Die Eignung dieser Techniken zu Pretest-Zwecken wurde im Grunde genommen erst später erkannt.

Zu den kognitiven Techniken im einzelnen zählen wir (s. Seite 38):

Kognitive und Labor-Methoden

Think-aloud-Technik mit zwei Varianten:

- Concurrent Think Aloud
- Retrospective Think Aloud

Probing-Technik mit den wichtigsten Varianten:

- Follow-Up-Probing
- Post-Interview-Probing
- Comprehension Probing
- Information Retrieval Probing

Memory Cues

Confidence Ratings

Paraphrasing

Sorting

- Free Sort
- Dimensional Sort

Think-aloud ist eine Neuentwicklung, bei der die Befragungsperson aufgefordert wird, „laut zu denken“. Die Think-aloud-Technik entstammt der stark kognitionspsychologisch orientierten Gedächtnisforschung und ist vor allem bei retrospektiven Fakt-Fragen sinnvoll einsetzbar.

Es gibt zwei wichtige Varianten:

- Bei der Concurrent Think Aloud denkt die Befragungsperson laut, während sie ihre Antwort formuliert.
- Bei der Retrospective Think Aloud denkt die Befragungsperson nach der Beantwortung der Frage laut darüber nach, wie ihre Antwort zustande gekommen ist.

Eine eher alte Technik ist das **Probing**: Darunter verstehen wir das explizite Nachfragen, um mehr über die Antwortstrategien der Befragungsperson zu erfahren.

Die wichtigsten Varianten des Probing sind:

- Follow-Up-Probing: Nachfrage sofort nach der Antwort
- Post-Interview-Probing: Nachfragen im Anschluß an das Interview
- Comprehension Probing: Nachfrage zum Fragenverständnis
- Information Retrieval Probing: Nachfrage zur Informationsbeschaffung

Vier weitere Neuentwicklungen im Bereich kognitiver Labormethoden zur Durchführung von Pretests sind

- **Memory Cues**: Erinnerungshilfen für die Befragungsperson
- **Confidence Ratings**: Befragungsperson bewertet den Grad der Verlässlichkeit ihrer Antwort.
- **Paraphrasing**: Befragungsperson soll die Frage mit eigenen Worten formulieren.
- **Free/Dimensional Sort**: Befragungsperson gruppiert vorgegebene Items entweder nach eigenen (Free Sort) oder vorgehenden Kriterien (Dimensional Sort) anhand vorgegebener Skalen.

Schließlich haben wir hier noch das **Focus Interview**, das im Grunde eine alte Technik ist, nämlich die relativ unstrukturierte Diskussion über das Befragungsinstrument mit Gruppen von Befragungspersonen oder mit einzelnen Befragten. Die focus group ist sozusagen das Äquivalent zur Gruppendiskussion mit dem Ziel, in der Diskussion nicht Informationen über Inhalte, sondern über das Befragungsinstrument selbst zu ermitteln.

Kennzeichnend für alle diese Techniken ist, daß sie einerseits Test-Charakter besitzen, andererseits aber auch gleichzeitig Instrumente zur Fragebogenentwicklung darstellen. Weiterhin ist festzuhalten, daß sich diese Techniken vorwiegend auf den Test einzelner Fragen beschränken und nicht auf den Fragebogen als Ganzes.

Dies bedeutet, daß wir Labor-Techniken zwar zur Vorbereitung einer Befragung empfehlen und einsetzen, daß sie aber unserer Ansicht nach nicht den Test des gesamten Instruments - in welcher Form auch immer - ersetzen können.

Zumindest einige dieser „Labor“-Methoden sind auch im normalen Interview einsetzbar, z.B. wurde die Think-Aloud-Methode von ZUMA bereits in Pretest-Interviews im Feld erfolgreich angewandt.

Bei den Verfahren, die im Prinzip ohne Befragungsperson auskommen, ist vor allem das Expertenrating zu nennen: Mehrere Experten stellen zunächst unabhängig voneinander die ihrer Ansicht nach vorhandenen Probleme eines Fragebogens fest und diskutieren anschließend gemeinsam darüber. Eine prominente Form des Expertenratings finden wir in der Fragebogenkonferenz des Instituts für Demoskopie (vgl. Noelle-Neumann 1996).

Bleibt schließlich das Behavior Coding. Grundlegendes Prinzip dieses Verfahrens ist die Klassifizierung von Verhalten.

Behavior Coding wurde ursprünglich eingesetzt, um Interviewerverhalten zu bewerten, z.B. bei Cannell, Lawson & Hausser (1975). In späteren Arbeiten wurde es dann auch auf Befragtenverhalten angewandt (z. B. bei Morton-Williams & Sykes 1984; Cannell, Kalton & Fowler 1985; Prüfer & Rexroth 1985).

Beim traditionellen Behavior Coding bewerten Coder nach dem Interview das Befragtenverhalten mittels eines Codesystems, das mehr oder weniger umfangreich sein kann, d.h. mehr oder weniger differenziert Verhalten erfassen kann. Der quantitative Häufigkeitswert adäquater und inadäquater Befragtenverhaltensweisen bei einer Frage wird als Qualitätsindikator dieser Frage gewertet, weil man davon ausgeht, daß bei inadäquatem Befragtenverhalten die Frage irgendeine Schwierigkeit, irgendein Problem aufweist.

In der ZUMA-Feldabteilung wird das Verfahren seit Jahren modifiziert eingesetzt (vgl. Prüfer & Rexroth 1985). Die Modifikation besteht in einer Verbindung zwischen traditionellem Behaviour Coding und Beobachtungspretest. Wie sieht das aus?

Die Bewertung wird nicht vom Coder nach dem Interview, sondern vom **Interviewer** während des Interviews vorgenommen. Bedingung ist, daß das Codesystem auf das Notwendigste reduziert ist. Das heißt, das Befragtenverhalten wird mittels einer Codeziffer nur im Hinblick darauf bewertet, ob es im Sinne der Fragestellung adäquat oder nicht adäquat ist. **Zusätzlich** beschreibt der Interviewer für jede inadäquate Verhaltensweise in einem Interviewererfahrungsbericht nach dem Interview möglichst detailliert das Befragtenverhalten.

Der Vorteil dieses Vorgehens besteht darin, daß man neben dem quantitativen Häufigkeitswert als Qualitätsindikator auch die konkrete Ursache für den Qualitätsmangel einer Frage kennt. Also: Quantifizierung von Verhaltensweisen plus qualitative Beschreibung des Problems (Prüfer & Rexroth 1985).

In der Literatur wird auf diese Art der Anwendung des Behaviour Codings nicht verwiesen, und unseres Wissens wird es auch außer bei ZUMA in der Praxis so nicht angewendet. Für den Interviewer bedeutet es eine hohe Anforderung, die nur durch entsprechende Schulungsmaßnahmen erfüllt werden kann. Das Verfahren hat sich bei ZUMA als sinnvolle Variante des Pretesting bewährt.⁵

⁵ Informationen zum aktuellen Stand des pretesting bei ZUMA finden sich bei Prüfer & Rexroth (1998)

6. Literatur

- Akkerboom, H. (1996): Labor für die Entwicklung und den Test von Erhebungsinstrumenten. S. 66 - 71 in: Statistisches Bundesamt (Hrsg.), Pretest und Weiterentwicklung von Fragebogen. Band 9 der Schriftenreihe Spektrum Bundesstatistik, Stuttgart: Metzler-Poeschel
- Alemann, H. v. (1984): Der Forschungsprozeß. Eine Einführung in die Praxis der empirischen Sozialforschung. 2. Auflage, Stuttgart: Teubner
- Allerbeck, K. und W. Hoag (1985): Zur Methodik der Umfragen. Frankfurt am Main: Johann Wolfgang von Goethe-Universität
- Atteslander, P. (1993): Methoden der empirischen Sozialforschung. Berlin, New York: De Gruyter
- Bungard, W. und H.E. Lück (1982): Nicht-reaktive Meßverfahren. S. 317-340 in: Patry, J.L. (Hrsg.): Feldforschung. Bern: Huber
- Berdie, D.R., Anderson, J.F. und M.A. Niebuhr (1986): Questionnaire: Design and Use. Metuchen, N.J.: The Scarecrow Press
- Bishop, G.F., Hippler, H.-J., Schwarz, N. und F. Strack (1988): A Comparison of Response Effects in Self-administered and Telephone Surveys. S. 321 - 340 in: Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls W.L. und J. Waksberg (Hrsg.): Telephone Survey Methodology. New York: Wiley
- Bradburn, N.M. (1983): Response Effects. S. 289 - 328 in: Rossi, P.H. und J.D. Wright (Hrsg.): The Handbook of Survey Research. New York: Academic Press
- Cannell, C., Kalton, G. und F. Fowler (1985): Techniques for Diagnosing Cognitive and Affective Problems in Survey Questions. Ann Arbor: The University of Michigan; Survey Research Center; Institute for Social Research
- Cannell, C., Lawson, S. und D. Hausser (1975): A Technique for Evaluating Interviewer Performance. Ann Arbor: The University of Michigan; Survey Research Center; Institute for Social Research
- Clark, H.H. und M. Schober (1992): Asking Questions and Influencing Answers. S. 15 - 48 in: Tanur, J.M. (Hrsg.): Questions about Questions. New York: Russell Sage
- Converse, J.M. (1984): Strong Arguments and Weak Evidence: The Open/Closed Questioning Controversy of the 1940s. *Public Opinion Quarterly*, 48, S. 267-282
- Converse, J.M. und S. Presser (1986): Survey Questions. Handcrafting the Standardized Questionnaire. Beverly Hills: Sage
- Cox, E.P. (1980): The Optimal Number of Response Alternatives for a Scale: A Review. *Journal of Marketing Research*, 17, S. 407-422
- Davis, J.A. (1980): General Social Surveys, 1972-1980: Cumulative Codebook. Chicago: National Opinion Research Center
- De Leeuw, E. und J. van der Zouwen (1988): Data Quality in Telephone and Face to Face Surveys. A Comparative Meta-Analysis. S. 283-300 in: Groves, R. u.a. (Hrsg.): Telephone Survey Methodology. New York: Wiley
- Diekmann, A. (1995): Empirische Sozialforschung. Reinbek: Rowohlt
- Dijkstra, W. und J. van der Zouwen (Hrsg.) (1982): Response Behavior in the Survey-Interview. London, New York, u.a.: Academic Press
- Dillman, D.A. (1978): Mail and Telephone Surveys. The Total Design Method. New York: Wiley
- Erbslöh, E. (1972): Das Interview. Stuttgart: Teubner

- Friedrichs, J. (1990): Methoden empirischer Sozialforschung. 14. Auflage. Opladen: Westdeutscher Verlag
- Fowler, F.J. (1992): How unclear Terms Affect Survey Data. *Public Opinion Quarterly*, 56, S. 218-231
- Fowler, F.J. und T.W. Mangione (1990): Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Newbury Park: Sage
- Frey, J.H. (1989): Survey Research by Telephone. 2. Auflage. Beverly Hills: Sage
- Frey, J.H., G. Kunz und G. Lüschen (1990): Telefonumfragen in der Sozialforschung. Opladen: Westdeutscher Verlag
- Grice, H.P. (1975): Logic and Conversation. S. 41-58 in: Cole, P. und J. L. Morgan (Hrsg.), *Syntax and Semantics 3: Speech Acts*. New York: Academic Press
- Groves, R.M. (1979): Actors and Questions in Telephone and Personal Interview Surveys. *Public Opinion Quarterly*, 43, S. 190-205
- Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L. und J. Waksberg (Hrsg.) (1988): *Telephone Survey Methodology*. New York: Wiley
- Hafermalz, O. (1976): Schriftliche Befragung - Möglichkeiten und Grenzen. Wiesbaden: Gabler
- Hippler, H.-J. (1988): Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven. *Planung und Analyse*, 6/88, S. 244-248
- Hippler, H.-J. und N. Schwarz (1987): Response Effects in Surveys. S. 102 - 112 in: Hippler, H.-J. Schwarz, N. und S. Sudman (Hrsg.): *Social Information Processing and Survey Methodology*. New York: Springer Verlag
- Hippler, H.-J. und N. Schwarz (1990): Die Telefonbefragung im Vergleich mit anderen Befragungsarten. S. 437-477 in: Forschungsgruppe Telefonkommunikation (Hrsg.): *Telefon und Gesellschaft*, Bd. 2, Berlin: Spiess
- Hippler, H.-J., Schwarz, N. und S. Sudman (Hrsg.) (1987): *Social Information Processing and Survey Methodology*. New York: Springer Verlag
- Holkamp, K. (1981): *Theorie und Experiment in der Psychologie*. Berlin, New York: De Gruyter
- Hopf, C. (1978): Die Pseudo-Exploration - Überlegungen zur Technik qualitativer Interviews in der Sozialforschung. *Kölner Zeitschrift für Soziologie und Sozialforschung* 7, Heft 2, S. 97ff
- Jabine, Th.B., Straf, M.L., Tanur, J.M. und R. Tourangeau (Hrsg.) (1984): *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington D.C.: National Academy Press
- Jobe, J.B. und E.F. Loftus (Hrsg.) (1991): Cognition and Survey Measurement. Special Issue of *Applied Cognitive Psychology*, 5, No. 3
- Kaase, M. Ott, W. und E.K. Scheuch (Hrsg.) (1983): *Empirische Sozialforschung in der modernen Gesellschaft*. Frankfurt: Campus
- Karakalos, B. (1979): Das narrative Interview als Instrument der Konstitution sozialwissenschaftlicher Daten. S. 227-242 in: Soeffner, H.-G. (Hrsg.), *Interpretative Verfahren in den Sozial- und Textwissenschaften*. Stuttgart: Metzler
- Karmasin, F. und H. Karmasin (1978): *Einführung in Methoden und Probleme der Umfrageforschung*. Wien, Köln, Graz: Hermann Böhlhaus Nachf.
- Krech, D. und R.S. Crutchfield (1948): *Theory and Problems of Social Psychology*. New York: McCraw Hill
- Lamnek, S. (1989): *Qualitative Sozialforschung*. München: Psychologie Verlags Union

- Lavrakas, P.J. (1987): Telephone Survey Methods. Beverly Hills: Sage
- Lazarsfeld, P.F. (1944): The Controversy over Detailed Interviews - an Offer for Negotiation. Public Opinion Quarterly, 8, S. 38-60
- Lucas, W.A. und W.C. Adams (1977): An Assessment to Telephone Survey Methods. Santa Monica
- Mayntz, R., Holm, K. und P. Hübner (1971): Einführung in die Methoden der empirischen Soziologie. Opladen: Westdeutscher Verlag
- Milgram, S., Mann, L. und S. Harter (1965): The Lost-Letter Technique: A Tool of Social Research. Public Opinion Quarterly 29, S. 437-438
- Mohler, P.Ph. und R. Porst (1996): Pretest und Weiterentwicklung von Fragebogen - Einführung in das Thema. S. 7 - 15 in: Statistisches Bundesamt (Hrsg.), Pretest und Weiterentwicklung von Fragebogen. Band 9 der Schriftenreihe Spektrum Bundesstatistik, Stuttgart: Metzler-Poeschel
- Molenaar, N.J. (1982): Response Effects of Formal Characteristics of Questions. In: Dijkstra, W. und J. van der Zouwen (Hrsg.): Response Behavior in the Survey Interview. New York: Academic Press
- Morton-Williams, J. und W. Sykes (1984): The Use of Interaction Coding and Follow-up Interviews to investigate Comprehension of Survey Questions. Journal of the Market Research Society, 26, S. 109-127
- Noelle, E. (1963): Umfragen in der Massengesellschaft. Hamburg: Rowohlt
- Noelle-Neumann, E. (1996): Die Fragebogenkonferenz. S. 55 - 65 in: Statistisches Bundesamt (Hrsg.), Pretest und Weiterentwicklung von Fragebogen. Band 9 der Schriftenreihe Spektrum Bundesstatistik, Stuttgart: Metzler-Poeschel
- Oksenberg, L., Cannell, Ch. und G. Kalton (1991): New Strategies for Pretesting Survey Questions. Journal of Official Statistics, 7, S. 349-365
- Payne, S. (1951): The Art of Asking Questions. Princeton: University Press
- Phillips, D.L. (1971): Knowledge from What? Chicago: Rand McNally
- Porst, R. (1985): Praxis der Umfrageforschung. Stuttgart
- Porst, R. (1996): Fragebogenerstellung. S. 737-744 in: Goebel, H., Nelde, P. H., Starý, Z. und W. Wölck (Hrsg): Kontaktlinguistik - Contact Linguistics - Linguistique de contact. Ein internationales Handbuch zeitgenössischer Forschung. Berlin und New York: De Gruyter
- Presser, S. und J. Blair (1994): Survey Pretesting: Do Different Methods Produce Different Results? Sociological Methodology, S. 73-104
- Prüfer, P. und M. Rexroth (1985): Zur Anwendung der Interaction-Coding-Technik. in: ZUMA-Nachrichten, 17, S. 2-49
- Prüfer, P. und M. Rexroth (1996): Verfahren zur Evaluation von Survey-Fragen. ZUMA-Arbeitsbericht Nr. 96/05, Mannheim: ZUMA
- Prüfer, P. und M. Rexroth (1998): Multi Method Pretesting. Ein neues Pretestkonzept (Arbeitstitel). Erscheint als ZUMA-Arbeitsbericht
- Scheuch, E.K (1967): Das Interview in der empirischen Sozialforschung. S. 66-190 in: König, R. (Hrsg.): Handbuch der empirischen Sozialforschung. 2. Auflage. Stuttgart: dtv
- Scheuch, E. K. (1973): Das Interview in der Sozialforschung. S. 66 - 190 in: König, R. (Hrsg), Handbuch der empirischen Sozialforschung, Bd. 2, 3. Auflage, Stuttgart: Enke
- Schnell, R., Hill, P.B. und E. Esser (1992): Methoden der empirischen Sozialforschung, 3. Aufl. München: Oldenbourg
- Schütze, F. (1983): Biographieforschung und narratives Interview. Neue Praxis 13, S. 283-293

- Schuman, H. und S. Presser (1977): Question Wording as an Independent Variable in Survey Analysis. *Sociological Methods and Research*, 6, S. 151-176
- Schuman, H. und S. Presser (1981): *Questions and Answers in Attitude Surveys*. New York: Academic Press
- Schwarz, N. (1990): Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction. S. 98 - 119 in: Hendrick, C. und M.S. Clark, Hrsg: *Research Methods in Personality and Social Psychology*. Beverly Hills, CA: Sage
- Schwarz, N. (1991): In welcher Reihenfolge fragen? Kontexteffekte in standardisierten Befragungen. ZUMA-Arbeitsbericht Nr. 91/16, Mannheim: ZUMA
- Schwarz, N. und H.-J. Hippler (1991): Response Alternatives: The Impact of their Choice and Ordering. In: Biemer, P.P., Groves, R.M., Mathiowetz, N.A. und S. Sudman (Hrsg.): *Measurement Error in Surveys*. Chichester: Wiley
- Schwarz, N., Hippler, H.-J., Deutsch, B. und F. Strack (1985): Response Categories: Effects on Behavioral Reports and Comparative Judgements. *Public Opinion Quarterly*, 49, S. 388-395
- Schwarz, N. und F. Strack (1988): The Survey Interview and the Logic of Conversation: Implications for Questionnaire Construction. ZUMA-Arbeitsbericht Nr. 88/03. Mannheim: ZUMA
- Schwarz, N. und S. Sudman (Hrsg.) (1992): *Context Effects in Social and Psychological Research*. New York: Springer Verlag
- Schwarz, N. und S. Sudman (Hrsg.) (1996): *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass Publishers
- Spöhring, W. (1989): *Qualitative Sozialforschung*. Stuttgart: Teubner
- Stadtler, K. (1983): *Die Skalierung in der empirischen Forschung*. München: Infratest Forschung
- Strack, F. und L.L. Martin (1987): Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys. S. 123 - 148 in: Hippler, H.-J., Schwarz, N. und S. Sudman (Hrsg.): *Social Information Processing and Survey Methodology*. New York: Springer Verlag
- Strack, F. und N. Schwarz (1992): Communicative Influences in Standardized Question Situations: The Case of Implicit Collaboration. In: Fiedler, K. und G.R. Semin (Hrsg.): *Language, Interaction, and Social Cognition*. Beverly Hills: Sage
- Sudman, S. und N.M. Bradburn (1983): *Asking Questions*. San Francisco: Jossey-Bass
- Tourangeau, R. (1987): Attitude Measurement: A Cognitive Perspective. S. in: Hippler, H.-J., Schwarz, N. und S. Sudman (Hrsg.): *Social Information Processing and Survey Methodology*. New York: Springer Verlag
- Webb, E.J., Campbell, D.T., Schwarz, R.D. und L. Sechrest (1966): *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago
- Wänke, M. (1996): Kognitionspsychologische Ansätze zum Pretest von Fragebogen. S. 46 - 54 in: Statistisches Bundesamt (Hrsg.), *Pretest und Weiterentwicklung von Fragebogen*. Band 9 der Schriftenreihe Spektrum Bundesstatistik, Stuttgart: Metzler-Poeschel
- Wiedemann, P.M. (1986): *Erzählte Wirklichkeit. Zur Theorie und Auswertung narrativer Interviews*. Weinheim, München: Psychologie Verlags Union
- Zimmermann, E. (1972): *Das Experiment in den Sozialwissenschaften*. Stuttgart: B.G. Teubner

