

Einstieg/Überblick

Paradigmen

Werturteile/Ethik

Forschungslogik

Hypothesen

Forschungsdesign

**Messung**

Standardisierte

Befragung

Qualitative

Befragung

Beobachtung

Nicht-reaktive

Methoden

Inhaltsanalyse

Stichproben

Herzlich willkommen zur  
Vorlesung

Methoden der empirischen  
Sozialforschung I

Vom Begriff zur Messung

## Begriff – Operationalisierung – Messung

- Zur Überprüfung von Hypothesen müssen die in ihnen verwendeten Begriffe empirisch „erfassbar“ gemacht werden. Dieser Vorgang heißt allgemein „Operationalisierung“.
- Ziel der Datenerhebung ist letztlich, den beobachteten Phänomenen Symbole (meist: Zahlenwerte) zuzuordnen; diese liegen der Datenauswertung zu Grunde. Dieser Vorgang heißt Messung.
- Vor der Operationalisierung und Messung müssen Begriffe definiert und u.U. in ihrem dimensionalen Gehalt geklärt werden (Konzeptspezifikation)
- In der qualitativen Forschung spielt dies Alles keine Rolle!

## Operationalisierung als wichtige Aufgabe

Ein häufiger „Anfänger-Fehler“ besteht darin, das Forschungsproblem nicht in eigene Operationalisierung zu „übersetzen“.

### Beispiel: Leistungsdruck in der Schule.

Hypothese (deskriptiv) : In der Schule (konkret: Oberstufe) herrscht so hoher Leistungsdruck, dass den SchülerInnen nicht ausreichend Zeit für Freizeitaktivitäten bleibt

Nicht: „Die Schule setzt mich häufig unter Leistungsdruck“  
„Die Schule behindert meine Freizeitgestaltung“

Sondern: Schulanforderungen möglichst genau messen  
Freizeitverhalten und Freizeitwünsche messen

## Begriff/Konstrukt/Konzept – Indikator

Häufig ist das Merkmal, welches wir erfassen wollen, nicht direkt beobachtbar. Wir sprechen hier von Konzept oder Konstrukt. Die empirische beobachtbaren Größen, anhand derer es erfasst werden soll, heißen Indikatoren.

### Beispiel: Armut

Arm sind Personen, die „... über so geringe (materielle, kulturelle und soziale) Mittel verfügen, dass sie von der Lebensweise ausgeschlossen sind, die in dem Mitgliedsstaat, in dem sie leben, als Minimum annehmbar ist“ (Definition der EU)

Indikatoren: Einkommen? Ausstattung mit Konsumgütern?  
Lebenschancen (Bildung, Lebenserwartung)? Partizipation?

# Begriff/Konstrukt/Konzept – Indikator

## Weiteres Beispiel: Arbeitslosenquote

### Zähler

- Alle beim Arbeitsamt als arbeitslos gemeldeten Personen?
- Alle Arbeitssuchenden, die keinerlei Erwerbstätigkeit nachgehen?
- Arbeitslos Gemeldete und „Stille Reserve“?

### Nenner

- Alle Erwerbspersonen?
- Alle zivilen Erwerbspersonen?

## Konzeptspezifikation

Viele Konstrukte sind mehrdimensional.

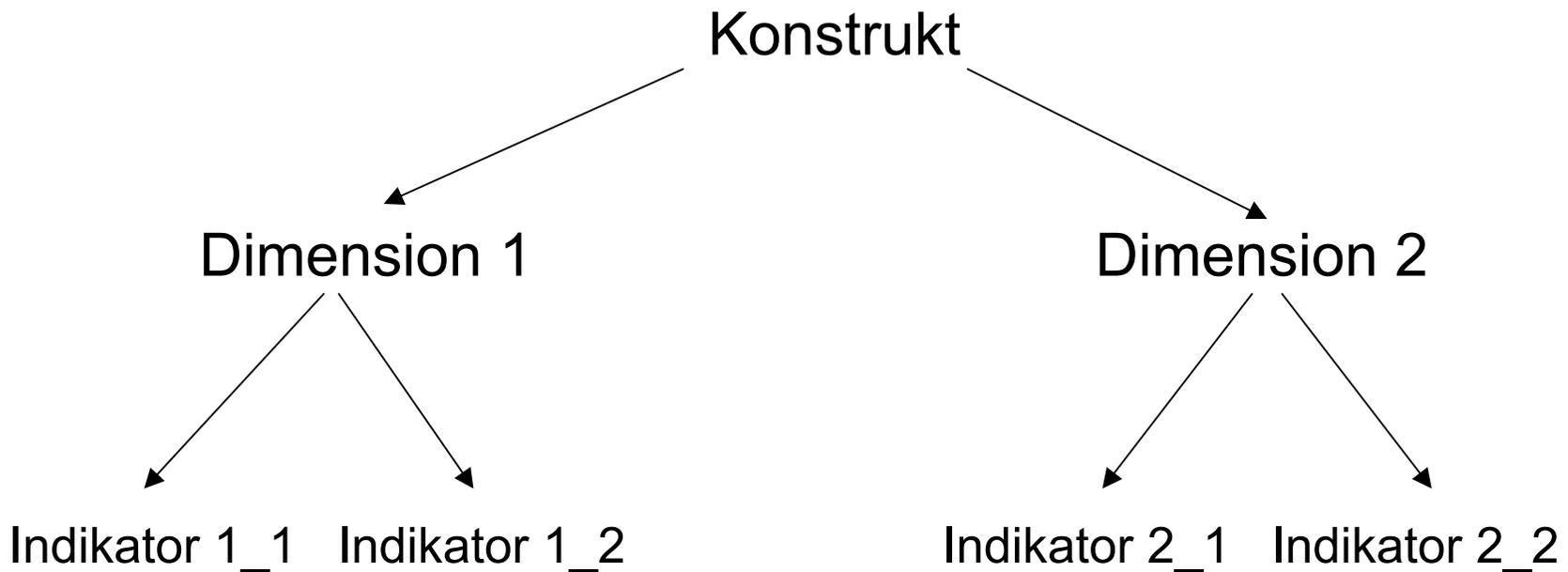
Sie beinhalten unterschiedliche, wenn auch zusammengehörige Dimensionen.

Beispiel: Armut (s.o.) – geringe materielle, kulturelle, soziale Mittel

Beispiel: Autoritarismus

- Neigung zu Unterwürfigkeit und Anpassung
- Feindseligkeit gegen Minderheiten und Schwächere
- Eintreten für „Ruhe und Ordnung“

# Konzept – Dimension – Indikator



## Messung mit mehreren Indikatoren

Es ist im Allgemeinen empfehlenswert, Merkmale durch mehrere Indikatoren zu messen – jedenfalls wenn Messfehler zufällig sind.

Solche Messungen führen (wenn korrekt durchgeführt) zu genaueren Messergebnissen.

Negativbeispiel: Prüfung, Klausur mit nur einer einzigen Frage (aus breitem Stoff).

Mehrere Indikatoren müssen natürlich auch herangezogen werden, wenn das Konstrukt mehrdimensional ist (am besten: für jede Dimension wiederum mehrere Indikatoren)

# Messung mit mehreren Indikatoren: Beispiel

[http://www.polwiss.fu-berlin/osz/dokumente/for\\_rechts.htm](http://www.polwiss.fu-berlin/osz/dokumente/for_rechts.htm)

## Rechtsextremismus-Skala (Auszüge)

- (Autoritarismus:) Wer seine Kinder zu anständigen Bürgern erziehen will, muss von ihnen vor allem Gehorsam und Disziplin verlangen.
- (Nationalismus:) Deutschland sollte wieder eine führende Rolle in der Welt übernehmen.
- (Ethnisch motivierte Fremdenfeindlichkeit:) Ausländer sollten so schnell wie möglich Deutschland verlassen.
- (Pronazistische Einstellungen:) Ohne Judenvernichtung würde man Hitler heute als großen Staatsmann ansehen.

## Messung mit mehreren Indikatoren: Index

Unter einem Index versteht man eine Messung anhand mehrerer Indikatoren, die jedoch manchmal relativ beliebig zusammengefügt werden.

Beispiel Schicht-Index nach Erwin K. Scheuch/Hans-Jürgen Daheim

- Einkommen, Beruf und Bildung (SHE, Kap. 4.4.1.2)

Gelegentlich wird unter Index auch eine einfache Messgröße verstanden (z.B. Big-Mac-Index, siehe Diekmann).

## Messung mit mehreren Indikatoren: Skala

Unter einer Skala versteht man eine Messung anhand mehrerer Indikatoren, wobei spezifische Annahmen über die Struktur der Beobachtungen gemacht werden.

Diese Struktur sollte sich empirisch prüfen lassen.

Nachfolgend einige Beispiele hierzu.

## Skalierungsverfahren I: Guttman-Skala

Eine Guttman-Skala besteht aus mehreren Items, wobei jedes Item eine „stärkere“ Ausprägung des jeweiligen Merkmals misst als das vorherige.

### Beispiel: (Konventionelle) politische Beteiligung

- Ich werde zur Wahl gehen *ja / nein*
- Ich werde Geld für den Wahlkampf spenden *ja / nein*
- Ich werde im Wahlkampf für eine Partei aktiv werben  
(Infostände, Verteilen von Materialien, usw.) *ja / nein*
- Ich werde für ein Mandat oder Amt kandidieren *ja / nein*

## Skalierungsverfahren II: Likert -Skala

Bei einer Likert-Skala werden mehrere Aussagen („Items“, „Statements“) vorgegeben; die Befragten geben auf einer mehrstufigen Antwortvorgabe (meist fünf bis sieben Stufen) den Grad ihrer Zustimmung an.

Beispiel: Siehe oben Skala „Rechtsextremismus“ (dort aber oft nur ein Item pro Dimension; besser: mehrere)

Beispiel: Studierendenbefragung im vergangenen Semester

Die Likert-Skalierung ist das häufigste Verfahren in den Sozialwissenschaften

## Messniveaus (Skalenniveaus)

- **Nominalskala:** Unterschiedliche Ausprägungen bedeuten Unterschiedlichkeit – sonst nichts (Bsp.: Parteien).
- **Ordinalskala:** Unterschiedliche Ausprägungen bedeuten Reihenfolge – aber keine Angaben über Größe der Abstände (Bsp.: Olympiamedaillen, allgemein: Rangfolgen. Auch: Guttman-Skala, Likert-Skala [letztere oft wie intervallskaliert behandelt]).
- **Intervallskala:** Die Differenzen der Messwerte sind aussagekräftig (oder sollen es sein) – Bsp. Temperatur in °C.
- **Ratioskala:** Die Verhältnisse der Messwerte sind aussagekräftig – Bsp. Temperatur in °K; evtl. Einkommen.

## Gütekriterien I: Validität

Unter V. versteht man die **Gültigkeit** einer Messung – anders gesagt: dass gemessen wird, was gemessen werden soll.

Verfahren zur Prüfung:

- Inhaltsvalidität (keine formale Prüfung – trotzdem wichtig)
- Kriteriumsvalidität: Bezug auf Außenkriterium
  - Übereinstimmungsvalidität („Concurrent validity“): Außenkriterium wird gleichzeitig gemessen (Bsp.: Autoritarismusgrad bei Nazis)
  - Vorhersagevalidität: Außenkriterium wird später gemessen (Bsp.: Studienerfolg nach Aufnahmeprüfung)

## Gütekriterien I: Validität, hier: Konstruktvalidität

- Aufgrund vorliegender Theorien wird Zusammenhang zwischen verschiedenen Konstrukten angenommen und geprüft.
- Bestätigung des angenommenen Zusammenhangs spricht für (nicht: beweist!) Validität.
- Variante: Multi-Trait-Multi-Method Matrix
  - Das gleiche Konstrukt, mit verschiedenen Methoden gemessen: Starker Zusammenhang (Konvergenz)
  - Verschiedene Konstrukte: Schwacher oder kein Zusammenhang (Diskriminanz)
  - Verschiedene Messungen des gleichen Konstruktes zeigen ähnliche Zusammenhänge mit anderen Merkmalen

## Gütekriterien II: Reliabilität

Reliabilität heißt **Zuverlässigkeit**.

Messinstrumente sollen bei wiederholter Messung am (unveränderten!) Objekt immer den gleichen „Messwert“ liefern.

Formal: Zusammenhang zwischen beobachteten und wahren Werten. Da man wahre Werte nicht erheben kann, behilft man sich mit

- Test-Retest-Methode
- Paralleltest-Methode
- Prüfung interner Konsistenz

## Gütekriterien III: Objektivität

Objektivität bezieht sich auf Messung durch Dritte und meint **Unabhängigkeit der Messung von der Person**, die sie durchführt.

Beispiele: Beurteilung von SchülerInnen durch LehrerInnen, PatientInnen durch ÄrztInnen, Medieninhalten durch CodiererInnen.

Die Objektivität wird durch die Übereinstimmung der Urteile geprüft (bitte für später merken: die Korrelation der Urteile – wie bei Diekmann zu lesen – ist hierfür nicht geeignet, siehe ILMES: Interraterreliabilität, ICC).