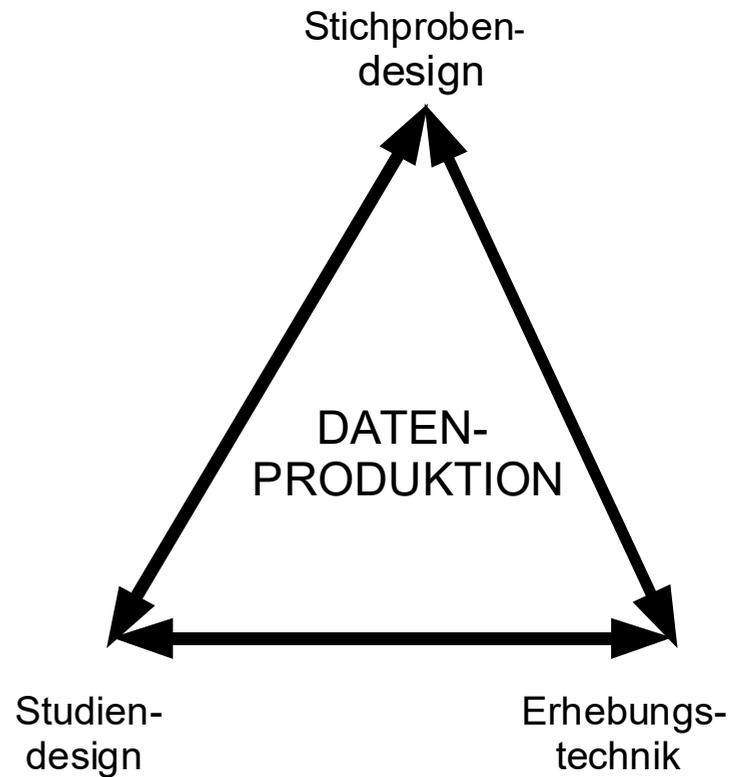


C2 Ausgewählte Aspekte von Studiendesigns und Erhebungstechniken

C2.1 Vorbemerkungen

Wesentliche Designentscheidungen

Uneinheitliche Bezeichnungen



C2.2 Studiendesigns

Im Folgenden bezeichnet:

n : Stichprobenumfang

E : Erhebung der zu untersuchenden Merkmale

I_i : Untersuchungseinheit i , mit $i \in \{1, \dots, n\}$, $n \in \mathbb{N}$

t_j : Erhebungszeitpunkt j , mit $j \in \{1, \dots, T\}$, $T \in \mathbb{N}$

Erhebungszeitpunkt	t_1	t_2	\dots	t_T
Untersuchungseinheit				

Querschnittsstudie

Zeitreihenstudie

Panel-Studie

Trendstudie

Ereignisanalytischer Blick

- Lebensdauern
- Konkurrierende Risiken
- Mehrepisoden-Modelle
- Zensierungsproblematik
- Beobachtung: direkt; kalendarisch; episodensbasiert

Experimentelle Designs

Im Folgenden bezeichnet:

R : Zuordnung der Untersuchungseinheiten zu den Versuchsgruppen per *Randomisierung* (d.h. zufällig)

X_k : Einsatz des k -ten experimentellen *Stimulus* bzw. der Behandlung k , mit $k \in \{1, \dots, K\}$, $K \in \mathbb{N}$ („*Treatments*“), schreibe X , wenn nur ein Stimulus

t_0	t_1	t_2

Klassisches Experiment ($K = 1$)

Experiment mit Vorher-Nachher-Messung ($K = 1$)

Experiment mit verschiedenen Behandlungen

Quasi-Experiment

C2.3 Erhebungstechniken/-methoden: Ein sehr kurzer Überblick

Beobachtung

- [Beobachtung weitgefasst: jede Form der systematischen Registrierung von Daten]
- hier: Beobachtung i.e.S.: unmittelbares, systematisches Erfassen der relevanten Sachverhalte
 - * Feld versus Labor
 - * teilnehmend versus nicht-teilnehmend
 - * offen versus verdeckt

Befragung (Überblick, siehe auch später)

Datenerhebung anhand eines Fragebogens

- * klassisch face-to-face: persönlich vs. telefonisch vs. schriftlich (per Post)
- * Nutzung von moderner Technologie: CAPI - Computer Assisted Personal Interview vs. CATI - Computer Assisted Telefon-Interview vs. Online-Befragung

Inhaltsanalyse

Systematisches Auslesen von Daten aus Texten und anderen Quellen

Prozessproduzierte Daten

Aufbereitung von Informationen, die nicht zu Forschungszwecken aufgezeichnet wurden, immer häufiger

Nicht-reaktive Verfahren i. e. S. („Spurensuche“)

C2.4 Fragebogengestaltung

C2.4.1 Allgemeine Überlegungen

Kernaufgaben eines Fragebogens:

- Datenerfassung
- Weiterverarbeitung der Antworten gewährleisten
- Ziel einer Befragung: korrekte und glaubwürdige Angaben
- Zahlreiche Daten zu einem Thema (sowie persönliche Angaben) werden gebündelt abgefragt
- Fragebogen müssen sehr gut strukturiert und durchdacht aufgebaut werden
- Viele Fallen und Tücken („gut gemeint ist oft das Gegenteil von gut gemacht“)

Überlegungen zu Beginn der Fragebogenerstellung

- Wer?
- Was?
- Wie?
- Womit?

Welcher Befragungstyp soll eingesetzt werden?

- Schriftliche Befragung
- Mündliche Befragung: Einsatz von Erhebungsbeauftragten (Interviewern)

Hilfsmerkmale (in amtlichen Erhebungen)

§12 Abs. 1 Bundesstatistikgesetz:

Hilfsmerkmale sind von den Erhebungsmerkmalen zum frühestmöglichen Zeitpunkt zu trennen und gesondert aufzubewahren.

Für die Gestaltung von Erhebungsunterlagen folgt daraus, dass Hilfsmerkmale bei Papierfragebogen

- entweder getrennt vom Frageteil
- oder auf einem gesonderten Bogen (z. B. Haushaltsmantelbogen)

anzuordnen sind.

C2.4.2 Entwicklung und Gestaltung eines Fragebogen

- Die Titelseite
- Hinweise zum Ausfüllen
- Hinweise für die Befragten / Interviewer
- Arten von Fragen
- Skalen
- Gebote der Fragenformulierung
- Filter
- Layout

Die Titelseite

- Beispiel für „reine“ Titelseite: Mikrozensus; Selbstausfüllerbogen

- Beispiel für „multifunktionale Titelseite:“ Zensus
 - Haushaltebefragung
 - Zweck der Erhebung
 - Zugangsdaten für Online-Ausfüllen (CAWI)
 - Hinweise zum Ausfüllen
 - Beginn Fragenteil

Hinweise zum Ausfüllen des Fragebogens I

- Hinweise für
 - die/den Befragte(n)
 - die/den Interviewer(in)
- Hinweise sind Erläuterungen / Definitionen, die für die möglichst exakte Beantwortung komplexerer Fragen erforderlich sind.
- Immer am Anfang des Fragebogens alle im Fragebogen vorkommenden Antwortmöglichkeiten (Kästchen (Einfach- versus Mehrfachantwort), numerische Eintragungsfelder, alpha-numerische Eintragungsfelder, . . .) aufführen.

Hinweise zum Ausfüllen des Fragebogens II

- Schriftliche Befragung / Selbstausfüller, drei Möglichkeiten der Anordnung:
 - * Direkt bei der Frage (Gefahr der Überfrachtung des Fragenteils; online durch Anklicken)
 - * Am Ende des Fragebogens (Wird oft nicht gelesen)
 - * Kombination (Kurzerläuterung im Fragenteil, ausführlichere Erläuterung am Ende)

- Persönliches Interview (PAPI)
 - * Anleitungen für den Interviewer (Generelle Schulung, für Details: Erläuterungen entsprechend dem Fragenverlauf plus Glossar)

Arten von Fragen

- Geschlossene Fragen Begrenzte und definierte Anzahl von Antwortkategorien Zwei Arten:
 - Einfachnennungen (nur eine Antwort zulässig):
 - Mehrfachnennungen (mehrere Antworten sind zulässig)
- Halboffene Fragen
- Offene Fragen

Anwendung geschlossener Fragen

- Merkmalsausprägungen sind klar definiert (bei Einfachnennungen disjunkt)
- Kleine Zahl von Merkmalsausprägungen (z.B. nur oberste Ebene bei Klassifikationen)
- Sonderform: offene Frage gegliedert, Befragter klassifiziert sich in einer vorgegebenen Liste
- Vor- und Nachteile:

Anwendung offener Fragen

- Große Zahl von Merkmalsausprägungen (z.B. tief gegliederte Klassifikationen)
- Alle Merkmalsausprägungen nicht vollständig bekannt
- Vor- und Nachteile:

Spezialform: Fragebatterien/Skalen

- Erhebung eines latenten Merkmals/ weniger latenter Merkmal durch Vielzahl von Fragen
- lokale stochastische (= bedingte) Unabhängigkeit, gegeben das Merkmal
- binär oder ordinale Zustimmungsskalen
- Schätzung der Güte der Messung, Crombachs α
- Skalierungsverfahren (z. B. Rasch-Modell)
- Siehe: Vorlesung Wirtschaftsstatistik, EMOS Seminar

Skalenniveaus: Problemfelder

- Ordinal- oder Intervallskala mit interpretierbaren Abständen
 - ? In den Sozialwissenschaften: Entwicklung von Fragen, die die Voraussetzungen einer Intervallskala „erfüllen“. Mit endpunktbenannten Skalen wird der Eindruck erweckt als ob alle Skalenpunkte gleichabständig wären.
 - + ordinale Regression für ordinale abhängige Variable (Tutz!)
 - (?) ordinale Kovariablen
- partiell geordnete Daten (Schollmeyer)

Checkliste zur Formulierung von Fragetexten I

- Anlehnung an ein Gespräch
- Einfache, verständliche und für alle Befragte eindeutige Formulierung!
- Grammatikalisch korrekte und vollständige Formulierung!
- Bei homogenen Gesamtheiten: Sprachebene und Kontext der Befragten berücksichtigen!
- Widerspruchsfreiheit und eindeutige Beantwortbarkeit
- Kurzer Fragentext!
- Keine doppelte Verneinung!

Checkliste zur Formulierung von Fragetexten II

- Keine Überforderung / Überbelastung der Auskunftgebenden
- Neutralität (üblicherweise keine Suggestivfragen, (außer als spezielle Technik))
- Eindimensionalität
- Im Allgemeinen: keine hypothetischen Fragen
- Eindeutiger Zeit- (und Orts-)bezug

Umgang mit heiklen Fragen I

- Tiefes Eindringen in Privatsphäre der Befragten nur aus inhaltlich absolut unverzichtbarem Grund
- Gefahr von Item-Nonresponse bzw. Abbruch

Umgang mit heiklen Fragen II

- inhaltliche Wichtigkeit der Frage verdeutlichen
- Geheimhaltung verdeutlichen
- Anonymität der Situation sicherstellen
- eventuell Mitläufereffekt, Suggestivfragen nutzen
- eventuell beschönigende/rechtfertigende Formulierung
- eventuell Einbettung in allgemeineren Kontext
- eher gegen Ende des Interviews
- schriftliche/Online-Befragung ?

Umgang mit heiklen Fragen III

- bei stetigen Merkmalen Kategorien bilden
- oder offen mit schrittweiser Verfeinerung nach eventueller Verweigerung der offenen Frage (Einkommensfrage ALLBUS, PASS)
- im Extremfall nur auf Verteilungsaussagen statt individueller Daten abzielen:
 - verwandte Merkmale (z.B. Ausgaben statt Mengen)
 - auch andere explizite Anonymisierung als Vergröberung (z B. Summe von Merkmalen)
 - Zufallsverschlüsselung (Randomized Response Techniken)
- Aber auch gegebenenfalls nicht vorhandene Antwortbereitschaft akzeptieren.

Checkliste zur Formulierung von Antwortvorgaben

- Erschöpfend
- überschneidungsfrei (disjunkt) können.
- Neutral
- Verzerrungsausgewogen / ausbalanciert

Filter

Allgemein: Fragen, die nicht oder nicht sinnvoll von allen Befragten beantwortet werden können, werden übersprungen.

- Punktueller Filter: Diejenigen werden gefiltert, die auf eine bestimmte Frage eine bestimmte Antwort gegeben haben.
- Globaler Filter: Alle Personen werden gemeinsam weitergefiltert, die eine bestimmte Frage beantwortet haben. Voraussetzung ist grundsätzlich ein punktueller Filter.

Layout

- Alle Aspekte, die die formale und äußere Gestaltung des Fragebogens betreffen.
- Weniger wichtig bei Befragungen durch Interviewer (PAPI, CATI)
- Sehr wichtig bei Selbstausfüllern. Ziele:
 - * Klare und übersichtliche Gestaltung zur Vermeidung von fehlerhaften Antworten
 - * Positive Motivation für den Befragten zum Ausfüllen
- die häufigsten Fehler:
 - Zu viele Fragen auf einer Seite
 - Zu viel Text auf einer Seite
 - Zu kleine Schrift
 - Zu enge und gedrängte Anordnung der Fragen
 - Unübersichtliche Darstellung

Pretests, allgemeine Ziele

Pretest: Testung und Evaluation eines Fragebogens oder einzelner seiner Teile vor ihrem Einsatz in der Haupterhebung.

- Aufdeckung von Schwierigkeiten bei der Beantwortung (z.B. auf Grund des Frageinhalts oder der Fragereihenfolge)
- Verbesserung der Fragebogen
- Entlastung der Befragten
- Erhöhung der Datenqualität der Erhebung

Ziele von Pretests: Auskunft über

- Die Verständlichkeit der Fragen
- Die passende Reihenfolge der Fragen
- Kontexteffekte
- Das Zustandekommen der Antworten
- Häufigkeitsverteilungen der Antworten
- Probleme der Befragten mit ihrer Aufgabe
- Interesse und Aufmerksamkeit der Befragten bei einzelnen Fragen
- Das Wohlbefinden der Befragten
- Probleme der Interviewer mit dem Fragebogen
- Technische Probleme bei Fragebogen und Befragungshilfen
- Die Zeitdauer der Befragung

Anwendung von Pretests

Pretests sollten auf alle Fälle durchgeführt werden, wenn

- gesetzliche Änderungen eine neue Erhebung verlangen
- neue Fragen, die noch nicht getestet wurden, in einen bestehenden Fragebogen aufgenommen werden
- das Erhebungsinstrument geändert wurde oder zusätzliche Erhebungsmodi eingeführt wurden
- mangelhafte Datenqualität festgestellt wurde

C3 Kurzausblick: Vertiefendes zu den messtheoretischen Herausforderungen

Formalisierung des Operationalisierungsproblems

Das Grundmodell der Klassischen Testtheorie (KTT)

- Grundgleichung

$$\text{Messung} = \text{wahrer Wert} + \text{Messfehler}$$

- Erst nichttrivial durch zusätzliche Annahmen, insbesondere Unabhängigkeit zwischen Messfehler und wahrem Wert
- Erweiterungen auf Wiederholungsmessungen und i.i.d.-Stichproben vieler Personen/Haushalte etc.

Formale Analyse der Messgüte von Frage(batterie)n

- Reliabilität
- Validität
- Cronbach's alpha

Fehler-in-den-Variablen Modelle

Korrekturverfahren für Regressionsmodelle mit fehlerhaft gemessenen Variablen