

Inhaltsverzeichnis

1.Allgemein	1
Forschung.....	1
Ziel der Forschung.....	1
Typisch dabei.....	1
Typen von Forschung.....	1
Explorative Untersuchung:	1
Deskriptive Untersuchung.....	1
Prüfung von Hypothesen und Theorien	1
Evaluationsstudien	1
Geschichte.....	1
Probleme	2
2.Methoden und Wissenschaftstheorie.....	2
Induktion	2
Deduktion.....	2
Falsifikationismus von Popper	2
Probleme.....	3
Wissenschaft und Wahrheit.....	3
3.Theorien.....	3
Grundsätze.....	3
Definition.....	4
Definition im weiteren Sinn:	4
Definition im engeren Sinn:.....	4
Formalisierung.....	4
Definition	4
Nutzen von Formalisierung	4
Forschungsfrage und Theorie	5
Warum Theorien wichtig?	5
4.Hypothesen	5
Definition.....	5
Typen von Hypothesen.....	5
Deterministische	6
Probabilistische	6
Wenn -dann.....	6
Je-desto	6
Konzepte.....	6
Konzeptdefinition:.....	6
Typen von Konzeptdefinition.....	7
Eigenschaften von Konzepten.....	7

Hypothesen und Kausalität	7
Operationalisierung	7
5. Variablen	8
Allgemein	8
Definition	8
Konstruktion von Variablen:.....	8
Typen von Variablen	8
Skalentyp.....	8
Nominal	9
Ordinal	9
Intervall.....	9
Ratio (Verhältnisskala).....	9
Variablen und Messungen	10
Messung	10
Objektivität.....	10
Reliabilität	10
Validität.....	11
Inhaltsvalidität.....	11
Kriteriumsvalidität	11
Konstruktsvalidität	11
Reaktivität von Messungen.....	12
6. Forschungsdesign	12
Allgemein	12
Definition	12
Überlegungen beantworten.....	12
Unbestimmtes Forschungsdesign	12
Untersuchungsebene und einheit	13
Zeitdimension	13
Querschnittsdesign.....	13
Trenddesign	13
Paneldesign.....	13
Informationsgehalt	14
Kohortendesign (Spezialfall von Längsschnitterh.).....	14
Kohorteneffekt.....	14
Lebenszykluseffekt	14
Kohortenfehlschluss	14
Varianz	15
Experiment.....	15
Design von Experimenten.....	15
Validität von Experimenten.....	16
Quasi-experimentelle Designs	17
Ex-post-facto-Design	18

7. Datenerhebung18

Stichprobe.....18

 Definition18

 Stichprobenverfahren 3 Typen.....19

 Wahrscheinlichkeitsauswahl (Zufallsauswahl)19

 bewusste Auswahl19

 willkürliche Auswahl19

 Wieso Zufallsstichproben?.....20

Umfrage.....20

 3 Arten von Fehlerquellen20

 Zufallsfehler der Stichprobe20

 Systematische Fehler aufgrund Verfahrens der Stichprobenauswahl.....20

 Verzerrung („nonsampling bias“)......20

 „nonsampling bias“20

 Messfehler20

 Fehlerquellen im Interview:.....20

 Diskrepanz zwischen Zielpopulation und Surveypopulation (Coverage-Fehler).....20

 Antwortverweigerung: “non-response”21

 Formen der Befragung21

 persönlichen Interviews21

 telephonischen Interviews (preiswert).....21

 schriftlichen Befragungen21

 Was misst man in Befragungen?21

 Erkenntnisse zur Fragenformulierung21

 Einflussfaktoren22

 Typen von Fragen22

 geschlossene Fragen:.....22

 Halboffene Fragen:23

 offene Fragen:.....23

Experteninterview23

 Infos aus Experteninterviews.....23

 Typen von Experteninterviews nach Art der erhobenen Info:23

 faktuelle Informationen.....23

 Einschätzungen:24

 Typen der Befragung24

 Fokussiertes Interview24

 Narratives Interview24

 Problemzentriertes Interview24

Textanalyse.....25

 Allgemein.....25

 Ziele der Analysen nach Semiotik.....25

 Syntaktik25

 Semantik25

 Pragmatik25

 Typen von Textanalysen.....25

 Systematische Textanalysen.....25

 Computerunterstützte Textanalysen:26

Qualitative Textanalysen:.....	26
Ablauf einer Textanalyse:	26
Probleme der Textanalyse.....	27
Vorteile.....	27
andere Daten.....	28
Allgemein.....	28
Prozessproduzierte Daten	28
Beobachtung	28
Teilnehmende - nichtteilnehmende Beobachtung	28
Offene - verdeckte Beobachtung.....	28
Feld-, Laborbeobachtung.....	29
Strukturierte - unstrukturierte Beobachtung.....	29
Fremd-, Selbstbeobachtung.....	29
Unterschied Sozialreportagen und Beobachtungsstudien	29
Nicht-reaktive Methoden.....	29
 8. Was man wählt, hängt von Frage und Gegenstand ab.	 31

9. Allgemein

9.1. Forschung

9.1.1. Ziel der Forschung

- deskriptive Inferenz

Aussagen über nicht beobachtbare Tatsachen (schliessen von Stichprobe auf Charaktereigenschaft der Grundgesamtheit (Population))

- kausale Inferenz

Aussagen über Ursache der Beobachtung (weniger Kriminalität weil Abtreibung legal geworden)

9.1.2. Typisch dabei

- Vorgehensweise öffentlich zugänglich
- Schlussfolgerungen prinzipiell unsicher

9.1.3. Typen von Forschung

- **Explorative Untersuchung:**

- Erforschender Bereich relativ unbekannt
- Ziel Hypothese finden

- **Deskriptive Untersuchung**

- Ist keine kausale, hat daher keine Hypothese von der ausgegangen wird

- **Prüfung von Hypothesen und Theorien**

- **Evaluationsstudien**

- Wirkung von etwas testen + Nebeneffekte, Bsp. Abfallkübel in ZH

9.2. Geschichte

- Sozialwissenschaftliche Daten schon früh erhoben → Volkszählungen für Steuer- und Militärzwecke

- Renaissance → zwei Traditionen

- Politische Arithmetik → mit Registern Erkenntnisse → eher quantitativ
- Unversitätsstatistik: → Beschreibung von Staatsmerkwürdigkeit → eher qualitativ

- 20 Jhd. Institutionalisierung

- Wiener Schule (Paul Lazarsfeld) → Arbeitslosigkeit Marienthal
- Chicago → Feldexperimente
- Columbia

- Behavioral revolution: Durchbruch der Umfragen und Verfeinerung der Umfragetechnik → Präsidentschaftswahl 1936 USA Gallup mit kl. Stichprobe vs. Literary Digest

9.3. Probleme

- Wir suchen immer nach Mustern
- Irreführende Interpretationen → Geschwindigkeitskontrolle 100 zu schnell, davon 60 Männern; sollte noch wissen Anteil Frauen + Männer aller Untersuchten
- Scheinkorrelationen → Regenschirme und Regenwürmer
- Selbstselektion → macht Gewaltsehen selektiv, oder schauen gewaltbereite Menschen mehr Gewaltfernsehen
- Versuchsleitereffekt → Erwartung des Versuchsleiter kann Versuchspersonen beeinflussen
- Werturteile

10. Methoden und Wissenschaftstheorie

10.1. Induktion

Von endlicher Anzahl von Beobachtungen macht man allgemeingültige Schlüsse → Schluss vom Besonderen auf das Allgemeine

Gute Induktion basiert auf

- Vielen Beobachtungen welche
 - Unter verschiedenen Bedingungen gemacht sind
 - Die dem Gesetz/Theorie nicht widersprechen

Problem: aus wahren Prämissen muss nicht zwingend eine wahre Konklusion resultieren

10.2. Deduktion

Von Theorie auf Beobachtung schliessen → Schluss vom Allgemeinen auf das Besondere

- Es entstehen keine neuen Wahrheiten
- Deduktion kann immer logisch korrekt erfolgen, heisst aber noch nicht, dass Schlussfolgerung aus den Prämissen auch wahr ist

10.3. Falsifikationismus von Popper

(20-ern suchte Alternative zum Induktivismus)

eine Induktionslogik im Sinne eines logischen Schlussverfahrens analog zur Deduktionslogik existiert nicht

Falsifikation: Nachweis, dass ein vermuteter Zusammenhang nicht existiert

Theorien machen Aussagen über eine unendliche Menge von Objekten gibt jedoch nur eine endliche Menge von Beobachtungen → Theorien lassen sich grundsätzlich nicht verifizieren

Hingegen: allgemeine Sätze sind falsifizierbar. Steht ein wahrer singulärer Satz im Widerspruch zu einer deterministischen Hypothese, so gilt die Hypothese als falsifiziert

10.3.1. Probleme

- viele neuen Theorien zuerst falsifiziert da
 - Untersuchungsinstrumente nicht präzise
 - Randbedingungen einen starken Einfluss hatten
- Ad-hoc Hypothesen können Theorie schützen → da kein neuer Informationsgehalt welcher der Falsifizierung unterzogen werden könnte

10.4. Wissenschaft und Wahrheit

Paul Feyerabend (1978) stellt in Frage, dass sich die Wissenschaft von anderen Arten von Erkenntnissen unterscheidet → alle Methoden gleichwertig: "Anything goes". Auch widersprach Feyerabend Poppers Idee der Annäherung an die Wahrheit. Konkret sagt Feyerabend

- in der Wissenschaft sei gegen jede methodologische Regel produktiv verstossen worden
- Überlappungsbereich von alten und neuen Theorien seien meist so schmal, dass dies keine fortschreitender Falsifikationismus sei
- wissenschaftlicher Fortschritt gebe es nur durch Lösen von den etablierten Wissenschaftstheorien.

Die Suche nach Wahrheit wird von Seiten der kritischen Psychologie und des radikalen Konstruktivismus als sinnlos eingestuft.

11. Theorien

11.1. Grundsätze

in wissenschaftlichem Forschen von zentraler Bedeutung. Ausgangspunkt ist eine Forschungsfrage. Die Theorie liefert eine mögliche Antwort auf diese Frage. Die Antwort wäre dann unsere Hypothese, die wir überprüfen wollen.

Theorien sollen

- falsifizierbar sein
- logisch konsistent sein (klappt besser mit Formalisierung)
- zu erklärende Variable (auch Konzept genannt) haben
- konkrete Beschreibungen enthalten (klare, spezifische Begriffe)
- möglichst umfassend sein (besser eine Theorie über alle Demokratie als nur über die Schweiz)

11.2. Definition

Theorien sind uneinheitlich definiert. Insbesondere in der Soziologie kann sich hinter einer Theorie alles mögliche verbergen.

11.2.1. Definition im weiteren Sinn:

- eine Menge miteinander verknüpfter Aussagen (mindestens 2), von denen sich
- eine nicht-leere Teilmenge auf empirisch prüfbare Zusammenhänge zwischen Variablen beziehen.

umfasst zu viel → bringt Theorien, welche keine Antworten auf Forschungsfragen liefern. Bsp. a) Morgen ist Donnerstag, am Donnerstag regnet es. b) kann man empirisch überprüfen → regnet es wirklich?

11.2.2. Definition im engeren Sinn:

- Grundannahmen, und zwar zentrale Hypothesen über Zusammenhänge, die empirisch meist schwer prüfbar sind. Definitionen der grundlegenden Begriffe.
- Aus den Grundannahmen abgeleitete Hypothesen sowie Regeln zur Messung der Variablen (Messhypothesen)

Das Donnerstags-Bsp. geht hier nicht mehr, man müsste erklären, warum es am Do. regnet. Sind Grundannahmen aus 1 mathematisch ausgedrückt, nennt man das die **Axiome** einer Theorie. Mathematisch-deduktiv können aus Axiomen Theoreme / Hypothesen abgeleitet werden, welche Modell genannt werden. Diese sollen so einfach wie möglich aber so komplex wie nötig sein. Die Modelle der Naturwissenschaften sind ausgereifter als jene der Sozialwissenschaften.

11.3. Formalisierung

11.3.1. Definition

- Theories are general statements (Grundannahmen) that describe and explain the causes and effects of classes of phenomena.
- Pfaddiagramm als Bsp. für Formalisierung einer Theorie: $A \rightarrow q \rightarrow r \rightarrow B$. bei $A(\text{Donnerstag}) \rightarrow B(\text{Regen})$ fehlen causes and effects.

Bsp.: Demokratien führen nicht Krieg gegeneinander. $A(2 \text{ Demokratien}) \rightarrow B(\text{kein Krieg})$. Es fehlen causes and effects → keine Theorie → Zusammenhang induktiv erzeugt. Theorie soll erklären, warum das so ist etwa im Sinn: Demokratien handeln häufiger miteinander, so dass der Krieg teurer wird als zwischen Ländern, die weniger Handel betreiben.

- A theory is a very general set of propositions from which others, including 'laws,' are derived.

11.3.2. Nutzen von Formalisierung

- Präzisierung von Theorien (z.B. Zusammenhänge)
- Gebrauch der Mathematik und anderer formaler Sprachen für Deduktions- und Ableitungsregeln

- die Sicherung der Konsistenz der Theorie, Regel der Sprache (z.B. Math) sichert die logische Konsistenz, welche bei nur verbaler Formulierung oft nicht gegeben ist.

Nicht nur formalisierte Modelle Theorie, auch verbal formulierte Zusammenhänge. Minimalanforderung: Aussagen über empirisch prüfbare Zusammenhänge zwischen zwei Variablen

Pfaddiagramm → Variablenzusammenhänge darstellen. Die in einem Pfaddiagramm behaupteten Zusammenhänge müssen durch Grundannahmen begründet sein.

11.4. Forschungsfrage und Theorie

11.4.1. Warum Theorien wichtig?

Weil sie mögliche Antworten (Hypothesen) zu unserer Forschungsfrage liefern.

Beispiel:

Forschungsfrage: Wieso werden Medikamente unterschiedlich schnell eingeführt?

Theorie: theoretische Überlegungen zur Verbreitung von Informationen

Forschungsfrage: Wieso führen Demokratien kaum (oder überhaupt nie) Krieg gegeneinander?

Theorie: z.B. theoretische Überlegungen zur Entscheidungsfindung in Demokratien. (Konfliktlösung anders, Handel intensiver (siehe oben), Niederlage kostet die nächste Wahl → Demokratien führen nur Krieg, wenn sie ziemlich sicher sind, auch zu gewinnen. Dafür wird der Krieg dann auch mit den nötigen Mitteln geführt. → Demokratie vs. Demokratie ist darum schlechteste Variante und wird vermieden.

12. Hypothesen

12.1. Definition

Hypothesen sind Vermutungen über einen Sachverhalt (Zusammenhang zw. Konzepten) welche theoretisch hergeleitet werden.

Gute Hypothesen:

- Immer falsifizierbar → was würde mich dazu bringen sie zu falsifizieren
- Richtung der (kausalen) Zusammenhänge darlegen → Ursachen und Wirkung
- Begriffe/Konzepte immer genau definiert werden

12.2. Typen von Hypothesen

- Deterministische und probabilistische → Zusammenhang
- Je-desto und Wenn-dann Hypothesen → Formulierungsmuster

1 + 2 schliessen sich nicht aus, können kombiniert sein

12.2.1. Deterministische

zwischen mindestens zwei Variablen ein deterministischer Zusammenhang behauptet wird → Bsp: Zusammenbruch des Staates und Bauernrevolten sind notwendige Bedingungen für soziale Revolutionen (Skocpol 1979)

12.2.2. Probabilistische

Wahrscheinlichkeitshypothese → Vorhersagewert (bei zwei Variablen ist das die zweite Variable der Hypothese) nur mit einer gewissen Wahrscheinlichkeit auftritt → Bsp: Frühehen weisen ein höheres Risiko der Ehescheidung auf als Spätehen

Wahrscheinlichkeit $p=1$ → deterministische Hypothese

12.2.3. Wenn -dann

Hypothese aus dichotomen Variablen, dann Zusammenhang als 'Wenn-dann-Aussage'
Wenn bestimmte Bedingungen erfüllt → dann bestimmte Konsequenzen → Bsp: Skocpol zwei Notwendige Bedingungen (Zusammenbruch des Staates + Bauernrevolte)

12.2.4. Je-desto

Argument der Ausprägung; Variablen nicht nominal, sondern mindestens ordinal skaliert
→ Zusammenhang zwischen zwei Variablen Form einer 'Je-desto'-Hypothese

folgenden Bedingungen: x die unabhängige und y die abhängige

- Wenn x zunimmt, dann nimmt auch y zu; oder
- wenn x abnimmt, dann nimmt auch y ab; oder
- wenn x zunimmt, dann nimmt y ab; oder
- wenn x abnimmt, dann nimmt y zu.

(1) und (2) → positiven Zusammenhang (monoton steigend)

(3) und (4) → negativen Zusammenhang (monoton fallend)

Bsp: Je höher die Anzahl Vetospieler, desto weniger Politikveränderung (Tsebelis 1999)

12.3. Konzepte

12.3.1. Konzeptdefinition:

- in Hypothese verwendeten Begriffe (Konzepte) müssen präzisiert/definiert werden
→ nicht einfach nur Definition sondern überlegen, was Konzept ausmacht (bei Krankheit sucht man nicht Symptom sondern Ursache, nicht Symptom von A wirkt auf B, sondern Charaktereigenschaft von A)

- Müssen auf ihre verschiedenen Ebenen hin erläutert → da verwendeten Begriffe meist mehrdeutig und vielschichtig

Ziel → eine genauere Definition des verwendeten "Konzepts"

Wichtig → verschiedenen Dimensionen des Begriffes nicht nur herausarbeitet, sondern auch Verbindungen, die zwischen diesen Dimensionen liegen

• Typen von Konzeptdefinition

- Definitionen mit **Kriterien** als notwendige (Multiplikation)/hinreichenden (Addition) Bedingungen
- **Ähnlichkeit** zu Familien von Objekten → oft hinreichend → da Konzept nicht bestimmte essentielle Eigenschaft, lässt sich aber Gruppe/Familie zuordnen

• Eigenschaften von Konzepten

Haben **positiven/negativen** Pol

- Krieg: negativer Pol → Frieden, Nicht-Krieg
- Demokratien: negativer Pol → Diktatur, Monarchie.

sollen beide Pole berücksichtigen → macht verwendetes Konzept (verwendeten positiven Pol) klarer. Ausserdem überlegen, was zwischen den Polen liegt. Wenn Demokratie dichotom → würde nichts zwischen den Polen liegen → entspricht meistens nicht der Realität (Grenzfälle). Dichotome Einteilungen können zu Messfehlern führen, weil dann alle Mitglieder einer Kategorie gleichsetzt → besser stetige Konzepte.

12.4. Hypothesen und Kausalität

Kausalität: Beziehung zwischen Ursache und Wirkung

Korrelation: Beziehung zwischen zwei oder mehr statistischen Variablen. Wenn sie besteht → nicht gesagt, ob eine Größe die andere kausal beeinflusst, ob beide von einer dritten Größe kausal abhängen oder ob sich überhaupt ein Kausalzusammenhang folgern lässt.

- meisten Hypothesen (oder deren theoretische Untermauerung) gehen von kausalen Zusammenhängen aus.
- stärkste Art von Kausalität in Wenn-dann-Hypothesen (mit notwendigen und hinreichenden Bedingungen)
- Probabilistische Hypothesen kommen öfter vor, weil in den Sozialwissenschaften die theoretischen Grundlagen oft für eine deterministische Hypothese nicht reichen.
- Bei der Überprüfung von deterministischen Wenn-dann-Hypothesen müssten eigentlich einzelne "falsifikatorische" Fälle zur Verwerfung der Hypothese führen. Ob dies der Fall sein soll oder nicht, wird in der Literatur heftig debattiert.

12.5. Operationalisierung

Messen der Variablen/Konzepte

13. Variablen

13.1. Allgemein

13.1.1. Definition

Eine Variable bezeichnet ein Merkmal von Merkmalsträgern.

- Variablen (Merkmale) z.B. Haarfarbe
- Ausprägungen von Variablen (Kategorien) z.B. braun, schwarz, grau, weiss, etc.
- Merkmalsträger z.B. Personen

13.1.2. Konstruktion von Variablen:

Kategorien einer Variable müssen

- **disjunkt** (alle Merkmalsträger können eindeutig einer - und nur einer - Kategorie zugeordnet werden) → nicht disjunkt: z.B. Frage nach Beruf, dann taucht als Antwortmöglichkeit "Student" auf aber auch "Teilzeitangestellter" - man kann beides sein.
- **erschöpfend** sein (alle Merkmalsträger können einer Kategorie zugeordnet werden) → nicht erschöpfend: Demokratie, Autokratie - es gibt Länder, die kann man weder zum einen noch zum anderen zählen kann (z.B. Russland).

13.1.3. Typen von Variablen

- nach Ausprägungen (kontinuierlich/ diskret (polytom, dichotom))
- nach Merkmalsebene (Individuen oder Kollektive, vor allem bezüglich Messung relevant)
- nach Skalenniveau (nominal, ordinal, Intervall- oder Ratio-niveau)

13.1.4. Skalentyp

(Messniveau, Skalenniveau): sagt aus, welche Transformationen und Rechenoperationen mit den Variablenwerten ausgeführt werden dürfen.

	• Absoluter • Nullpunkt	• Abstände	• Ränge	• Identität	• Beispiel
Nominal	Nein	Nein	Nein	Ja	Familienstand
Ordinal	Nein	Nein	Ja	Ja	Zufriedenheit
Intervall	Nein	Ja	Ja	Ja	Temperatur in C
Ratio	Ja	Ja	Ja	Ja	Länge

•Nominal

sagt nur etwas über Gleichheit oder Verschiedenheit aus (=Äquivalenzklasse) z.B. Geschlecht: Mann=0, Frau=1. Transformation: alles, was die Unterscheidbarkeit nicht zerstört. Rechenoperationen: fast keine z.B. Mittelwert von 0.5 ist sinnlos. Einzig Modus (Modalwert) = Wert der Klasse, mit der höchsten Besetzungszahl.

•Ordinal

Zufriedenheit mit der Regierung kann man in eine Reihenfolge (Ränge) von sehr zufrieden bis sehr unzufrieden bringen. Transformation: alles, was die Reihenfolge beibehält. Rechenoperation: Median = Wert der die Gruppe in je 50% aufteilt.

•Intervall

Der Temperaturabstand (Intervall) von 20°C zu 21°C ist gleich wie jener von 10°C zu 11°C. Eine Strecke kann doppelt so lange wie eine andere sein. 10°C ist aber nicht doppelt so warm wie 20°C. Transformation: Abstände müssen gleich bleiben → $f=aX+b$. Rechenoperation: arithmetisches Mittel (aber keine prozentualen Aussagen möglich).

•Ratio (Verhältnisskala)

Transformation: $f=aX$. Rechenoperation: geometrisches Mittel → Quotientenbildung möglich = Aussagen über Verhältnis von Werten zu einander.

Häufig wird einfach angenommen, Daten hätten Intervall-Niveau (ohne genau Prüfung), insbesondere bei Nutzenmessungen umstritten, wo man auch mit Framing-Effekten zu tun hat = Leute entscheiden unterschiedlich, je nach dem, wie - an sich gleiche - Nutzen dargestellt werden. Skalenniveau entscheidet über die möglichen statistische Analyseverfahren → vorher überlegen, wie man Daten auswerten möchte.

Skalierungen = Messung auf der Basis eines Skalierungsmodells (es gibt eine Annahme über die Struktur der Beobachtungen).

Likert-Skalen (Bsp. Umweltmessung oben) kann man nutzen, um ungeeignete Indikatoren (=Fragen z.B. zu Umweltbewusstsein) zu erkennen. Vorgehen: von 1 (kein Umweltbewusstsein) bis 5 (hohes Umweltbewusstsein) können Befragte ihre Zustimmung ausdrücken. Sollte eigentlich für alle Fragen ähnlich ausfallen = konsistente Antwortreaktionen. Frage kann nun oft stark von übrigen Antworten abweichen. Grund: Frage enthält Fremddimension, die nichts mit Umwelt zu tun hat. Man bildet nun Summe pro Befragter für alle Indikatoren. Hohe Summe = hohe Umweltbewusstsein. Die Antworten zu einzelnen Fragen müssten hoch mit den Summen korreliert sein. Falls das nicht so ist → Indikator ungeeignet zur Messung von Umweltbewusstsein.

Guttman-Skalen (Indikatoren werden in eine Hierarchie gestellt (z.B. Teilnahme in Wahlen, Mitglied einer Partei, Teilnehmer an Demonstrationen, etc.)

Index = Variable, deren Wert sich aus Rechenoperation mit anderen Variablen ergibt:
 $I=f(X_1, X_2, \dots, X_n)$

13.2. Variablen und Messungen

13.2.1. Messung

- Basierend auf der Operationalisierung der Konzepte
- Zuordnung von Zahlen nach bestimmten Regeln
- 1 Indikator oft ungenügend → mehrere Indikatoren verwenden
- Qualität der Messung hat 3 Kriterien → Gütekriterien
 - Objektivität
 - Reliabilität
 - Validität

13.2.2. Objektivität

Gleiches Resultat egal wer Messinstrument benützt; Distanz messen, wenn Massstab stimmt (Messinstrument) gemessene Strecke immer gleich lang

Problematische Beispiele:

- Umfragen → Haushaltsarbeit von Männern, andere Antwort ob Interviewer Mann/ Frau
- Kodierung von Medienberichten

13.2.3. Reliabilität

Zuverlässigkeit, Reproduzierbarkeit von Messungen → Anwendung des gleichen Messinstrumentes auf gleiche Untersuchungseinheit sollte gleiche Resultate geben

Wird kontrolliert durch:

- **Paralleltest-Methode:** 2 Messinstrumente parallel → Messung erfolgt mit zwei vergleichbaren Messinstrumenten (2 verschiedene Fragebögen) Korrelation Instrument A mit Instrument B informiert über Reliabilität.
- **Test- Retest-Methode:** 1 Messinstrument 2mal → Messinstrument wird nach einem Zeitintervall wiederholt angewendet. Messung Korrelation der Zeitpunkte. Problem hier: die zeitliche Stabilität muss gegeben sein. Bsp. Berner Umweltstudie, Retest 1/2 Jahr nach dem Test. Allenfalls hat sich Umweltbewusstsein der Befragten zwischenzeitlich verändert.
- **Methode Testhalbierung:** Aufteilung von multiplen Indikatoren → z.B. Einstellung zur Einbürgerung messen. Kann man mit verschiedenen Indikatoren tun → teilt diese Indikatoren in 2 Gruppen. Man fragt z.B. was halten Sie für wichtig:
 - 1 Schweizer Vorfahren
 - 2 in der Schweiz geboren
 - ...
 - 9 Schweizer Gesetze werden befolgt.

- → Die beiden Gruppen sollten dann das gleiche Ergebnis liefern. Da mit höherer Anzahl Indikatoren auch Reliabilität steigt, ist sie in zwei halb so grossen Gruppen dann tiefer.

13.2.4. Validität

Gültigkeit von Messungen: misst das Messinstrument, was gemessen werden muss? Stützt sich mehr auf Argumentation als (wie bei Reliabilität) auf Empirie.

Unterschieden werden:

• Inhaltsvalidität

Es gibt unendlich viele Indikatoren. Die gewählten sollten die zu messenden Eigenschaften repräsentieren = repräsentative Stichprobe von Indikatoren. Der Polity IV hat drei Dimensionen. Wenn man eine einfach weglässt, misst man nicht mehr alles gemäss Definition des Konzeptes.

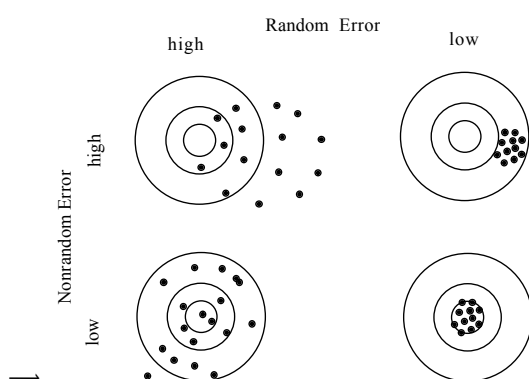
• Kriteriumsvalidität

Was man misst, soll hoch korreliert sein mit einer anderen Messung des gleichen Konzepts. Z.B. man will man mit Test Depression messen an Menschen die nachweislich Depression haben. Das Testergebnis vergleicht man dann mit Ergebnissen anderer Messungen (z.B. Einschätzung des Therapeuten)

• Konstruktvalidität

Man hat Theorie und Hypothese mit abhängiger und unabhängiger Variable und den Wirkungsmechanismus. Was wir messen muss auch diese Kausalität ausweisen. Test: z.B. misst man Umwelteinstellung und Kooperationsverhalten mit Umfrage und mit Verhaltensanalyse → Ergebnisse von Verhaltensanalyse und Interview über Umwelteinstellung muss höher korreliert sein als die beiden Befragungen zu verschiedenen Themen → man kann zwischen verschiedenen Konstrukten differenzieren.

Für das gleiche Konzept bestehen oft unterschiedliche Indikatoren, einige sind nicht valide. Der **BigMac-Index** = Indikator für Kaufkraft. Es fehlt die Inhaltsvalidität, weil er nur ein einziges Gut benutzt (anstelle eines Warenkorb). Der Index zeigt auch Unter- und Überbewertung von Währungen. Es entsteht ein Konflikt zur Kriteriumsvalidität. Die Konstruktvalidität ist verletzt, weil z.B. der BigMac in der Schweiz aufgrund hoher Mietpreise teurer ist. Man will aber mit dem Index nicht Mieten messen.



Messen ist mit dem Schiessen auf Zielscheibe vergleichbar. Ziel ist möglichst immer die Mitte zu treffen. Nur unten rechts ist valide und reliable. Unten links ist im Durchschnitt im Zentrum (=valide) aber nicht reliabel, weil zur grosse Streuung. Oben rechts ist nicht valide aber reliabel, weil man zwar falsch misst, dafür aber das gleiche Ergebnis erzielt.

13.2.5. Reaktivität von Messungen

Problem auf, die mit "unobtrusive measures" (unauffällige Messungen) vermieden werden können = Messung, die der Proband nicht wahrnimmt.

- Umfragen sind kaum unauffällige Messungen
- Wasserverbrauch oder Bodenabnutzung in Museum sind unauffällige Messungen
- Messung der Abwassermenge kann Aufschluss geben über Benutzung von Zweitwohnungen in Touristengebieten: unauffällige Messung

14. Forschungsdesign

14.1. Allgemein

14.1.1. Definition

Erhebungsdesign und Forschungsdesign nicht → genau das gleiche

- Forschungsdesign: allgemeiner gehalten
- Erhebungsdesign: die Erhebung der Daten → Mittel zum Zweck der Sammlung aussagekräftiger Daten.
- Theorie, Hypothesen und Operationalisierung beeinflussen → welche Forschungsdesign am besten
- Forschungsdesign und Operationalisierung wählt man darum oft parallel → Design sicherstellen → bestmöglich Weg wählt, die Forschungsfrage zu beantworten.

Design Einfluss → Aussagekraft: wenn man z.B. einfach 5% Arbeitslosigkeit misst → noch nicht, wie viele Prozent pro Jahr von Arbeitslosigkeit betroffen → könnten 10% sein

14.1.2. Überlegungen beantworten

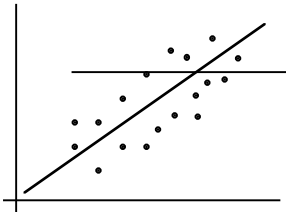
- Untersuchungseinheit - und ebene
- Zeitdimension (Querschnitt, Trend, Panel)
- Varianz (Experiment, Quasiexperiment, Ex-Post-Facto Design)

14.2. Unbestimmtes Forschungsdesign

Auswahl eines Forschungsdesigns → acht geben, erlaubt unsere Fragen zu beantworten

- **Anzahl erklärende Konzepte/Variablen, Anzahl Fälle** → zusätzliche unabhängige Variable braucht grössere Anzahl Fälle, ansonsten unbestimmtes Forschungsdesign
- **Beziehung zwischen erklärenden Konzepten/Variablen (Multikollinearität)** → Variablen dürfen nicht vollständig korrelieren. Bsp: Demokratie, Handel, Krieg. Problem: wenn nicht-demokratische Länder immer auch wenig Handel haben (aber manchmal Krieg), dann können die beiden Effekte (welche beiden??) nicht auseinander gehalten werden.

- **Auswahl der Fälle** → nicht zufällige Auswahl der Fälle verzerrt Ergebnis. Bsp: Umfrage in CH nur unter Stimmberechtigten → die Grundgesamtheit klar, die Stichprobe zufällig. Aber: zufällige Auswahl oft nicht möglich. Bsp. Frage nach Demokratie und Krieg ist unklar, welche Länder überhaupt Grundgesamtheit → Zufallsauswahl unmöglich wenn Grundgesamtheit unklar.



Auswahl der Fälle anhand abhängiger Variable: Wenn man z.B. nur Fälle auswählt, deren abhängige Variable ein gewisses Niveau überschreiten, dann fallen mehr Fälle weg, die nach unten streuen → der Effekt schwächt sich ab, die Regressionsgerade wird flacher. Die Fälle sollen daher anhand der unabhängigen Variable ausgesucht werden. Problem: manchmal ist unabhängige Variable erst nach der Auswahl der Fälle bekannt.

14.3. Untersuchungsebene und einheit

- Theorie und Hypothese → Untersuchungseinheit bestimmen
- Jedoch bei gewissen Fragen → müssen andere Ebenen gewählt werden, Bsp: wer wählte Hitler
 - Kaum Umfrage/Individualdaten vorhande
 - Daher Aggregatsebene → Daten von Gemeinden, Bezirken → neue Untersuchungseinheit

14.4. Zeitdimension

14.4.1. Querschnittsdesign

- 1 Zeitpunkt oder kurze Zeitspanne → einmalige Erhebung
- Für Eigenschaft (Variablenwert) bei N Untersuchungseinheit

14.4.2. Trenddesign

- Werte der gleichen Variable
- Zu mehreren Zeitpunkten
- Mit unterschiedlichen Stichproben

→ Veränderung auf Aggregatsebene

14.4.3. Paneldesign

- Werte der gleichen Variable
- Zu mehreren Zeitpunkten
- Mit gleicher Stichprobe

→ Veränderung auf der individuellen Ebene; Trends können mit Panel besser geschätzt werden, da kein Stichprobenfehler wie beim Querschnitt existiert. Beim Panel hat man immer Ausfälle über die Zeit hinweg - Ausfälle meist nicht zufällig sondern systematisch (z.B. Berufstätige) → systematische Stichprobenfehler. Ziel: Panelmortalität tief halten, ist aufwendig durch Pflege der Adressen.

14.4.4. Informationsgehalt

Panel > Trend > Querschnitt. Bei einem Panel z.B. über Jahre hinweg hat man mal eine repräsentative Stichprobe aus der Bevölkerung gezogen. Die Bevölkerung ändert mit den Jahren (Zusammensetzung z.B. Ausländer), das Panel aber nicht → Trends der Gesamtbevölkerung besser mit Trenddesign, weil in diesem immer wieder neue Stichproben gezogen werden.

14.4.5. Kohortendesign (Spezialfall von Längsschnitterh.)

Kohorte = Bevölkerungsgruppe, die durch ein gemeinsames Starterereignis definiert ist (z.B. Geburt, Eheschließungszeitpunkt, Berufseintritt). Kohorten können ex ante (z.B. gezielte Auswahl bestimmter Jahrgänge) oder ex post (Daten nach Jahrgang sortieren) gebildet werden.

• Kohorteneffekt

Kohorte ist gleichen Einflüssen ausgesetzt, sollte sich z.B. auf den Lebenslauf auswirken.

• Lebenszykluseffekt

Merkmal hängt von der Zeit ab, die seit Starterereignis verstrichen ist (z.B. Wahrscheinlichkeit zu sterben, Ehescheidungsrisiko ab Zeitpunkt Eheschließung). Man kann Lebenszykluseffekt bei mangelnden Daten aus Querschnittsdaten schätzen, wenn man glaubt, es liege kein Kohorteneffekt vor (aber: Gefahr eines Lebenszyklus-Fehlschluss). Bsp.: Querschnittsdaten ergaben abnehmende Intelligenz mit zunehmendem Alter → Schlussfolgerung war, Intelligenz nimmt ab, wenn Mensch älter wird = Lebenszykluseffekt. Bei vielen Dimensionen stimmt das nicht, die Älteren hatten einfach eine geringere Bildung (=Kohorteneffekt), Abnahme wegen Alter kann oft nicht gezeigt werden.

• Kohortenfehlschluss

Ein Lebenszyklus-Effekt wird als Kohorteneffekt gedeutet. Bsp. Inglehart brachte 1977 die Postmaterialismushypothese, wonach das Geburtsjahr (und die entsprechende Sozialisierung) einen Einfluss auf die Werte hat. Querschnittsdaten könnten das bestätigen. Problem: Es könnte auch sein, dass man Werte im Laufe des Lebens ändert (jünger: postmaterialistisch, älter: materialistisch), das würde bedeuten, dass nicht das Geburtsjahr entscheidet. Fazit: Kohorteneffekt, Lebenszykluseffekt können nur mit Längsschnittdaten geprüft werden.

14.5. Varianz

Varianz: Maß, wie die einzelnen Daten um Mittelwert verteilt sind, d.h. wie stark die Daten um Mittelwert streuen.

14.5.1. Experiment

- Kann die unabhängige Variable gezielt manipulieren
- Mindestens zwei experimentelle Gruppen
- Einfluss von Störgrößen durch Konstanthaltung der Versuchsbedingungen kontrollieren
 - Elimination
 - Randomisierung (Personen zufällig auf Gruppen verteilt → sonst Problem der Selbstselektion)
 - Parallelisierung
- 2 Arten von Experimenten:
 - Laborexperimenten: Versuchsbedingungen vollständig kontrolliert → hohe interne Validität
 - Feldexperimente: aufgrund der natürlichen Umgebung → hohe externe Validität

• Design von Experimenten

„Normales Design“:

Schreibweise: **O** = Observation/Beobachtung, **X** = Intervention (Eingriff), **R** = Random

Frage: Beeinflusst Bericht Wahrnehmung?

X → Bericht über den Irakkrieg

O → die Befragung nach Wichtigkeit des Thema?

1.Design:

X O

Problem: Thema für wichtig gehalten → weiss nicht, ob Person schon vorher für wichtig hielt → deshalb Kontrollgruppe = Zufallsmässige Zuteilung in zwei Gruppen

2. Design:

R X O (Versuchsgruppe)

R O (Kontrollgruppe) → müsste nun Thema als unwichtiger einstufen

Kontrollgruppen → kann eine Variable zu beeinflussen während alles andere konstant

Problem: - durch zufallsmässige Zuteilung → können andere Erklärungsfaktoren (Drittvariablen) neutralisiert

- **Hawthorne-Effekt** → durch Untersuchung wird Untersuchung beeinflusst → Bias (Verzerrung durch Reaktivität) → Leute geben sich mehr Mühe

3. Design:

R O1 X O2 (Versuchsgruppe)

R O3 O4 (Kontrollgruppe)

misst Kontrollgruppe und Versuchsgruppe zweimal → vor und nach der Intervention (Licht verändern). Kontrollgruppe gab es keine Licht-Veränderung → durch zweimalige Messen → Effekt durch Experiment entsteht kontrollierbar

Solomons Vier-Gruppen Design:

R O1 X O2

R O3 O4

R X O5

R O6

- O2 und O5 unterscheidet → ob vor Intervention schon eine Messung
- O1 X O2 sollte eine Erhöhung der Produktivität zeigen
- O3 O4 hat auch eine Produktivitätserhöhung (Effekt, der nicht interessiert).
- Vergleich O2 zu O4 → Effekt der Intervention
- O5 zu O6 Kontrolle → Wert sollte gleich sein wie O2 zu O4. → verhindert, dass Verzerrung durch Reaktivität vernachlässigt wird

• Validität von Experimenten

Interne Validität: kann der Effekt auf die manipulierten Variablen (X) eindeutig zurückgeführt werden? Ist es also die Intervention, die zur Veränderung der abhängigen Variablen führt?

gefährdet durch:

- **Geschichte:** Experiment über gewisse Zeit hinweg → Zwischenzeit geschieht etwas → Verhalten der Leute beeinflussen können.
- **Maturation:** Experiment verschiedene Phasen → Leute gewöhnen sich an Untersuchungssituation
- **Test:** weiss, dass man Forschungsobjekt ist
- **Instrumente:** Wie beeinflusst mich Instrument → Bsp: Briefe meist als unwichtig und persönliche Information als wichtiger eingestuft werden.
- **statistische Regression:** Messbild über Zeit wird nicht gleich sein → Verbesserung der Messung
- **Regressionseffekt:** Tendenz zur Mitte. Bsp.: Kinder grosser Eltern → grösser als Durchschnitt → aber kleiner als Eltern, = Tendenz zur Mitte, sonst gäbe es plötzlich Riesen und Zwerge. Bsp. Schlechte Schüler werden mit Massnahme gefördert. Danach gibt's Leistungssteigerung. Aber: Schlechte Schüler bringen entweder gleich schlechte oder bessere Leistung = Tendenz zur Mitte. Somit ist Verbesserung nicht unbedingt Resultat der Massnahme.

- **Auswahl:** Die Zuteilung zu Versuchsgruppe und Kontrollgruppe ist nicht zufällig, was dann Ergebnisse beeinflusst. Je grösser die Gruppen, desto eher hat man eine Zufälligkeit.
- **Experimentalsterblichkeit:** die Leute steigen aus dem Experiment aus
- **Interaktion zwischen Auswahl und Maturation:** die nicht zufällige Zuteilung in Gruppen kann auch zu unterschiedlich ausgeprägter Maturation führen.

Externe Validität: sind die aufgezeigten Effekte auch ausserhalb des Experimentalkontextes zu erwarten? Ist also eine Generalisierung möglich?

externe Validität ist gefährdet durch:

- **reaktive Effekt auf Experiment:** das sind Effekte, die nur durch das Experiment selbst entstehen
- **Interaktion zwischen Auswahl und Stimulus:** Gruppen/Personen können unterschiedlich auf Stimulus reagieren → wichtig Aufteilung von Personen in Gruppen
- **reaktive oder Interaktionseffekte:** Serie von Interventionen → wie ich auf 3. Intervention reagiere abhängig von 1. + 2. Intervention
- **Einfluss anderer Stimuli:** es können ganz andere Dinge den beobachteten Effekt ausgelöst haben → Salamiegeschmack in Luft

14.5.2. Quasi-experimentelle Designs

- Zuweisung der Versuchspersonen zu Kontroll- und Experimentalgruppe in Untersuchungen → nicht durch Randomisierung oder Parallelisierung → 2 Probleme
 - Selbstselektion in beiden Gruppen → Einfluss auf andere Variablen
 - Ausfall von Untersuchungseinheit → kann systematisch sein (leistungsschwache Schüler brechen ab)

Lösung: kann versucht werden für andere unabhängige Variablen zu kontrollieren (vergleichbare Fälle, statistische Kontrollen) → Man bildet Gruppen mit Ähnlichkeiten, so dass man vergleichbare Fälle (im Bsp. bezüglich Bildung) hat. Die Wirkung des Stimulus wird dann innerhalb der vergleichbaren Fälle betrachtet.

	• Gelesen	• Nicht gelesen
wichtig	20	30
unwichtig	20	30

Man befragt Leute, ob sie etwas über das Thema gelesen hätten, und ob sie das für wichtig oder nicht halten. Nebenstehendes Ergebnis zeigt keinen Effekt auf die Einstufung der Wichtigkeit.

Unterteilung Personen in Bildungsniveau:

	• hoch		• niedrig	
	Gelesen	N. Gelesen	Gelesen	N. gelesen
Wichtig	5	20	15	10
unwichtig	15	10	5	20

Es zeigt sich nun, dass bei den Leuten mit tiefen Bildungsniveau die Wichtigkeit bei 75% der Leute gegeben ist, die auch etwas darüber gelesen haben. Bei den besser gebildeten gibt es einen gegenteiligen Effekt.

- 2 quasi-experimentelle Untersuchungsanordnungen.
 - gemessener Eigenschaften der Versuchsobjekte Bsp: Viel- und Wenigsehern; Personen mit hohem Fernsehkonsum → Experimentalgruppe, Personen mit geringem Fernsehkonsum → Kontrollgruppe
 - "natürlichen Experiment": natürlichen, nicht kontrollierten Ereignissen Bsp: Einführung neuen Schultyps/ der Verbreitung eines neuen Mediums.
- keinen Rückschluss auf kausale Zusammenhänge → nicht sicher ob Unabhängige die Abhängige umgekehrt beeinflusst

14.5.3. Ex-post-facto-Design

- Nicht Anforderung für experimentelle/ quasi-experimentelle
- unabhängige und abhängige Variablen gemessen
- Störvariablen nicht kontrollierbar
- nur korrelative Aussagen
- Vorteil: geringem finanziellen und personellen Aufwand

15. Datenerhebung

15.1. Stichprobe

15.1.1. Definition

Stichproben: erlauben Rückschlüsse auf die Grundgesamtheit (Population)

Grundgesamtheit (Population): alle Einheiten, die uns für unsere Untersuchung interessieren

Stichprobe (sample): Anzahl von Elementen aus der Grundgesamtheit

Erhebungseinheit: Elemente der Grundgesamtheit

Untersuchungseinheit: auf diese beziehen sich erhobene Daten, muss nicht mit Erhebungseinheit übereinstimmen, z.B. Lehrer (Erhebungseinheit) werden über Schüler (Untersuchungseinheit) befragt)

Zielpopulation: angestrebte Grundgesamtheit. Man wählt leider oft aus einer Gesamtheit aus, die nicht mit Zielpopulation übereinstimmt (=Auswahlgesamtheit, frame population). Schnittmenge der beiden ist survey population.

15.1.2. Stichprobenverfahren 3 Typen

Vorschrift, die festlegt, wie Elemente aus der Population ausgewählt werden

• Wahrscheinlichkeitsauswahl (Zufallsauswahl)

genau sagen, mit welcher Wahrscheinlichkeit ein Element in der Stichprobe auftaucht.

- **einfache:** einstufige Auswahl aus Gruppe → jedes Element mit gleicher Wahrscheinlichkeit in Stichprobe; nur möglich → Grundgesamtheit bekannt.
- **mehrstufige:** zuerst eine Auswahl von Gruppen getroffen → danach aus den Gruppen zufällig Elemente gezogen (z.B. 5 Kantone, danach je 200 Leute → nicht gut, weil kleiner Kanton zu stark vertreten)
- **PPS-samples (probability proportional to size):** die Auswahlwahrscheinlichkeit eines Kantons (oder Gemeinde) → proportional zur Grösse, danach kann pro Kt. gleich viele Leute
- **Klumpensamples:** Bsp. zufällig Strassenabschnitte einer Stadt, danach alle Leute dort befragt. Ist zweistufiges Verfahren mit Berücksichtigung aller Elemente der zweiten Stufe.
- **Schichtung:** Bsp. Durchschnittseinkommen ermitteln - Wahrscheinlichkeit, mit einfacher Auswahl aMillionäre hat → sehr gering → geschichtete Zufallsauswahl = Auswahl nach Einkommensgruppen. Bedingt aber Kenntnisse über die Grösse dieser Gruppen

• bewusste Auswahl

- **Quotenauswahl (Gallup):** Stichprobe gewisse Charakteristika (z.B. 50% Frauen) abzubilden → keine Zufallsauswahl mehr. Wahrscheinlichkeit → Individuum aus der Grundgesamtheit in Stichprobe → nicht mehr gewährleistet. Setzt Wissen über die Grundgesamtheit voraus. Quote garantiert nicht, dass Merkmale, für die man sich interessiert (z.B. Wahlverhalten) unverzerrt abgebildet werden. **Gallup vs. Literary Digest**

• willkürliche Auswahl

- **Schneeballauswahl:** Umfrage unter Obdachlosen, kennt fast keine → Deshalb befragten Obdachlosen nach weiteren. Wahrscheinlichkeit, dass jemand in Stichprobe auftaucht, kann nicht bestimmt werden.

15.1.3. Wieso Zufallsstichproben?

Zufallsstichproben → Rückschlüsse auf Grössen (Parameter) der Grundgesamtheit.

- n von N Personen sind Wähler von Ségolène Royal
- die Wahrscheinlichkeit in Zufallsstichprobe der Grösse 1, einen Wähler von Ségolène Royal haben ist n/N .
- grösseren Zufallsstichproben können wir n schätzen → wissen **Unsicherheit dieser Schätzung** → deshalb wichtig zu wissen, mit welcher Wahrscheinlichkeit ein Individuum in der Stichprobe auftaucht, d.h. deshalb braucht man überhaupt eine zufällig Stichprobe.

Unsicherheit hängt von 2 Faktoren ab:

- Stichprobengrösse
- Wie Charakteristik in Stichprobe verteilt ist

Beispiel:

- bei 100 Befragten → 50% für Segolène → grössere Unsicherheit als
- Bei 10000 Befragten → 50% für Segolène → kleinere Unsicherheit
- 5% für Segolène → kleinere Unsicherheit als
- 50% für Segolène → grössere Unsicherheit

Fazit: mit Stichprobengrösse und Heterogenität nimmt Unsicherheit ab

15.2. Umfrage

15.2.1. 3 Arten von Fehlerquellen

• Zufallsfehler der Stichprobe

→ auch in der perfekt zufällig gezogene Stichprobe → nur zu vermeiden → ganze Population befragt

• Systematische Fehler aufgrund Verfahrens der Stichprobenauswahl

→ bestimmte Personen mit höheren Wahrscheinlichkeit in Stichprobe als andere

• Verzerrung („nonsampling bias“)

→ keinen Zusammenhang mit Stichprobenauswahl

15.2.2. „nonsampling bias“

• Messfehler

z.B. sagt die befragte Person die Wahrheit

• Fehlerquellen im Interview:

Interview bildet nicht ab, was eigentlich interessiert

• Diskrepanz zwischen Zielpopulation und Surveypopulation (Coverage-Fehler)

z.B. Leute anhand von Telefonregistern → stimmt Zielpopulation nicht mit Surveypopulation überein → weil Personen ohne Telefonanschluss wegfallen

- **Antwortverweigerung: "non-response"**

Leute geben insgesamt keine Antwort (unit non response) oder verweigern die Auskunft zu bestimmten Fragen (item-non-response) **Ausschöpfungsquote**: Anzahl ausgewertete Interviews im Verhältnis zur bereinigten Stichprobe (ist Stichprobe - stichprobenneutrale Ausfälle wie Tod).

15.2.3. Formen der Befragung

- **persönlichen Interviews**

→ für komplexe Sachverhalte geeignet z.B. wenn man Bilder, Grafiken zeigen will, aber teuer

- **telephonischen Interviews (preiswert)**

- **schriftlichen Befragungen**

→ Person entscheidet selbst zum mitmachen, wann sie die Befragung durchführt, hat keinen Kontakt zu Befrager → für heikle Themen geeignet, geringe Kosten Nachteile: Rücklaufquote tief, Fragebogen muss einfach sein, Verständnisprobleme können nicht gelöst werden, füllt wirklich Zielperson aus?

strukturierte Befragung: Fragen und Antwortmöglichkeiten sind festgelegt (erhöht die Objektivität gegenüber unstrukturierter Befragung, aber keine Info jenseits der Antwortmöglichkeiten).

unstrukturierte Befragung: Befrager weiss, welche Punkte er anzusprechen hat (Leitfaden). Ansonsten → Freiheit in der Interviewgestaltung.

15.2.4. Was misst man in Befragungen?

Befragungen → die Antworten auf bestimmte Fragen erfasst.

- Die Fragen (und die Befragungssituation) werden am besten als Stimulus betrachtet auf welchen der Befragte reagiert. (Stimulus → Reaktion: misst diese (verbale) Reaktion)
 - Stimulus je nach Kultur sehr unterschiedlich Bsp. hohe Neigung zur inhaltsunabhängigen Zustimmung in manchen Kulturen.
- Antworten enthalten fast gezwungenermassen → systematische und zufallsabhängige Fehler.
- letzteren versucht man möglichst durch gute Frageformulierungen zu vermindern.

15.2.5. Erkenntnisse zur Fragenformulierung

- praktischen Erfahrungen
- Experimenten (Vergleich von verschiedenen Fragenformulierungen in Teilstichproben)

Grundsätzliches Problem: man nimmt an → gebe "richtige" Antwort einer Person auf eine Antwort. Item will eine Zieldimension messen. Richtige Antwort kann verzerrt werden durch Fremddimension und soziale Erwünschtheit.

15.2.6. Einflussfaktoren

- **Entscheidungstheorien** → wie entscheidet der Befragte, was er antwortet. Manchmal anhand von Überlegungen, die nicht zur wahren Antwort führen (z.B. Risikoeinschätzung, dass Steueramt davon erfährt).
- **Hypothese zur Informationsverarbeitung:** wie werden Infos vom Befragten verarbeitet / bzw. abgerufen → kann zu Verzerrungen führen.
- **Fragestellung selbst:**
 - Frageformulierungen
 - neutrale Formulierungen → inhaltliche identische Begriffe können unterschiedlich wirken z.B. nicht erlauben vs. Verbieten
 - kurz, verständlich und nicht suggerierende Fragen → eine Antwortmöglichkeit gegenüber als erwünscht erscheint.
 - Mittelkategorien bei Skalen
 - Möglichkeit der Antwortverweigerung
 - Rating oder Ranking

Rating: man fragte Leute nach verschiedenen Politikbereichen, die sie in Skala von "ernstes Problem" bis "kein ernstes Problem" einteilen mussten → generell erhalten Themen höhere Wichtigkeit, weil nicht einordnen müssen

Ranking: die Befragten müssen die Politikbereiche sortieren nach Wichtigkeit → kann zu Unterschieden führen welche Methode

- **Mittelkategorie in Skala:** beeinflusst 1) die Antworten 2) ob die Leute überhaupt antworten → keine Mitte antworten manche lieber gar nicht
- **Möglichkeit der Antwortverweigerung:** "weiss nicht"/"Frage gar nicht beantworten." Problem: erzwingt man Antwort → unsicher, ob überlegte Antwort → möchte reflektiert antworten
- **Inhalt der Antwortmöglichkeit:** Schätzen/Vorgegeben
- **Interviewsituation:** Mann + Mann, Mann + Frau; Partner zuhause → kann weniger lügen

15.2.7. Typen von Fragen

• geschlossene Fragen:

bestimmte Anzahl von Antwortmöglichkeiten z.B. 5 Möglichkeiten von "sehr wichtig" bis "überhaupt nicht wichtig" → Ausprägungen müssen 2 Kriterien erfüllen

- 1) jede Person kann sich einer Kategorie zuordnen - das heißt die Antwortmöglichkeiten überschneiden sich nicht (=disjunkt)

- 2) es muss jede Möglichkeit in den Antwortmöglichkeiten enthalten sein (=erschöpfend).

Vorteil: Vergleichbarkeit der Antworten, Objektivität, leicht auszuwerten → ist dominierender Fragetyp.

• Halboffene Fragen:

hat Möglichkeit „andere Antwort“ zu geben als vorgegebene → da Problem, alle Möglichkeiten zu erfassen → deshalb Kategorie "andere"

Nachteil: schlicht einfacher, jene zu wählen, die schon da stehen → andere wird selten ausgefüllt

• offene Fragen:

Bsp: Vox-Analyse: Welche Argumente waren im Abstimmungskampf für sie entscheidend? → Wäre schwierig alle Antworten zu finden → Kategorien würden zu stark die Antworten kanalisieren.

Problem: wie Antworten auswerten

15.3. Experteninterview

Umfragen bis jetzt zielen auf → Informationen zu Grundgesamtheit → diese aber zu gross → daher Stichprobe

Experteninterview/ Grundgesamtheit spezieller Gruppierung

- Unstrukturierte (qualitative) Befragung in Grundgesamtheit
- Strukturierte Befragung in Grundgesamtheit

→ bei vielen Experteninterview strukturiert → mind. Leitfaden

15.3.1. Infos aus Experteninterviews

Meisst in gewissen Situationen nötig

- Entscheidungsprozess (z.B. Frage: Wie kommt man in der Schweiz zu bestimmten Entscheidungen? Man befragt Experten zum Prozess (Ablauf, Akteure, Interessen).
- Charakteristiken (z.B. Ideologie, etc.) von Parteien
- Regierungspositionen (z.B. welche Positionen bringen Regierung in internat. Verhandlungen ein?)

15.3.2. Typen von Experteninterviews nach Art der erhobenen Info:

• faktuelle Informationen

- wie Entscheidungsprozesse in der Schweiz durch den verstärkten Einfluss der Globalisierung und Europäisierung verändert (z.B. welche Ämter sind in Entscheidungen involviert) → Info Person aus dem Bundeshaus geben.

- wie wurde der im Konvent erarbeitete Entwurf des Verfassungsvertrag für die EU in den Regierungen der 25 Mitgliedstaaten behandelt (z.B. in welcher Reihenfolge und welche Ministerien wurden gebeten Stellung zu beziehen).
- welche Positionen haben die Regierungen der 25 EU Mitgliedsländer in den Verhandlungen zum EU Verfassungsvertrag eingenommen. (hat auch Teile von b) weil Positionen zur Diskussion stehen. Gehört etwas zu a) oder b) entscheidet sich auch an den Antworten → fast alle gleich, dann a)

• **Einschätzungen:**

- In mehreren Experteninterviews wurden die ideologischen Ausrichtungen von Parteien erfragt (Experten geben unterschiedliche Einschätzungen)

15.3.3. Typen der Befragung

Weitere qualitative unstrukturierte Befragungen (Expert/Nicht-Exp.)

• **Fokussiertes Interview**

Person wurde einem Stimulus ausgesetzt (=bestimmte Situation, Film, usw) → Erfahrung schildern = subjektive Einschätzung → Interview offen, aber mit Leitfaden, der Interview aber nicht strukturieren soll. Kann z.B. der Hypothesenbildung dienen.

• **Narratives Interview**

Erzählweise des Befragten im Vordergrund → gibt aber Regeln, wie man Dinge darstellt. Diese Regeln befolgen (unaufgefordert) viele Leute, z.B. will man vollständig oder verständlich darstellen

• **Problemzentriertes Interview**

ähnlich wie narratives Interview aber mit aktiveren Rolle des Befragers, der während der Erzählphase Frage stellt, die aber der Erzähllogik folgen.

Probleme dieser Arten:

- Erkenntnisse nachvollziehbar → Aufzeichnung, Leitfaden etc.
- Analyse dieser Aufzeichnungen und Informationen → oft sehr schwierig. (wie soll man genau Daten als Sekundärdaten zur Verfügung stellen? → mehr Datenarchive mit quantitativen Daten)
- Deshalb sehen Schnell, Hill und Esser (1999, 355) den Nutzen solcher Befragungen vor allem in folgenden Situationen:
 - zur Exploration, als Pretest, zur Hypothesenentwicklung, zur Systematisierung vorwissenschaftlichen Verständnisses
 - zur Analyse seltener oder interessanter Gruppen, die in grossen Stichproben nur in kleiner Zahl repräsentiert sind → man kann dann mehr Details fragen.
 - als Ergänzung und zur Validierung anderer Forschungsinstrumente
 - als Instrument einer qualitativen Sozialforschung

15.4. Textanalyse

15.4.1. Allgemein

Befragungen (Stichproben, Experten) sind reaktiv, weil die Befragten wissen, dass mit Ihnen etwas erhoben wird. Textanalyse hat dieses Problem nicht. Erste Anwendungen auf Kriegspropaganda im I WK.

Definition: Die Inhaltsanalyse (Textanalyse) ist eine empirische Methode zur systematischen, intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen.

Untersuchungsobjekte: Texte, welche Beobachtungen liefern

Ziel: Inferenzen - bezogen auf Sender, soziales Umfeld, Empfänger oder Texteigenschaften selbst (z.B. Analyse der Parteiprogramme mit dem Ziel der Inferenz bezüglich ideologischer Ausrichtung der Partei)

15.4.2. Ziele der Analysen nach Semiotik

(= die Frage, welche Elemente der Sprache man berücksichtigen will)

• Syntaktik

Ausschliesslich formale Elemente des Textes werden analysiert (z.B. wie häufig werden bestimmte Wörter benutzt?) → Zeichen-Zeichen-Relation.

• Semantik

Inhaltliche Elemente werden analysiert, sprich Bedeutung → Zeichen-Bedeutung-Relation

• Pragmatik

Aufnahme der Inhalte wird problematisiert → Sender-Zeichen + Zeichen-Empfänger-Relation.

15.4.3. Typen von Textanalysen

• Systematische Textanalysen

- **Frequenzanalyse** :Textteile, z.B. Worte, Quasi-Sätze, etc. werden gezählt
- **Kontingenzanalyse**: auch Betrachtung von Häufigkeiten aber mit Ermittlung von Assoziationsstrukturen → sind bestimmte Wortkombinationen miteinander verbunden?)
- **Bewertungsanalyse**: Valenzanalysen (Bewertung wird erfasst), Intensitätsanalysen (Intensität der Bewertung wird erfasst) - wenn z.B. vom Militär die Rede ist kann man erfasse ob dafür oder dagegen und auch wie stark negativ oder positiv.

•Computerunterstützte Textanalysen:

- Computer können mit spezifischen Programmen die Datenerhebung vereinfachen.
- Dabei können unterschiedliche Ansätze unterschieden werden:
- Analysen welche auf den Wortfrequenzen basieren
- Analysen welche auf Kodierungsregeln basieren (z.B. Analyse von Wahlprogrammen)
- Andere

sind auch systematisch, Syntaktik-Untersuchungen können sehr schnell durchgeführt werden → günstiger

Wordscores-Analyse der Schweizer Parteiprogramme → Referenztexte die Parteiprogramme der "extremsten" Parteien

Probleme: Begriffen ändert sich über die Zeit hinweg → macht Vergleich über die Zeit hinweg schwierig.

Regelbasierte Analysen währenddem für eine Wordscores-Analyse nur Referenztexte und entsprechende "Referenzwerte" nötig sind, basieren andere computerunterstützte Textanalysen auf "Wörterbüchern" und "Kodierungsregeln" (welche sprachenspezifisch sind).

•Qualitative Textanalysen:

- legen grösseres Gewicht auf die semantische und pragmatische Seite einer Textanalyse.
- grosses Gewicht der Hermeneutik (Auslegung des Textes, Verständnis)
- Beispiel einer qualitativen Textanalyse: Untersuchung zur Anpassung des Diskurses der sozialistischen Partei Griechenlands (PASOK) bis zur Machterlangung (Papadopoulos, 1989)

15.4.4. Ablauf einer Textanalyse:

- Fragestellung, Hypothese
 - z.B. Wandel der Werte Freiheit und Gleichheit in der Presse? (=keine Wenn-Dann-Hypothese)
- Grundgesamtheit und Stichprobe
 - z.B. Wahlprogramm und Stichprobe keine weil alle Wahlprogramme) → alle Leitartikel der 5 auflagenstärksten Zeitungen der letzten 30 Jahre, Stichprobe: jede achte Ausgabe.
- Festlegung der Analyseeinheit
 - Text als ganzes, Quasi-Sätze, Wörter → einzelne Wörter und Wortverbindungen, die etwas mit Freiheit zu tun haben
- Konstruktion des Kategoriensystems
 - zuordnen von Sätzen zu Kategorien → entspricht der Operationalisierung der Variablen

- Kodierbogen und Kodieranweisungen
 - Wie sollen Quasi-Sätze einer Kategorie zugeordnet werden? Der Kodierer benutzt das dann. Texte werden im Kodierbogen erfasst.
- Schulung der Kodierer
 - damit alle für den gleichen Text möglichst das gleiche Resultat liefern)
- Pretest, Prüfung der Reliabilität
 - gleiches Wahlprogramm mit gleicher Kodierungsanweisung muss bei verschiedenen Kodierern zum gleichen Resultat führen), eventuell Revision von Kategoriensystem und Kodierregeln
- Haupterhebung
- Datenübertragung und Datenauswertung

15.4.5. Probleme der Textanalyse

Wie alle wissenschaftlichen Vorgehensweisen sollten auch Textanalysen nachvollziehbar sein:

- computergestützten Textanalysen →ok
- systematischen Textanalysen, von Kodierern durchgeführt werden → Nachvollziehbarkeit → in Reliabilität der Messung wieder → Wichtigkeit eines eindeutigen Kodierschemas und Objektivität
- qualitativen Textanalysen → detaillierte Dokumentation der Vorgehensweise unabdingbar

Systematische, und computergestützte Textanalysen

- eingeschränkt in Analyse der Semantik +Pragmatik.
- Validitätsproblem: zeigen die vermuteten Ursachen ihre beobachtbaren Folgen auch in den Texten? Textanalysen zwar nicht reaktiv, aber Text immer in bestimmten Kontext.

15.4.6. Vorteile

- gut geeignet für Erforschung von sozialen / kulturellen Werten + deren Wandel.
- Texte z.T. über sehr lange Zeit noch vorhanden (NZZ seit 1780) → lange zurückliegende Materialien auswerten, bei Befragung geht das nicht.
- Nichtreaktiv.

15.5. andere Daten

15.5.1. Allgemein

grosse Anzahl zusätzlicher Methoden der Datenerhebung → suchen oft Inspiration bei nicht-reaktiven Messmethoden und experimentellen Untersuchungsanordnungen (oft Experimente oder Feldexperimente).

15.5.2. Prozessproduzierte Daten

- Daten unabhängig von spezifischen Forschungsabsichten erhoben (z.B. öffentliche Register (Einbürgerungen), Bevölkerungszählungen)
- Es ist ein sozialer Prozess, der die Daten produziert → Kontext der Erhebung nie unsere Forschungsfrage → Sekundäranalysen
- Notwendig: Validität (und Reliabilität) der Messung beachten Qualität prozessorientierter Daten ist oft schlecht (Kriminalitätsstatistik: nur angezeigte Delikte, Steuerdaten: Einkommen nach unten verzerrt, Arbeitslosenstatistik: nur gemeldet Leute).

15.5.3. Beobachtung

Grundsatz: jede Empirie ist Beobachtung z.B. Ablesen eines Messgerätes

In Sozialforschung → **Beobachtung spezifisch direkte Beobachtung menschlicher Handlungen** (inkl. Merkmale wie z.B. Kleidung).

• Teilnehmende - nichtteilnehmende Beobachtung

nimmt die beobachtende Person am Ereignis teil oder nicht - die Beobachteten wissen davon meist nichts. Teilnehmende Beobachtung wurde wissenschaftlich bereits von der "Chicago-Schule" in den 20er Jahren benutzt.

Beispiele für nicht-reaktive Beobachtung

- Bsp.: Günter Walraff - teilnehmende Beobachtung als angeblicher Journalist
- Bsp.: Teilnahme an Anti-WEF-Demo zur Beobachtung der Demo.

Problem Teilnehmende Beobachtung: zu starke Identifikation des Beobachtungsgegenstandes + Beeinflussung der Geschehnisse.

Nicht-Teilnehmend: man muss nicht agieren nur beobachten (protokollieren)

• Offene - verdeckte Beobachtung

- Problem verdeckt: ethisch
- Nicht-reaktiv → Beobachtete nicht ihr Verhalten ändern → weil wissen, dass beobachtet einzig verdeckt-nichtteilnehmend ist absolut einflusslose Beobachtung.

•Feld-, Laborbeobachtung

- Labor: Stimulus genau vorgegeben → Störfaktoren kontrolliert
- häufig experimentelles Design → hohe interne Validität.
- Feld → hohe externe Validität , gewisse Dinge im Labor nicht testen (z.B. langfristige Wirkungen)
- Feldexperiment = Kombination Feldbeobachtung und Experiment.

•Strukturierte - unstrukturierte Beobachtung

- Strukturierung soll Objektivität erhöhen → Beobachten immer selektiv → Problem der Verzerrung
- "unstrukturiert" bis "hochstrukturiert" → Mitte Beobachtungsleitfaden.
- Strukturierte Beobachtungen testen → überhaupt genügend Zeit um alles zu erfassen?
- Verhaltensbeobachtungen oft höherer Validität als Befragungen → wenn möglich Beobachtung statt Befragung
-

•Fremd-, Selbstbeobachtung

Erfinder des LSD (Hoffmann) → verschiedene Substanzen im Selbstversuch getestet und e Wirkung dokumentiert. Ist aber nicht intersubjektiv überprüfbar.

15.5.4. Unterschied Sozialreportagen und Beobachtungsstudien

bei teilnehmender Beobachtungen:

- Theorie geleitet (Hypothesen)
- Systematisches Vorgehen

beides notwendig → Studie nachvollziehbar

Ausserdem schränkt systematisches Beobachten die Gefahr selektiver Wahrnehmung ein (z.B. nur hypothesen-bestätigende Dinge wahrnehmen).

Beispiel für systematische Beobachtung: Die Arbeitslosen von Marienthal. Walraffs Beobachtungen waren nicht systematisch → Sozialreportagen

15.5.5. Nicht-reaktive Methoden

Problem: Verhalten mit wenig beeinflussbar → reaktive Methoden → Verhalten, welches die zu untersuchende Hypothese bestätigt, möglicherweise selbst erzeugen (=methodischer Artefakt). Reaktiv = Messvorgang beeinflusst Messergebnis.

Es gibt verschiedene Abhilfen:

- Versuchsleiter 1) kennt Hypothese nicht 2) weiss nicht, welches Kontrollgruppe und Experimentalgruppen sind (→ Doppelblindversuch: auch Versuchsperson weiss beides nicht). Problem: Leute bilden Annahmen über Hypothese.

Grundlegende Alternativen sind nicht-reaktive Methoden (heisst aber nicht unverzerrt).

- Feldexperimente (Forscher beeinflusst eine Variable aktiv)
- Unaufdringliche Beobachtung
- Verhaltensspuren (Forscher ist passiv, nimmt ohnehin produzierte Verhaltensäusserungen auf)
 - z.B. Anzahl Pizza-Bestellungen im Weissen Haus → nehmen vor einem Kriegsausbruch zu.
 - Bei historischen Ereignissen gibt es oft nur Verhaltensspuren.
 - z.B. aus dem Abfall auf Ernährungsgewohnheiten schliessen.
 - Prinzip: Verhaltensweisen liefern indirekte Hinweise auf Ausmass / Qualität sozialer Aktivitäten.

Das Internet schuf zusätzliche Quellen von (digitalen) Verhaltensspuren.

- Sekundäranalysen prozessproduzierter Daten
- Inhaltsanalysen

Probleme: Validität? Reliabilität? Ethisch? Wie kommt Stichprobe zustande?

Feldexperimente

- experimentelle Untersuchungen in natürlicher Umgebung
- Falls nichtreaktiver Erhebungsmethode → nichtreaktives Feldexperiment + Randomisierung ist oft gut möglich = interne und externe Validität
- Aber: ob gemessener Indikator wirklich das Konzept darstellt, ist oft nicht klar.

Beispiele:

- UniZH Spendenverhalten in Abhängigkeit zu Informationen die man den Studenten gibt (siehe frühere Vorlesung)
- Einfluss der Medien auf die öffentliche Meinungen (Iyengar und Kinder (1987): News that matters)
- Einfluss von Wahlmobilisierungen auf Stimmbeteiligung (Gerber und Green (2000))
- Einfluss der Ethnizität auf die Zurverfügungstellung von öffentlichen Gütern (Weinstein, Habyarimana, Humphreys und Posner (2006))
- Hypothese: Menschen in guter Stimmung helfen anderen eher als Menschen in schlechter Stimmung: Experiment: man liess Leute eine Münze finden, danach liess jemand in der Nähe Akten fallen ☒ Münzfinder helfen häufiger. Kann man nicht durch Befragung testen.
- experimentelle Briefe: Beschwerde an Behörde, Inhalt gleich, nur Absender unterschiedlich (Mann, Frau, Dr. med.). Unterscheiden sich Behördereaktionen je nach Absender?
- sozial erwünschtes Verhalten ist durch Befragung kaum zu erheben (wird nach oben verzerrt). Beispielsweise Umweltverhalten muss daher beobachtet werden, wodurch das Resultat einer (reaktiven) Befragung auf Validität überprüft werden kann.

16. Was man wählt, hängt von Frage und Gegenstand ab.

Marienthal-Studie zu Arbeitslosigkeit:

- Grundfrage war, ob AL zu Radikalisierung oder zu Apathie der Betroffenen führt.
- setzte viele verschiedene Methoden ein
- Beobachtungen erfolgten in natürlicher Situation (=Feldbeobachtung)
- Forscher nahmen manchmal an Interaktion teil (teilnehmende Beobachtung)
- Messen der Gehgeschwindigkeit = nicht-teilnehmend, nicht-reaktiv, quantitativ, verdeckt, strukturiert
- Es wurden auch "normale" Methoden (Fragebogen) eingesetzt

Multi-Methoden-Mix brachte bestätigende Ergebnisse mit unterschiedlichen Erhebung → Verstärkt das Gewicht der Aussage.