

ERKENNTNISGEWINNUNG & sonstiges

empirische Wissenschaft

- > entwirft Hypothesen&Theorien und konfrontiert diese mit der Realität
- > Vergleich von gedanklichen Antworten auf Fragen mit den in der Realität vorfindbaren Sachverhalten

Ziel wissenschaftlicher Tätigkeit

- >Generieren von beständigem Wissen

Alltagspsychologie

- > Überzeugungsstrategie
 - man gibt sich bei Beantwortung einer Frage sehr überzeugt, obwohl man die Antwort nicht wirklich kennt (andere glauben das)
- > Berufung auf Autoritäten
 - man gibt einen berühmten Wissenschaftler als Quelle seiner Aussage an, um diese glaubhafter zu machen (Problem: aus dem Zusammenhang gerissen / auch Autoritäten können falsch liegen)
- > individuelle Erfahrung
 - aus eigenen Erfahrungen werden falsche Tatsachen abgeleitet
 - unsere „Datenbasis“ ist lückenhaft und die Schlussfolgerungen deshalb falsch

Psychologie als Wissenschaft

- > Problemlösen
 - „ein gegebener Ausgangszustand (vorläuf. Antw.) soll in einen Zielzustand (geprüfte Antw.) überführt werden, wobei eine Barriere diese Überprüfung erschwert“
 - wiss. Methoden benötigt, um zuverlässige und gültige Ergebnisse zu generieren
- > psychologische Methoden
 - Methoden, mit deren Hilfe wir Antworten auf Fragen aus der Psychologie erhalten, um langfristig gültige Antworten zu finden
- > Erkenntnis durch Sicherung und Prüfung
- > abstrakte und allgemeine Formulierungen

Induktion

- > ausgehend von Einzelbeobachtungen/ Einzelfallanalysen wird auf eine generelle Regel geschlossen
- > Schlüsse haben Wahrscheinlichkeitscharakter
- > Grundlage für andere Methoden

Deduktion

- > aus allg. Zusammenhängen (Theorien) werden weitere Hypothesen logisch abgeleitet
- > wenn Theorie stimmt, muss auch Hypothese

Falsifikationismus nach Popper (kritischer Rationalismus)

- > Beweis einer Theorie durch endliche Anzahl von Beobachtungen nicht möglich
- > Theorien=vorläufige, spekulative Vermutungen
- > Theorien müssen falsifizierbar sein, also müssen sie empirisch prüfbar sein(Erfahrbarkeit)

hypothetiko-deduktiver Ansatz

- > 1) aus Beobachtungen/Intuition entsteht eine Theorie (Induktion)
- > 2) aus der Theorie werden Hypothesen abgeleitet > empirisch getestet > neue Ergebnisse

entstehen (Deduktion)

> beide Formen kombiniert als zykl. Prozess

Qualitative Forschung

> sinnverstehende, unstandardisierte Verfahren

> oft zur Generierung von Forschungsfragen (Induktion), selten zur Hypothesenprüfung

> deskriptive Feldforschung, biografische Methode, gegenstandsbezogene Theoriebildung

> unstandardisierte Interviews, Gruppendiskussionen, teilnehmendes Beobachten

> Inhaltsanalyse, Hermeneutik, Diskursanalyse

Vorteile:

> subj. Erfahrung des Untersuchungsobjekts

> unerwartete Einsicht durch offene Ergebniserwartung

> erlegt Teilnehmern keine spezielle Sichtweise auf

Nachteile:

> Validität und Reliabilität der Daten problematisch

> Generalisieren und Vorhersagen nicht mögl.

> fehlende Replizierbarkeit

> geringe Effizienz durch hohe Kosten und Zeitaufwand bei größeren Stichproben

Quantitative Forschung

> objektiv gemessene, standardisierte Verfahren

> oft zur Hypothesenprüfung (Deduktion)

> Ziel: allgemeingültige Aussagen über Zusammenhänge usw.

> Experimente, Quasiexperimente, Korrelationsstudien, Metaanalyse..

> Beobachten, Zählen, Testen,...

> beschreibende/schlussfolgernde Methoden, multivariate Methoden, Modelltests

Vorteile:

> hohe Präzision & Kontrolle

> Effizienz (große Stichproben umsetzbar)

> Anspruch auf Kausalaussagen

> prädiktive Kraft

Nachteile:

> Übervereinfachung der komplexen menschlichen Natur

> Vernachlässigung der subjektiven Natur der Psychologie

> Vernachlässigung der Individualität und Autonomie der VPn

Schrittweise Erkenntnisgewinnung

1) anfängl. Beobachtung/Frage

2) Hypothenbildung

3) Konzeption der Untersuchung (Design)

4) Datenauswertung und Schlussfolgerungen

5) Veröffentlichung der Ergebnisse

6) Diskussion der ungelösten Fragen

7) Lösung offener Fragen

Themenbereiche der Psychologie

> klinische, biologische, emperimentelle, kognitive, Persönlichkeit, Sozial, Entwicklung, Arbeit-Organisation-Wirtschaft, pädagogische, forensische, Sport

Angemessenheit wissenschaftlicher Aussagen

> logische Konsistenz

> Fundierung auf Erfahrung

- > Überprüfbarkeit der Beobachtungen
- > intersubjektiv nachprüfbar

Wissenschaftstheorie

- > Wissenschaftsbereich, der sich mit Voraussetzungen, Methoden und Zielen von Wissenschaft und ihrer Form der Erkenntnisgewinnung beschäftigt

Methodologie

- > Theorie wissenschaftlicher Methoden

Entwicklung von Wissenschaft

- > Theologie > Metaphysik > Wissenschaft

Idealismus

- > materielle Dinge existieren nur in sofern, als sie an „Ideen“ teilhaben – ideale Wirklichkeit

Realismus

- > es gibt ein vom Individuum unabhängiges Sein – Erkenntnis ist möglich

Rationalismus

- > die Welt ist nach vernunftmäßigen Prinzipien aufgebaut, Erkenntnis weitgehend unabhängig von Erfahrung möglich

Empirismus

- > es gibt eine externe, unmittelbar über Erfahrung zugängliche Wirklichkeit
- > am Anfang aller Erkenntnis steht Erfahrung

Wissenschaftsarten

- Formalwissenschaft
- > Mathematik, Logik
- **nomothetische** Realwissenschaft
- > Erforschung von allgemeingültigen Gesetzmäßigkeiten
- > experimenteller Ansatz der Psychologie
- **ideografische** Realwissenschaft
- > Untersuchungen, die einmalige Ereignisse und Sachverhalte beschreiben
- > korrelativer Ansatz der Psychologie
- ideografische Geisteswissenschaft
- > auch Teil der Psychologie

Bayesianismus

- > man kann Annahmen weder durch Gegenbeispiele widerlegen, noch durch pos. Beispiele beweisen
- > keine Falsifizierung / Verifizierung möglich
- > spezielle Prüfkriterien (Signifikanz)
- > die Wahrscheinlichkeit wird nach ihrer Prüfung entsprechend angepasst und kann zur Vorhersage verwendet werden

Basisziele der Psychologie

- > **Beschreiben**
- > alle Begriffe in Fragestellung definieren
- > benennen, ordnen, klassifizieren, definieren, auszählen (Häufigkeiten), messen (Ausprägung)
- > Operationalisierung der Merkmale
- > keine Kausalaussagen, höchstens Assoziationen (Korrelation)
- > deskriptive Statistik
- > **Erklären**

- > Richtung von Zusammenhängen angeben (Kausalrelation: UV bedingt AV)
- > positive/negative Zusammenhänge, Stärke der Zusammenhänge angeben
- > **Vorhersagen**
- > Prognosen = vorwärts gerichtete Erklärungen, die dazu dienen, das Eintreten eines zukünftigen Sachverhalts vorherzusagen
- > UV = **Prädiktor**; AV = **Kriterium** (Regression)
- > mehrere, präzise Prädiktoren benötigt, um genau vorherzusagen (Komplexität)
- > Prognosemodell: Prädiktoren + Gewichtung
- > neue Studien können Prognosemodell überprüfen
- > **statistisches Prognosemodell**: auf Stichprobe bezogen
- > **individuelle Prognose**: je höher der stat. Prognosewert, desto wahrscheinlicher, dass korrekt
- > Prognosezeitraum: je länger, desto ungenauer
- > **Verändern**
- **Korrektur**
- > Ausgangszustand: problematisch
- > Schwerpunkt: Negatives aufheben (Therapie)
- **Förderung**
- > Ausgangszustand: normal
- > Schwerpunkt: Optimierung (Unterricht)
- **Prävention**
- > Ausgangszustand: normal – gefährdet
- > Schwerpunkt: Schlechtes vermeiden
(Alkoholpräventionsaktionen in Schule)

Theorie

- > Annahmen, Aussagen, Definitionen, die Sachverhalte beschreiben
- > Zusammenstellung von Wissen (Überlegung/Erfahrung)
- > ständiger Überprüfung ausgesetzt → kann nie wahr sein, sondern nur empirisch belegt

Bewertungskriterien quantitative Daten

>>Objektivität

- > intersubjektive Übereinstimmung/Nachprüfbarkeit
- > theoretische und empirische Nachvollziehbarkeit
- > schlecht: unterschiedliche Versuchsbedingungen / Versuchsleitereinflüsse; keine Festlegung für Auswertung,..
- > Voraussetzung für Reliabilität

Durchführungsobjektivität

- > Güte der Datenerhebungsverfahren
- > Daten müssen unabhängig vom Forscher zustandekommen
- > Versuchsleitertraining, Leitfäden, einheitliche Instruktionen

Auswertungsobjektivität

- > gleiche Daten = gleiche Testergebnisse unabhängig vom Forscher
- > Auswertungsregeln – Vorgehen genau beschreiben und festlegen

Interpretationsobjektivität

- > Interpretation und Bewertung der Ergebnisse unabhängig vom Forscher
- > Vergleichsmaßstab für Ergebnisse vorher explizit festlegen

>>Reliabilität

- > Zuverlässigkeit, Beständigkeit der Untersuchung
- > Voraussetzung für Validität

Test-Retest-Reliabilität

- > Stabilität - Messgenauigkeit
- > Ausmaß der Übereinstimmung bei wiederholter Messung an gleicher Stichprobe (durch Korrelation gemessen)

>Schwankungen

- > Probanden: Tagesform und Erinnerungseffekt
- > Messinstrumente: Fehler bei Bedienung
- > Auswertung: Fehler bei Durchführung
- trotz Standardisierung wird es immer Fehlerquellen geben, bei reliabler Messung sollten sie sich jedoch in Grenzen halten
- besser: Messwiederholung und Werte mitteln > Fehler mitteln sich raus, wenn nicht systematisch

Interrater-Reliabilität

- > Höhe der Übereinstimmung der Einschätzungsergebnisse unterschiedlicher Beobachter/Testanwender (Rater)
- > hoch, wenn verschiedene Rater gleiche Vpn relativ gleich bewerten
- > durch Beurteilerschulungen erhöhen

Paralleltest-Reliabilität

- > gibt an, inwieweit Ergebnisse von zwei vergleichbaren/parallelen Testversionen miteinander korrelieren, welche an der gleichen Personenstichprobe zum selben Zeitpunkte erhoben werden (Testergebnisse werden verglichen (korreliert))
- > prüft Verlässlichkeit und Genauigkeit des Erhebungsverfahrens

Interne Konsistenz

- > Überprüfung der Verlässlichkeit und Genauigkeit eines Verfahrens
- > eine Stichprobe wird mit einem Instrument zu einem Zeitpunkt untersucht
- > danach vergleicht man einzelne Teile des Tests miteinander
- 1) **Split-half-Reliabilität:** Test in zwei Hälften teilen und miteinander korrelieren
- 2) **Cronbachs Alpha:** Test in so viele Teile zerlegen, wie er Items hat und korrelieren

>>Validität

- > misst der Test, was er messen soll?
- > Ziel: Bedeutungskern einer Variablen beim Operationalisieren erfassen
- > kein dichotomes, sondern kontinuierliches Merkmal

Konstruktvalidität

- > Relationen zu anderen Konstrukten
- 1) **konvergent:** Übereinstimmung mit Konstrukten, die das gleiche messen
 - Beispiel: zwei Tests zu Ängstlichkeit messen das gleiche
- 2) **diskriminant:** Abweichung von Konstrukten, die etwas anderes messen
 - Beispiel: Intelligenztest korreliert nicht mit Frustrationstest

Kriteriumsvalidität/externe/ökologische Validität

- 1) **simultan (concurrent):** Vergleich mit Realität zum selben Zeitpunkt
 - Einsamkeitsfragebogen erfasst aktuell einsame Personen
- 2) **Vorhersagevalidität (predictive):** Vorhersage des Kriteriums
 - Berufseignungstest sagt Berufserfolg voraus

Inhaltsvalidität

- > Methode erfasst Konstrukt in seinen wichtigsten Aspekten (subjektiv eingeschätzt)

> „face-validity“

interne Validität

> Störvariablen in Versuch kontrollieren

unter „Experiment“:

interne, externe (Populations-, Situations-, Variablenvalidität)

unter „MGV“

normative, soziale Validität

>> Spezialformen, deshalb nicht hier

Bewertungskriterien für qualitative Forschung

Verfahrensdokumentation

> alle Schritte, Instrumente... nachvollziehbar dokumentieren

argumentative Interpretationsabsicherung

> Interpretation nachvollziehbar begründen

Regelgeleitetheit

> geplante Analyseschritte dürfen modifiziert werden, aber systematisch

Nähe zum Gegenstand

> Personen sollten in natürlicher Umwelt befragt werden > Achtung: Verzerrungen

kommunikative Validierung

> VP Ergebnisse zeigen: Zustimmung = Absicherung

Triangulation

> unterschiedliche Lösungswege zur gleichen Frage entwerfen

> muss nicht komplett übereinstimmen – Gesamtbild erhalten

Methoden außerhalb der Forschung

> Diagnostik

- Erkenntnisse > Urteil > Gutachten > Entscheidung über Maßnahmen

- oft qualitativ (Individuum)

> Intervention

- gezielte Maßnahmen zur Prävention/Rehabilitation

> Evaluation

- Intervention prüfen, verbessern > Optimierung

HYPOTHESEN

> vorläufige (vermutete) Antworten, die Forscher auf ihre Fragen geben, nennt man Hypothesen

> Hypothesen = Wahrscheinlichkeitsaussagen → Signifikanz

> Vermutungen so lange bis wissenschaftlicher Beweis erbracht ist

Entstehung

> Ungereimheit in Befunden

> Anfrage eines Auftraggebers

> Beobachtungen im Alltag ohne Erklärung

> viele unterschiedliche Erklärungen in Literatur ...

→ erst verfügbares Wissen sammeln: vllt schon beantwortet

>>Kriterien

Formulierung

- > Art der Hypothese soll offensichtlich sein

Widerlegbarkeit

- > logische Konsistenz (Widerspruchsfreiheit)
- > Hintergrundwissen schon einbezogen (kein Widerspruch zu vorherigen Erkenntnissen)

Operationalisierbarkeit

- > abstrakte Begriffe definieren & beobachtbar / messbar machen

Begründbarkeit

- > Stand der Forschung berücksichtigen
- > wenn neues Gebiet: logisch nachvollziehbar

Falsifizierbarkeit

Konditionalsatz möglich

- > wenn dann oder je desto

Allgemeingültig

- > nicht nur auf eine Person bezogen

>>Hypothesenarten

universelle Hypothese

- > ohne jede Einschränkung formuliert (generelle Allgemeingültigkeit)
- > kann nur widerlegt werden (1 Gegenbeispiel genügt), kann nie bestätigt werden
- > „Intelligenz bedingt Änglichkeit“

beschränkt universelle Hypothese

- > raumzeitlich/personenbezogen eingeschränkt
- > größere WS nicht widerlegt zu werden (trotzdem reicht ein Gegenbeispiel)
- > nie entgültig zu bestätigen
- > „Intelligenz bedingt bei Erwachsenen Änglichkeit“

quasiuniverselle Hypothese

- > in Psychologie verwendet
- > Einschränkungen auf Wahrscheinlichkeit bezogen > Regehaftigkeiten finden
- > keine Verifikation/Falsifikation > immer nur für jeweilige Bedingungen
- > Verallgemeinerung durch Replikationen mit neuen Bedingungen
- > Prüfung durch Inferenzstatistik
- > „Intelligenz bedingt zumeist Ängstlichkeit.“

Zusammenhangshypothesen

- > zwischen 2 oder mehr Merkmalen besteht ein Zusammenhang

Unterschiedshypothesen

- > zwei oder mehr Populationen unterscheiden sich in einer oder mehreren Variablen

Veränderungshypothesen

- > durch einen Einfluss lässt sich eine Veränderung feststellen
 - **Kausalhypothesen**
 - > die UV bewirkt Veränderungen in der AV → verlangt nach Experiment

Einzelfallhypothesen

- > sind Vermutungen über individuell bedeutsame Merkmal (z.B. über unterschiedliche Testverläufe)

>>Untersuchungsarten

hypothesenprüfende Untersuchung/Deduktion

- > Hypothese schon vorher gebildet (a-priori)
- > Prüfsituation entscheidet über Annahme oder Zurückweisung
- > Ergebnis: geprüfte Aussage
- > oft quantitativ
- > Deduktion

hypothesengenerierende Untersuchung/Induktion

- > Forschungsgegenstand nicht sehr differenziert
- > Aufstellung einer begründeten Hypothese erschwert > Erkundungsversuch
- > Ergebnis: Grundlage für Hypothese (a-posteriori)
- > Hypothesen nie im selben Versuch generieren&prüfen!! (Unabhängigkeit)

Mischform

- > in einem Versuch wird eine Hypothese geprüft und gleichzeitig nach neuer Systematik gesucht

>> **Hypothesenebenen**

- > Umformungen resultieren in statistischer Prüfbarkeit

1) theoretisch-inhaltliche Hypothese (TIH)

- > Bsp: „Intelligente Personen reagieren zumeist ängstlich“

2) empirisch-inhaltliche Hypothese (EIH)

- > Bsp.: „Wenn Vpn höhere Werte im IQ-Test erzielen, dann erhalten sie auch höhere Werte im Änglichkeits-Fragebogen als weniger intelligente Vpn“
- > Operationalisierung der TIH

3) statistische Vorhersage (SV)

- > Parametrisierung: ein statistisches Maß finden, das zur Hypothese passt
- > Bsp: IQ-Test Ergebnis und Änglichkeitsfragebogen Ergebnis hängen positiv zusammen:
 $H_0: r(x,y) \leq 0$ / $H_1: r(x,y) > 0$

4) Testhypothesen (TH)

- > Signifikanzprüfung (beweisen, dass Unterschiede nicht nur auf Zufall beruhen und nicht eigentlich doch H_0 stimmt)

5) Rückschluss

- > Übersetzung der Bedeutung der Ergebnisse in die theoretisch-inhaltliche Hypothese
- > Achtung! Signifikanz ungleich Effektstärke
- Ereignisse zunächst nur für Stichprobe gültig

VARIABLEN

- > veränderliche Beobachtungsgrößen
- > mindestens 2, meist mehr Ausprägungen (Gegenteil: Konstante)
- > operationalisiertes Merkmal
- > Verhältnisse darstellbar durch Pfaddiagramme

>> **Unterscheidungen**

qualitative Variablen

- > Beschaffenheit (Qualität) mit endlicher Anzahl von Ausprägungen
- > Zuordnung von Zahlenwerten hat wenig-keine Bedeutung
- > Bsp: Augenfarbe (grün, blau..)
- > Nominalskala (auch Ordinalskala?)

quantitative Variablen

- > Unterscheidung nach Zahlenwerten
- > interpretierbar im Sinne der Intensität/Größe
- > systematische Abbildung auf Skala
- > Intervall-, Ratio-, Absolutskala

konkrete (manifeste) Variablen

> direkt beobachtbar (Reaktionszeit)

abstrakte (latente) Variablen

> nicht direkt beobachtbar, müssen operationalisiert werden (Intelligenz)

einfache Variablen

> ein (oder wenige) Bedeutungsaspekt (Reaktionszeit)

komplexe Variablen

> hohe Anzahl an Bedeutungsaspekten/ Facetten(Intelligenz)

unabhängige Variablen (UV)

> in wenn-dann-Hypothesen: wenn-Komponente

> Prädiktor, wenn Vorhersage

> Ursache oder Faktor, wenn ursächlicher Einfluss

abhängige Variablen (AV)

> in wenn-dann Hypothesen: dann-Komponente

> Folge oder Konsequenz der UV

> auch Kriterium genannt

intervenierende Variablen (IV)

> gleichzeitig Ursache und Wirkung

> auch: vermittelnde Variable, Mediator

> Bsp: UV: gewalttätige Computerspiele AV: Aggressivität IV: erhöhte Verfügbarkeit aggressionsbezogener Gedanken

Moderatorvariablen (MV)

> kann Störvariable sein, wenn nicht kontrolliert

> im oberen Beispiel: Alter würde Effektstärke beeinflussen

exogene Variablen

2 Definitionen:

> von außen kommend (Bsp: sozialer Kontext der Person) – Wirkung nach innen oder außen

> UV

endogene Variablen

2 Definitionen:

> von innen kommend – Wirkung nach innen oder außen (Bsp: psych.Störung ist dann endogen, wenn keine äußeren Einflüsse dafür erkennbar sind)

> AV

diskrete Variablen

> endlich viele (Parteien bei Wahl) oder abzählbare unendlich viele Ausprägungen (Versuche bis Aufgabe richtig gelöst wird)

- **dichotome Variablen** = nur zwei diskrete Ausprägungen

- **polytome Variablen** = mehr als zwei diskrete Ausprägungen

stetige Variablen

> unendlich viele Ausprägungen innerhalb eines Intervalls (Gewicht; Reaktionszeit,...)

> meist diskret gemessen (es gibt keine kleinste Einheit, es geht immer genauer)

organismische Variablen

> können nicht verändert werden (Geschlecht,...)

>>Operationalisierung von Variablen

> ein latentes Merkmal beobachtbar/messbar machen

> durch „Messbarkeit“ verlieren Variablen an semantischem Gehalt, wenn komplex

> Ziel: zentralen Aspekt finden, um dieses Problem einzudämmen

Beispiel:

theoretisch-inhaltliche-Hypothese (TIH)

„ Intelligente Personen reagieren zumeist ängstlich“

> empirisch-inhaltliche Hypothese (EIH)

„Wenn Vpn höhere Werte im IQ-Test erzielen, dann erhalten sie auch höhere Werte im Ängstlichkeits-Fragebogen als weniger intelligente Vpn.
→ Intelligenz wurde auf Test reduziert und Ängstlichkeit auf Fragebogen
→ jede TIH hat mehrere Operationalisierungsmöglichkeiten (EIH)

STICHPROBEN (&Population)

Population: Grundgesamtheit / Menge aller potenziellen Untersuchungsobjekte für gegebene Fragestellung

Stichprobe: Teilmenge aus einer Grundgesamtheit, die unter bestimmten Gesichtspunkt (Fragestellung) ausgewählt wurde

→ Ziel: Ergebnisse der Stichprobe auf Population übertragen können (Stichprobenziehung!)
→ Kriterium: Merkmalsadäquanz (optimale Repräsentation)

>>Arten von Stichproben

Zeitstichprobe (Beobachtung)

> Beobachtungen werden in festgelegtem Zeitabstand protokolliert
> breiterer Überblick über Gesamtverlauf

Ereignisstichprobe (Beobachtung)

> Ereignis definieren
> Häufigkeit, Art, Dauer protokollieren
> bei gezielter Hypothesenprüfung, aufeinanderfolgenden Ereignissen, seltenen Ereignissen

Quotenstichprobe

> Forscher gibt Merkmale/Kombinationen vor
> gezielte Auswahl (nicht zufällig) > Repräsentativität fraglich

angefallene Stichprobe

> aktuell leicht verfügbare Populationsmitglieder (oft Studenten)
> noch schlechter generalisierbar als Quotenstichprobe
> auch: ad-hoc, systematic oder Pseudozufallsstichprobe

wiederverwendete Stichproben

> Panels = feste Stichprobe wird wiederholt befragt
> Untersuchung von Veränderungen
> Einschränkung der Aussagekraft durch Veränderung des Antwortverhaltens (Reifung, Einfluss vergangener Testungen)

>> Zufallsstichproben

> am besten für Merkmalsadäquanz
> jede Person der Population hat gleiche Chance
> wesentliche Kriterien auf Population übertragbar (**Populationsvalidität**)
> Gesetz der großen Zahl: je größer, desto repräsentativer

einfache Zufallsstichprobe (random)

> jedes Populationsmitglied hat gleiche Chance gezogen zu werden → selten
> (Urnen-) Ziehung → zufällige Entnahme von Einheiten
> nicht automatisch repräsentativ
> z.B. Personen aus Telefonbuch zufällig auswählen

> Schritte:

- 1) **Zielgesamtheit** (über sie soll die Aussage getroffen werden)
- 2) **Auswahlgesamtheit** (Personen, die eine Chance haben, in Stichprobe zu kommen)
- 3) Befragte nach Zufallsprinzip wählen

Probleme:

> **overcoverage** = Elemente, die es in der Zielgesamtheit nicht gibt, in der Auswahlgesamtheit schon

> **undercoverage** = Elemente, die in der Zielgesamtheit gibt, in der Auswahlgesamtheit aber nicht
> fehlende Werte

> Zufallsfehler bei Zufallsauswahl

geschichtete Zufallsstichprobe (stratified)

> Untergruppen der Population sind auch in Stichprobe repräsentativ vertreten

> bei Merkmalen in ausgefallenen Kombinationen

> Schritte:

1) Grundgesamtheit in Untergruppen aufteilen

2) für Stichprobe: gleiche Untergruppen, ähnliche Relationen

3) aus jeder Untergruppe einfache Zufallsstichprobe ziehen

> Bsp: 8 Klässler → Hauptschule/Realschule/Gymnasium

mehrstufige Zufallsstichprobe

> bei Schwierigkeiten direkt eine Stichprobe zu ziehen (fehlende Listen, weiträumige Verteilung)

> Schritte:

1) Klumpenstichprobe mit großen Klumpen

2) aus diesen Klumpen werden durch Zufallszahlen einige Klumpen ausgewählt

3) aus allen Personen in diesen Klumpen, wird eine Zufallsstichprobe gezogen

> Bsp: Studentenwohnheime in Berlin > einige Wohnheime ziehen (Klumpenstichprobe) > daraus einige Studenten ziehen

Klumpenstichprobe (cluster)

> Beispiel für Klumpen: Schule

> jede VP gehört nur zu einem Klumpen

> Klumpen sind relativ homogen, aber in sich heterogen

> Stichprobenfehler wird größer, wenn Klumpen in sich homogener und Klumpen untereinander heterogener sind

> Schritte:

1) Population besteht aus kleinen relativ homogenen Gruppen (cluster)

2) Untersuchung einer zufälligen Stichprobe von Klumpen

3) einzelne Klumpen vollständig erheben

matched samples

> wenn geringe Größe der Untergruppen in Kontrollpopulation eine vernünftige Zufallsauswahl nicht mehr zulässt

> jeder Einheit aus der Untersuchungsstichprobe wird eine Einheit aus der Kontrollpopulation zugeordnet, die der Einheit im Bezug auf das Merkmal (Alter, Geschlecht, Nationalität, Qualifikationsniveau) möglichst ähnlich ist

> Ziel: Übereinstimmung der Experimental- und Kontrollstichprobe

typische Fehler

Auswahlverzerrung (Methode der Datenerhebung)

Selbstselektion (angefallene Stichprobe)

extrem unterschiedlich große Stichproben

zu große/kleine Stichproben (Interpretation von Signifikanztests)

Kontrollgruppen

Null-Kontrollgruppe

> kein Treatment

Alternativ-Kontrollgruppe

> alternatives Treatment

Warte-Kontrollgruppe

> kein Treatment – Wartezeit – Treatment

Placebo-Kontrollgruppe

> Placebo-Treatment (sollte nicht wirken)

FORSCHUNGSETHIK

>>Perspektiven

zweckrational

> Erkenntnisgewinn steht im Vordergrund

> Druckmittel (in gewissem Ausmaß) dürfen eingesetzt werden

wertrational

> betonen starke psychische Belastung, die durch Druck entsteht

> Person > Erkenntnis

→ **institutionelle Maßnahmen: Ethikkommission**

>>Prinzipien

1) Gewährleistung psych.&phys Unversehrtheit&Integrität

2) Transparenz der Untersuchung

3) Vermeidung von Täuschung

4) Freiwilligkeit der Teilnahme

5) Vertraulichkeit der Untersuchungsergebnisse

6) Aufklärung nach Untersuchung

7) Vertrag über Rechte und Pflichten

> einige Kriterien (Transparenz und Täuschungsvermeidung) beeinträchtigen die Ergebnisse der Untersuchung: hier gilt es Kosten und Nutzen abzuwägen und die Probanden danach aufzuklären

QUANTITATIVE DATENERHEBUNG

Vorteile

Präzision

> genaue Angaben über Ausprägungsgrad eines Merkmals

Vergleichbarkeit

> Aussagen über Unterschiedsgrößen und direkte Vergleiche möglich

Sparsamkeit

> Rechenoperationen ermöglichen einfache Verknüpfungen/Zusammenfassung vieler Werte (Mittelwert usw)

Übersicht Methoden

1) **Selbstberichtsverfahren** → Fragebögen, Ratings, Tests

2) **Beobachtung**

3) **biopsychologische Messungen**

4) **Mixed-Methods/Triangulation**

Wahl der Methode:

> inhaltliche Fragestellung

> Merkmale des Untersuchungsobjekts

> Rahmen: zeitlich, finanziell, personell

> Qualitätsansprüche an Information

> Art des Verhaltens/Erlebens

1) quantitative SELBSTBERICHTSVERFAHREN

- > Mensch = Black Box
- > offenes Verhalten bedeutungslos, wenn nicht interpretierbar
- > wenn psych. Prozess für VP zugänglich ist: Selbstberichtsverfahren möglich

>> Schritte zur Auskunft

Interpretation der Frage

- > wichtige Aspekte

Bildung eines Urteils

- > Einflüsse: voraktivierte Infos (Priming), Wortwahl der Frage, Situation,...

Übersetzung in kommunizierbare Auskunft

- > schwierig bei vorgegebenen Antwortformaten, da diese wiederum interpretiert werden müssen
→ jeder Schritt ist fehleranfällig

>> Varianten&Konstruktion

schriftlich vs mündlich

- > schriftlich: Fragebogen, meist quantitativ
 - höher standardisiert, Vorwissen zu Konstruktion nötig, größere Anonymität (weniger Reaktivität)
- > mündlich: Interview, meist qualitativ
 - oft unstrukturiert, mehr sozial erwünschte Antworten, mehr Beeinflussungs- und Übertragungseffekte, einfacher für einige Befragte, höherer Aufwand, flexibler

standardisiert vs unstandardisiert

- > Standardisierung = Freiheitsgrade der befragten Person
- > standardisiert: geschlossene Fragen (vorgegeben Antworten)
- > unstandardisiert: offene Fragen (Essay, freie Assoziation)
- > halb standardisiert: halboffene Fragen (z.B. grafisch)

strukturiert vs unstrukturiert

- > Strukturierung = Freiheitsgrade des Forschenden
- > strukturiert: Wortlaut und Reihenfolge der Items vorgegeben
- > unstrukturiert: keine Vorgaben in Formulierung und Reihenfolge

Anzahl VPn

- > Einzelbefragung: nur eine VP
- > Gruppenbefragung: kleine Gruppe (5-15VPn)
- > Survey (Umfrage): hohe Anzahl VPn

Analyseeinheit

- > Individuen
- > Dyaden (Ehepaare, Interaktionspartner...)
- > soziale Systeme (Familien, Kulturen..)

transparente vs intransparente Erhebung

- > transparent: Person kennt Verwendungszweck der Daten, Verfälschungsgefahr (Meinungsumfrage, Eignungstest..)
- > intransparent: Person kennt weder Bedeutung ihres Verhaltens noch Zweck der Daten (biopsych. Messung, nonreaktive Erhebung, IAT)

typisches vs maximales Verhalten

- typisch: fast alle Persönlichkeitstests
- maximal: IQ-Tests, Leistungsvermögen wird erfasst

>>Reaktivität

- > Veränderung/Verzerrung der erhobenen Daten allein aufgrund der Kenntnis der VP, dass sie Untersuchungsgegenstand ist → unerwünschte Verhaltensveränderung
- > Befragungen verändern Denken / Beobachtungen verändern Verhalten
- > Personen wollen positiven Eindruck hinterlassen

>> Gegenmaßnahmen

Unkenntnis der VP

- > effektivste Möglichkeit: VP weiß nicht, dass sie untersucht wird
- > bei Feldstudie möglich, bei Experiment nicht
- > Vorsicht: Ethik

Zusicherung der Anonymität

- > verbessert Ehrlichkeit und Kooperation
- > besonders wichtig bei Umfragen (Meinungen usw)

Formulierung einer Coverstory

- > gezielte Induktion bestimmter Erwartungen über Fragestellung/Hintergrund, obwohl in Wahrheit etwas anderes untersucht wird
- > Ehtikproblem: Täuschung der VP

Verwendung nichtreaktiver Messverfahren

- > Maße, die die VP nicht beeinflussen kann (biopsychologische Maße)
- > selbst wenn VP weiß, was gemessen wird, kann sie es nur schwer steuern

Verwendung indirekter/impliziter Maße

- > anderer Zweck der Messung wird suggeriert (Reaktionszeit bei IAT)
- > Auswertung von Verhaltensaspekten, die die VP nicht mit Messung in Verbindung bringt

> mündliche Befragung/Interview

- > häufig weniger standardisiert
 - offene Fragen, höhere Reaktivität, stärkere Gefahr VL-Effekte
- > aufwändig und kostenintensiv
 - Einzeltestung, Schulung der Interviewer, Kodierung von Antworten
- > erfordert nicht die Kompetenz zu schriftlicher Ausdrucksfähigkeit
- > Nachfragen möglich
- > häufig flexibler
- > unterscheidet sich im Aktivitätsgrad des Interviewers, Anzahl der Personen und im Strukturierungsgrad des Interviews

> Schriftliche Befragung

- > Befragungssituation kontrollieren: sonst leidet Validität (online schlecht)
- > Problem: Systematisch fehlende Daten

>>Vorraussetzungen

- > Menschen beobachten sich selbst
- > Menschen erwerben im Zuge der Selbstbeobachtung Wissen über sich
- > Menschen sind fähig und gewillt, dieses Wissen dem VL mitzuteilen

>>Formulierung der Fragen

- > nicht überfordern durch Formulierung oder zu genaue Gedächtnisleistung (Fähigkeiten der Befragten beachten)
- > keine doppelten Verneinungen

- > nicht mehr als 1 Aspekt pro Frage
- > keine forced choice, wenn sich Antworten nicht logisch ausschließen
- > keine Fragen, die von allen einheitlich beantwortet werden: keine Streuung = keine Aussage
- > mehrere Items pro Aspekt = höhere Reliabilität
- > Reihenfolge: Anfang einfacher als Ende
- > klare Instruktion am Anfang

lexikale Ebene

- > ungebräuchliche Begriffe/Fachwörter vermeiden

syntaktische Ebene

- > keine langen, verschachtelten Sätze

semantische Ebene

- > keine abstrakten und komplexen Sachverhalte

>>Ratings

- > Befragte geben Urteile auf numerisch interpretierbarer Skala ab (Beurteilung eines Gegenstands hinsichtlich eines Merkmals)
- > gleich große Abschnitte: **Äquidistanzannahme**(Intervallskalenniveau kann angenommen werden)
- > meist mit **Likert-Skala**
- > beachten: **Diskriminationsfähigkeit VP** (wie genau kann die VP unterscheiden/abstufen)
- Differenzierbarkeit des Sachverhalts** (wie genau ist der Sachverhalt abstufbar)
- > mögliche Skalen: Häufigkeiten, Intensitäten, Bewertung, Wahrscheinlichkeit (äquidistant)

Konstruktion

- > Items: Frage oder Aussage
- > Endpunkte: uni-oder bipolar
 - **unipolar**: ruhig/unruhig – Gegenteil/natürlicher Nullpunkt
 - **bipolar**: ruhig/angespannt – wechselseitige Definition
- > Abstufung der Skala
 - meist 4-9 Stufen
 - zu wenige = Informationsverlust / zu viele = übersteigerte Differenzierung
- > Stufenanzahl
 - **gerade**: neutrale Option, problematisch wenn zu häufig
 - **ungerade**: Tendenz erzwungen
- > Kennzeichnung der Abstufungen
 - **numerisch**: eindeutig, gleiche Abstände
 - **verbal**: leicht verständlich, Abstände schwierig gleich zu halten
 - **sprachfrei (grafisch)**: Bsp:verschiebbarer Strich auf Skala; leicht verständlich
 - **sprachfrei (ikonisch)**: Bsp: Smileys; leicht verständlich

>> Semantisches Differenzial

- > durch Antworten auf mehrere bipolare Items entsteht ein Polaritätsprofil, das die mentale Repräsentationen des Gegenstands in wesentlichen Merkmalen wiedergibt
 - zwei Objekte im Vergleich
 - schneller Überblick: zentrale Merkmale und Unterschiede

>> Urteilstendenzen bei Beantwortung

Tendenz zur Mitte

- > Vermeidung von Extremurteilen, v.a. wenn wenig mit Gegenstand vertraut
- > Problem: keine Streuung = keine Aussagekraft
- > Ambivalenz-Indifferenz-Problem: Ist VP unentschlossen oder gleichgültig?
- > Lösung: bipolare Skala (Verankerung) & gerade Stufenanzahl

Gedankenlose Reproduktion

- > ähnlich erscheinende Items werden genau gleich beantwortet
- > Lösung: Umpolung (Benennung vertauschen)

Primacy-Effekt

- > anfängliche (extreme) Urteile beeinflussen nachfolgende Urteile
 - wahrgenommene Ähnlichkeit von Items = ähnlich extreme Bewertung
 - wahrgenommene Unterschiedlichkeit der Items = Kontrastierung verzerrt Wertung
- > Lösung: Reihenfolge interindividuell (Stichprobe) ausbalancieren → Mittelung

Halo-Effekt

- > Beurteilung eines Objektes hinsichtlich verschiedener Merkmale aufgrund eines zentralen Merkmals
- > Bsp: gutmütige Personen werden auch als freundlich, zuverlässig, usw wahrgenommen
- > v.a. wenn Schlüsselmerkmal unklar/ungewöhnlich definiert
- > Lösung: klare Infos über Unterschiede zwischen Merkmalen geben

>Testen

- > v.a. diagnostisch auf individueller Ebene (ideografischer Ansatz), aber trotzdem auf Gruppenebene verglichen (nomothetischer Ansatz), um Referenzpopulation zu erhalten

Definition

- > wissenschaftliches Routineverfahren zur Erfassung eines oder mehrerer empirisch unterschiedlicher Persönlichkeitsmerkmale
- > genaue quantitative Aussage über relativen Grad der Merkmalsausprägung
- > Bedingung: Items werden von verschiedenen Menschen mit verschiedenen Merkmalsausprägungen unterschiedlich beantwortet
- > standardisierte Verhaltensstichprobe, die zu einem Testwert der VP aggregiert wird

Leistungstests vs Persönlichkeitstests

Leistungstests

- > objektiver Maßstab: Güte der Antworten (richtig/falsch)
- > entscheidend: kognitive Prozesse & Leistungsmotivation
- > Bsp: Intelligenz & Eignungstests
 - **Speed-Tests**: zu knappe Bearbeitungszeit, gleiche Schwierigkeit (Konzentrationstests)
 - **Power-Tests**: Aufgabenniveau sukzessive gesteigert (Intelligenztest HAWIE)

Persönlichkeitstests

- > emotional, motivational, sozial relevante Persönlichkeitseigenschaften (kein richtig/falsch)
- > Bsp: NEO Five Factor Inventory
 - **subjektive Persönlichkeitstests**: leicht durchschaubarer Zweck
 - **objektive Persönlichkeitstests**: Zweck verschleiert > geringere Reaktivität

klassische Testtheorie

- > Wert einer Person = wahrer Testwert + Fehleranteil
 - wahrer Wert praktisch nicht erzielbar, sondern nur empirisch nach Abzug des Fehlers schätzbar
 - mehrere Items pro Merkmal > Mittelung > Fehlerminimierung
- > **Axiome** (Vorraussetzungen)
 - Ergebnis: wahrer Wert + Fehler
 - mittlerer Messfehler = 0
 - Messfehler nicht mit anderen Tests, Messfehlern aus anderen Tests und wahren Wert korreliert

probabilistische Testtheorie

- > Antworten auf Items sind Indikatoren für latente Merkmale
 - die WS, dass eine Person Frage 1 auf Merkmal 1 positiv beantwortet, hängt von Ausprägung des Merkmals 1 ab
- > statistische Beschreibung der Beziehung zwischen Antwortverhalten (Lösungs-WS) und latentem Merkmal
- > grafisch: ICC (Item-Characteristic-Curves)

adaptives Testen

- > Aufgaben ändern sich in ihrer Schwierigkeit – optimal der Fähigkeit einer Person entsprechend
- > präzise Schätzung anhand weniger Items
- > hohe Testmotivation
- > computerbasiert
- > **tailored** passt sich nach jeder Frage an
- > **branched** erst nach Fragenblock

wünschenswerte Eigenschaften von Test(-items):

> Kriterien zur Individualdiagnose

Homogenität

- > Items auf Skala = homogen → erfassen nur ein Merkmal
(wenn mehrdimensional: mehrere Skalen)

Differenzierung

- > differenzierte Erfassung des Merkmals → möglichst viele Ausprägungen realisierbar

Distinktionsfähigkeit

- > eindeutige Unterschiede zwischen Personen mit hoher und niedriger Merkmalsausprägung

> allgemeine Kriterien zu Güte von Datenerhebung

Objektivität/Reliabilität/Validität (oben beschrieben)

> Kriterien innerhalb eines Items

Schwierigkeit

- > Prozentsatz aller Personen die das Item gelöst (Leistungstest) oder positiv beantwortete haben (Persönlichkeitstests)
- > breite Streuung erwünscht, um gesamtes Spektrum zu differenzieren

Trennschärfe

- > wie gut sagt ein Item das gesamte Testergebnis vorher?
- > trennscharfe Items haben einen hohen Wert, wenn das Gesamtergebnis hoch ist
- > hohe Trennschärfe grundsätzlich erwünscht
- je größer die Streuung der Schwierigkeit, desto geringer die Trennschärfe – Problem

Arten von Testitems

offene Beantwortung

- > Bsp: freie Assoziation bei Tintenkleks
- > Gütekriterien schlechter

halboffene Beantwortung

- > Bsp: vorgegebene Formulierung vervollständigen
- > Gütekriterien schlechter

Antwortvorgaben

- > Multiple Choice (MC)
- > weit verbreitet, da Gütekriterien meist erfüllt

Leistungstests – Verfälschungen & Gegenmaßnahmen

Raten

- > VP rät die richtige Antwort, ohne sie zu kennen
 - **Distraktoren:** logisch wirkende Antworten (richtige A. nur erkennbar, wenn wirklich man sie wirklich kennt)
 - **Ratekorrektur:** bei Auswertung wird Punktzahl, die durch Raten erreicht werden kann abgezogen

Persönlichkeitstests – Verfälschungen & Gegenmaßnahmen

Verfälschungen:

Selbstdarstellung (impression management)

- > Person möchte sich positiv darstellen oder eine bestimmte Rolle (nicht) einnehmen
- > gezielte Verfälschungen

soziale Erwünschtheit

- > Person antwortet so, wie sie den angenommenen Erwartungen und Normen der Gruppe entspricht, die das Ergebnis kennt
- > unterscheidet sich je nach Kontext / Ziel: Annerkennung

schematische Antworttendenzen

- > Tendenz zur Mitte
- > Ja-/Nein-Sage-Tendenz
- > systematisches Überspringen von Items
- teilweise unbewusst

Gegenmaßnahmen:

ausbalancierte Antwortvorgaben

- > beide Antwortrichtungen = gängige Normen

Aufforderung zu korrektem Testverhalten

- > bitten, ehrlich zu antworten (Druck ausübbar: „ich erkenne Lügen“ > Ethikproblem)

Kontrollskalen/Lügenskalen

- > Verhaltensweise, die negativ bewertet ist, aber trotzdem jeder macht
- > Bsp: Lügen schlecht → aber jeder nutzt manchmal Notlügen
- > wenn Person immer wieder mit „nie“ antwortet, ist es wahrscheinlich, dass sie lügt (soziale Erwünschtheit)

Randomized – Respons – Technik

- > Personen neigen weniger zu Verfälschung, wenn sie denken, dass ihre korrekten Antworten unbekannt sind (v.a. bei peinlichen, strafbaren Handlungen)
- > Wahrheitsgehalt abhängig von Zufallsereignis (Würfeln: bei 3&5 lügen), das nur Proband kennt
- > Testleiter unwissend, Proband sagt eher Wahrheit
- > WS für Verfälschung geschätzt, Vergleich zwischen „ehrlicher und unehrlicher Stichprobe“
 - wenn unterschiedlich, dann hat unehrliche Stichprobe (ohne Technik) verfälscht

2) Beobachten

- > systematische und regelgeleitete Registrierung des Auftretens bzw. der Ausprägung von ausgewählten Merkmalen/Ereignissen
- > in quantitativer Forschung: Zählen

Aufgaben des Beobachters

- > **Selektivität:** Information aus Wahrnehmungsstrom filtern und in bedeutungsvolle, stabile Einheiten organisieren
- > **Konstruktivität:** jeder Beobachter erzeugt ein Bild der Wirklichkeit
 - beobachterabhängige Selektivität und Konstruktivität kontrollieren!

wissenschaftliche Beobachtung = Beobachtungsplan

- > Kategorien bilden
- > relevante Aspekte finden
- > Interpretationsspielraum festlegen
- > Zeitpunkt/Zeitraum/Situationen der Beobachtung festlegen
- > Registrierung und Protokollsystem festlegen
- > Stichprobe: Ereignis- oder Zeit?
- Ziel dieser Regeln: Gütekriterien erfüllen

Modellierungsregeln

Selektion

- > Auswahl bestimmter Beobachtungsgegenstände bzw Herausfiltern bestimmter Reize

Abstraktion

- > Ereignis wird aus seinem jeweiligen konkreten Umfeld herausgelöst

Klassifikation

- > Zuordnung von Zeichen und Symbolen zu bestimmten Ereignis-/Merkmalsklassen

Systematisierung

- > Aufbereitung der Beobachtungsdaten

Relativierung

- > Aussagegehalt einer Beobachtungsstudie

Ordnungsmöglichkeiten

systematisch vs unsystematisch

- systematisch: Gegenstände und Einheiten festgelegt → Protokollsystem
- unsystematisch: alles verfolgen/filmen → wenn vorher noch unklar über mögl. Ergebnisse

natürliche Bedingungen vs Labor

- natürlich: normale Umgebung → gut: Generalisierbarkeit; schlecht: Störvariablen
- Labor: standardisiert → hohe interne, niedrige externe Validität (unnatürlich)

Beteiligungsgrad des Beobachters

- aktiv: anwesend, interagiert
- passiv: anwesend, macht nichts
- nicht anwesend

Art der Protokollierung

- **isomorphe Aufzeichnung:** Geschehen ohne Informationsverlust aufzeichnen (Video,...)
- **reduktive Aufzeichnung:** bei Beobachtung wird ein Teil der Infos ausgesondert

> **Zeichensysteme:** best. Verhalten wird analog festgehalten (gezählt) – Verteilung des Verhaltens über Zeit

> **Kategoriensysteme:** gliedern Verhalten in Einheiten und ordnen sie Verhaltenstypen zu – Häufigkeitsprofil der Verhaltenstypen

> **Ratingsysteme:** Verhalten wird über definierten Zeitraum beobachtet und hinsichtlich best. Merkmale eingeschätzt

teilnehmend vs. nicht teilnehmend

offen vs verdeckt

Konvergenz (am Beispiel der Fremd- und Selbstbeobachtung)

- erhöht durch hohe Beobachtbarkeit
- erniedrigt durch soziale Erwünschtheit > Verzerrung
- erhöht durch Konsistenz und Alltagsnähe der Merkmale
- Personen, die sich konsistent verhalten, sind einfacher einzuschätzen als variable Personen

- Personen, die sich selbst darstellen wollen oder selbst täuschen, stimmen mit dem Urteil anderer weniger überein
 - höher, wenn Beurteiler sehr fähig ist, selbst sein Urteil nicht verzerrt, keine Beurteilerstile hat
- > Halo-Effekt: ungerechtfertigte Generalisierung aufgrund eines bestimmten Merkmals über alle Facetten hinweg
- erhöht, wenn Dyade (Beurteiler und Beurteilender) unter einem Merkmal dasselbe verstehen
 - höher, wenn Dyade sich gut kennt
- > Selbstbeurteilung lässt sich nicht durch eine einzige Fremdbeurteilung validieren
- > dazu mehr Methoden benutzen (Beobachtung..)
- > Fremdbeobachter sollte geschult sein
- > Vergleich von erwarteter und erhaltener Konvergenz
- > äußere Umstände berücksichtigen

3) Biopsychologische und neurowissenschaftliche Messungen

- > mentale und psychische Vorgänge sind auf funktionelle biologische (v.a.) neuronale Systeme angewiesen
- > **wechselseitiger Einfluss** Biologie <> Psychologie
- > (beabsichtigte) Verfälschungen der Messung, aber Körper reagiert auf ungewohnte Situation
- > meiste Verfahren: **nichtinvasiv** (kein Eindringen in Gewebe)
- > **Spezifitätsproblem** – individualspezifische Reaktionen
- > **Auswertungsproblematik** – Veränderungsmessungen können systematische Fehler enthalten

Messmöglichkeiten

> **autonomes Nervensystem**

elektrodermale Aktivität (EDA)

- > Tätigkeit der Schweißdrüsen → Hautleitfähigkeit
- > zwei Elektroden an zwei Fingern
- > tonische Hautleitfähigkeit = Niveau der Aktivität
- > reaktive Hautleitfähigkeit = Reaktion auf bestimmte Stimuli
- > je höher Aktivität, desto höher Stress bei neuen Reizen
- > Indikator für Orientierungsreaktion, Habituation, Therapiekontrolle

Elektrokardiogramm (EKG)

- > elektrische Ströme der Herzmuskulatur gemessen (Herzrate, Herschlagfrequenz → Herzratenvariabilität)
- > Messung: Brustwandableitung oder Einthoven Ableitung 2 (linker Arm/rechtes Bein)
- > Blutdruckmessung durch Manschettendruckverfahren
- > Korrelate: Stress, Emotionen, Depression, Aufmerksamkeit

auch möglich: Atemfrequenz und Pupillenweite

- > Pupille: Emotionen, Aktivierung, Ausmaß Informationsverarbeitung

> **somatisches Nervensystem**

Elektromyografie (EMG)

- > Registrierung von Aktionspotenzialen der Muskelzellen (Gesicht)
- > Indikator für Emotionen (7 Grundemotionen: Wut, Ekel, Angst, Überraschung, Freude, Trauer, Verachtung)

Elektrookulogramm (EOG)

- > wohin blickt Person wie lange
- > Indikator für Aufmerksamkeit und Zuwendung kognitiver Ressourcen
- > Messung: Spannungsdifferenzen über Elektroden

Eyetracker

- > Messung: Kameras (Augen) & Headset (Kontrolle der Bewegungsartefakte des Kopfes)
- > Stellung und Bewegung der Augen: Fixation & Sakkaden (rasche Bewegungen zwischen Punkten)
- > Indikator für Aufmerksamkeit und Zuwendung kognitiver Ressourcen

Lidschlag

- > Orientierungsreaktion, Schutzreflex

Hormonsystem (endokrines System)

- > Hormone koordinieren Steuerung der Aktivität verschiedener Organe
- > Indikatoren: erlebter (Kortisol) und körperlicher (Adrenalin) Stress, Bindung (Oxytocin),...
- > Messung: Blut, Urin, Speichel

Immunsystem

- > Antikörper, immunaktive Zellen im Blut, Immunglobulin im Speichel
- > Indikator: Entspannung/Anspannung, psychosomatischer Gesamtzustand, Erschöpfung

> zentrales Nervensystem (ZNS)

Elektroenzephalogramm (EEG)

- > an Schädeloberfläche werden per Elektroden elektr. Potenziale von (Kortex-)Neuronen gemessen
- > Rückschlüsse auf Hirnaktivität > nur ganze Neuronenverbände messbar (Signalstärke)
- > Artefakte durch Augenbewegung (EOG nutzen)
- > Spontanaktivität (ständige Potenzialveränderungen) verglichen mit evozierten Potenzialen (kurze Reaktionen auf äußere Reize)
- > Indikator: allg. Aktivierung, Wachheit, Schlafforschung
- > hohe zeitliche, niedrige räumliche Auflösung

Positronenemissionstomografie (PET)

- > radioaktive Markierungssubstanzen (tracer) in Blut zeigen aktive Strukturen an
- > bildgebendes Verfahren: welche Areale sind an welchen Funktionen beteiligt

Magnetresonanztomographie (MRT/MRI)

- > nutzt Kernspin von Wasserstoffatomen
- > starke Magnetfelder stellen Struktur und Funktion des Gehirns dar
- > Dichte und Art des Gewebes erkennbar

funktionelles MRT (fMRT/fMRI)

- > nutzt sauerstoffreiches Hämoglobin > aktive ZNS-Regionen erkennen
- > aktive Hirnareale werden farbig in Darstellung markiert
- > Indikator: Verarbeitung von Reizen
- > Problem: laut, teuer, Bewegungsartefakte, nicht bei Metallimplantaten

Computertomografie (CT) – Röntgen

- > bildhafte Darstellung des Gehirns
- > Röntgenstrahlen durchdringen verschiedenes Gewebe verschieden stark
- > Problem: schädliche Strahlung

Artefakte

- > Signale, die einen anderen Ursprung als interessantes Biosignal haben

physiologisch

- > Grundaktivitäten des Körpers (Atmen beeinflusst Herzfrequenz) → kaum kontrollierbar
- > kann durch Filter rausgerechnet werden

Bewegung

- > Bewegung der VP, unkontrollierte Bewegungen der Apparatur
- > v.a. bei bildgebenden Verfahren problematisch

externe Einstreuung

- > elektrische/magnetische Felder in Umgebung (Radio, Leitungen..)
- > Lösung: Raum möglichst komplett abschirmen

Implizite Messungen

IAT (impliziter Assoziationstest)

- > Messung von Assoziationsstärken zwischen mentalen Repräsentationen von Objekten im Gedächtnis, häufig Messung zu Einstellungen
- > Beispiel: man muss per Knopfdruck positive und negative Wörter Bildern von Homo- und Heterosexuellen zuordnen: man wird die positiven Wörter schneller und fehlerfreier der Gruppe zuordnen, die man präferiert und umgekehrt

Priming Paradigmen

Emotionale Stroop-Aufgabe

→ beide nicht auf Folien erwähnt, nur im Eid

4) MIXED METHODS & andere

Arten von Methodenvergleichen:

- > Verfahren: Test vs. Fragebogen
- > Perspektive: Selbst- vs. Fremdeinschätzung
- > Forschungstradition: qualitativ vs. quantitativ
- > Zeitpunkt: quer- vs. längsschnittlich

Triangulation

> Erfassung eines Gegenstandes durch (mindestens) zwei verschiedene Methoden (Campbell & Fiske, 1959)

Varianten nach Denzin

> Einnahme verschiedener Perspektiven oder Anwendung verschiedener Zugangsweisen auf einen Forschungsgegenstand

Datentriangulation

> Heranziehen mehrerer Datenquellen (Interviewmaterial, Dokumente)

Investigatortriangulation

> Untersuchung eines Phänomens durch mehrere Forscher (Beobachtungsteam)

Theorientriangulation

> Heranziehen mehrerer Theorien als „Filter“, durch den Phänomene betrachtet werden

Methodentriangulation

> Verwendung verschiedener Zugänge zur Erfassung eines Phänomens, um Grenzen der Methoden zu überwinden

> in Datenerhebungsphase

- within-methods-Triangulation

- > mehrere Operationalisierungen eines Konzepts kommen in einer Methode zum Einsatz
- > Bsp: innerhalb eines Fragebogens gibt es standardisierte und offene Fragen

- between-methods-Triangulation

- > Verwendung unterschiedlicher Methoden auf einen Gegenstand
- > Bsp: Interview und Beobachtung
- > häufigste Triangulationsform

→ Methodentriangulation ist nur dann Mixed-Methods wenn eine qualitative und eine quantitative Methode benutzt wurden

→ Probleme: woher weiß man, ob Methoden konvergent sind oder ob sie mehrere Facetten messen

sollen = Interpretation von unterschiedlichen Ergebnissen erschwert
→ mehr Methoden sind nicht automatisch besser (Verdoppelung der Fehler)
→ bei konvergenter Validität: wenig Abweichungen durch Methodeneffekte

Mixed Methods

- > Kombination von Methoden und Ansätzen qualitativer und quantitativer Forschungstradition
- > Bsp: explorativ-qualitativer-Fragebogen; danach quantitative Survey-Studie

sonstige Varianten (nicht in Folien)

Mixed-Methods-Studie

> Kombination qualitativer und quantitativer Elemente in einer Untersuchung oder mehreren aufeinanderfolgenden Untersuchungen

- Multimethod-Design

> Kombination qual.+quant. Forschungsphasen, die in sich eigene Studien sind

- Mixed-Models-Design

> Kombination qual. + quant. Elemente über verschiedene Untersuchungsphasen hinweg

- Methodentriangulation

> sofern quant. + qual. Methoden enthalten

Multimethod-Studie

> Kombination mehrerer Forschungsmethoden in einer Phase (alle qualitativ, alle quantitativ oder gemischt)

Monomethod-Studie

> nur 1 Forschungstradition bestimmt Methoden und pro Phase nur eine Methode genutzt

Hybride Methoden

> in Konzeption der Methoden ist bereits eine Mischung beider Forschungstraditionen enthalten

Systematisierung von Mixed-Methods

- > Reihenfolge: simultan oder sequenziell
- > Gewichtung: gleichgewichtet oder eine übergeordnet
- > Auswertung: nur eine Tradition oder beide
- > hypothesentestend oder -generierend
- > Verbindung: eingebettet, vermischt, (sequenziell) verbunden

Messtheorie

>> **Unterschied Zählen <> Messen**

Zählen

- > diskrete Merkmale (Anzahl Geschwister...)
- > keine feineren Abstufungen als ganze Zahlen

Messen

- > Erfassung von stetigen Merkmalen (Größe...)
- > unendlich viele mögliche Ausprägungen/Intensitäten/Abstufungen

>> **Grundlagen der Messtheorie**

- > vor Erhebung: Skalenniveau bestimmen
- > Zuordnung: empirisches Relativ – numerisches Relativ

empirisches Relativ

> Menge von unterschiedlichen Objekten und Untersuchungsergebnissen

numerisches Relativ

> Menge von zugeordneten Zahlen

→ Ziel: **homomorphe Abbildung** (Relationen sollen erhalten bleiben)

Voraussetzungen

> Merkmal festlegen

> Operationalisierung

> geeignetes Messinstrument

>> Skalenniveaus

> entscheidet über mögliche Aussagen/Interpretationen der Werte und Anwendungen der statistischen Verfahren

> Transformationen werden verwendet, damit statistische Verfahren anwendbar sind

Nominalskala

> Verschiedenheit \leftrightarrow Äquivalenz

> Werte nicht interpretierbar und beliebig

> Beispiel: demografische Merkmale (Geschlecht..)

> alle eindeutigen Transformationen (nicht: $f(x)=100$)

Ordinalskala

> Rangordnung (größer/kleiner)

> Abstände nicht interpretierbar

> Beispiel: Platzierung, eigentlich auch Schulnoten

> alle streng monotonen Transformationen (nicht: $f(x) = 1/x$)

> Achtung: **Intransitivität** bei Rangfolge > niedrigeres Niveau annehmen oder Bezeichnung ändern

Intervallskala

> Größe/Relationen von Differenzen

> Skala festgelegt, Nullpunkt nicht

> keine direkten Vergleiche zwischen den Werten

> Beispiel: IQ

> alle Operationen, bei denen Differenzen erhalten bleiben (nicht quadrieren/Wurzel ziehen/logarithmieren)

> Grauzone: bei Ratingskalen/Schulnoten usw. wird häufig Intervallskalenniveau angenommen, obwohl die Merkmale eigentlich nur auf Ordinalniveau gemessen wurden > besser für statistische Verfahren (wichtigste hier anwendbar)

Verhältnis-/Ratioskala

> direkte Vergleiche zwischen Werten

> natürlicher Nullpunkt festgelegt

> Beispiel: Messungen in Meter

> alle Operationen, die Verhältnisse nicht verändern (keine Addition, nur Multiplikation)

Absolutskala

> Anzahlen

> nur Identitätstransformationen

→ alles, was bei niedrigeren Skalen interpretierbar war, ist auch bei höheren erlaubt

→ alle Transformationen, die man bei niedrigeren Skalenniveaus nicht durchführen durfte, sind auch bei höheren nicht erlaubt

Experiment

Merkmale

> 2 Klassen von Variablen: $X = UV = \text{Explans} / Y = AV = \text{Explandum}$

> die unabhängige Variable geht der abhängigen Variablen zeitlich voraus

- > systematische Beobachtung einer abhängigen Variablen unter verschiedenen Bedingungen einer unabhängigen Variablen
- > Störvariablen werden kontrolliert (hohe interne Validität)
- > zufällige Zuordnung der Probanden zu experimentellen Bedingungen
- > nur eine UV wird gleichzeitig variiert
- > unterscheiden sich die MW zweier Versuchsgruppen, weist das auf die Wirksamkeit der UV hin
- > nur, wenn Experimentalbedingungen erfüllt, kann eine **Kausalaussage** getroffen werden
- > hohe Präzision
- > geringe externe Validität
- > hoher zeitlicher und personeller Aufwand

Design

within/Wiederholungsmessung

- > AV prä → Manipulation UV → AV post
- > eine Gruppe wird sämtlichen experimentellen Bedingungen unterzogen

between/unabhängig

- > AV Kontrollgruppe ohne Manipulation <> AV Experimentalgruppe mit Manipulation
- > interindividuelle Unterschiede verteilen sich zufällig (randomisiert) auf Bedingungen

Blockbildung

- > bekannte Störfaktoren werden erfasst und VP danach aufgeteilt
- > Parallelisierung (Erfassung bekannter Störvariablen und Aufteilung in homogene Gruppen)

Störvariablen und Kontrolle

Störvariablen

- > andere Einflussgrößen, die systematisch mit den Stufen einer UV variieren und auf die AV einwirken „**Konfundierung**“

> Klassen

Versuchspersonenmerkmale

- > VP unterscheiden sich in verschiedenen experimentellen Bedingungen der UV hinsichtlich eines oder mehrerer Merkmale (Alter, Bildung..)

Situationsmerkmale

- > Situationen für Probanden unterscheiden sich in verschiedenen Stufen der UV systematisch

Versuchsleitermerkmale

- > VL verhält sich systematisch verschieden (oder anderer VL) in unterschiedlichen Stufen der UV
- > Motive, Erwartungen, physische und psychische Eigenschaften

äußere Störvariablen

- > Zeitgeschehen, Reifung
- > Instrumentierung (anders gemessen)
- > statistische Regression zur Mitte (mangelnde Reliabilität des Messinstruments)
- > Auswahlverzerrung
- > Ausfalleffekte

Kontrolle

- > wenn man sie kontrolliert, wirken sie zwar noch, aber konfundieren nicht mehr (Effekt der UV wieder klar nachweisbar)

Konstanthaltung

- > man hält Situations- und VL – Merkmale in allen Stufen gleich
- > bezogen auf alle Standardisierungsmaßnahmen für Situation und Versuchsleiterschulung

Elimination

- > Einflüsse einer Störvariablen werden auf 0 gesetzt (Lärm > schalldichter Raum; Versuchsleiter > Tonband)

systematische Variation

- > Störvariable wird zu UV gemacht (Kontrollfaktor)

zufällige Variation

- > Realisation möglichst vieler Ausprägungen der Störvariablen und zufällige Zuordnung zu den Probanden

- > geht nur in großen Stichproben, wenn in jeder Gruppe etwa gleich viele Vpn sind (Mittelung)

Randomisierung (VP)

- > Versuchspersonen werden den Versuchsbedingungen randomisiert zugeordnet
- > Idealfall: statistische Äquivalenz der Gruppen hinsichtlich aller denkbaren Störvariablen
- > Vorteil: Störvariablen müssen vorher nicht bekannt sein und beliebig viele können kontrolliert werden

Parallelisieren

- > bei kleinen Stichproben, wenn Randomisierung nicht geht
- > erst wird interessante AV gemessen > Rangordnung erstellen
- > immer zwei hintereinanderliegende Vpn werden per Zufall beiden Versuchsbedingungen zugeordnet (oder auch 3 Personen und 3 Bedingungen)
- > vergleichbarer MW der Störvariablen in allen Gruppen
- > Nachteil: man muss Störvariablen kennen, höchstens 3 können kontrolliert werden

Blindversuche

- > VL kennt Hypothese nicht
- > keine systematische Erwartungshaltung und -übertragung
- > **Doppelblindversuche:** sowohl VL als auch VP kennt Zweck nicht

spezielle Störvariablen – Messwiederholung

- > Probanden werden in verschiedenen Bedingungen mehrfach beobachtet
- > Vorteil: weniger Vpn, höhere Präzision, weniger Fehler

Positionseffekte

- > Ermüdungs-, Übungs-, Sensibilisierungs- und Erinnerungseffekte
- > je nachdem, wann die Messung in der Stufe erfolgt, wird AV positiv oder negativ beeinflusst
- > je später, desto stärker Effekt > je größer zeitlicher Abstand, desto schwächer
- > schlechter bei Ermüdung / besser bei restlichen Effekten

Übertragungseffekte

- > inhaltliche Wirkung von zeitlich früheren Bedingungen beeinflussen AV in späteren Bedingungen
- > Bsp: man lernt eine Taktik zum Problemlösen
- > unabhängig von Experimentdauer bis zur Anwendung (ungleich Positionseffekt)
- > je größer zeitlicher Abstand, desto schwächer

zwischenzeitliches Geschehen

- > Einflussgrößen auf die AV, die sich nicht aus dem Versuch, sondern aus dem Alltag der Vpn ergeben

>>Kontrolle

- > vollständige interindividuelle Ausbalancierung
- > alle möglichen Reihenfolgen der experimentellen Bedingungen werden realisiert
- > jeder Bedingungsreihenfolge wird per Zufall mindestens 1 Vp zugeordnet

Gütekriterien

interne Validität

- > erfasst das Ausmaß der Kontrolle der Störvariablen
- > hoch, wenn Veränderungen der AV ausschließlich durch UV bedingt
- > wenn niedrig: Kausalität nur vorläufig annehmbar

externe Validität

> **Populationsvalidität**

- > Übertragbarkeit der Stichprobe auf Population
- > hoch = adäquate Stichprobenwahl > große Zufallsstichprobe (nur hier komplett sicher)

> **Situationsvalidität**

- > Übertragbarkeit auf normale Situation
- > bei (Labor-)Experiment i.d.R. nicht gegeben, da künstliche Umgebung
- > Feldexperiment/-studie besser geeignet

> **Variablenvalidität**

- > Übertragbarkeit auf andere Formen der Operationalisierung
- > hoch, wenn Operationalisierung Bedeutungskern getroffen hat

Folgen von Einschränkungen der externen Validität

- > eingeschränkter Geltungsbereich der Hypothesen
- > Lösung: **Replikationsstudien**: danach auf den neuen Bereich übertragbar

Ableitungsvalidität

- > gegeben wenn Konkretisierungsschritte $TIH > EIH > SV > THn$ gelingen ohne inhaltliche/formale/statistische Veränderung der Ausgangshypothese

statistische Validität

- > Beurteilung der Güte der Auswahl und Anwendung statistischer Verfahren und Interpretation der Ergebnisse
- > niedrig bei verletzten Voraussetzungen (ungleiche Zellhäufigkeiten...)
- > Folgen: unkontrollierte Veränderungen der Fehler-WS und falsche Hypothesenentscheidungen

Präzision

- > Wahrscheinlichkeit, dass ein in der Population tatsächlich vorhandener Zusammenhang bzw. Unterschied durch die Untersuchung auch wirklich nachgewiesen werden kann
- > **Primärvarianz** = Anteil der Varianz in der AV, der auf die Uvn zurückgeht
- > **Sekundärvarianz** = Anteil der Varianz in der AV, der durch Störvariablen zustande kommt

Max-Kon-Min-Prinzip: Primärvarianz maximieren und Sekundärvarianz minimieren durch Kontrolle der systematischen Fehlervarianz

Varianten des Experiments

Bedingungen für Kausalschluss:

- > Ursache geht Wirkung zeitlich voraus
- > Ursache und Wirkung haben erkennbaren Zusammenhang
- > alternative Erklärungen (Drittvariablen) ausgeschlossen (Störvariablen, Randomisierung)
- ansonsten nur eingeschränkte Kausalität (Quasiexperiment,..)

Experiment

- > UV-AV getrennt (zeitliche Sequenz, systematische Variation)
- > Randomisierung
- > Ort: Labor = Kontrolle der Störvariablen
- > interne Validität = hoch / externe Validität = niedrig

Feldexperiment

- > gleiche Eigenschaften wie Experiment außer Durchführungsort
- > wird in normaler Umgebung durchgeführt
- > Störvariablen schwerer kontrollierbar

> interne Validität = niedriger / externe Validität = höher

Quasiexperiment

- > gleich wie Laborexperiment, nur ohne Randomisierung
- > bei Messwiederholungen, organismischen Variablen, angewandten Fragestellungen
- > interne Validität = niedriger / externe Validität = niedrig
- > Zeitreihenversuchspläne
- > Einzelfallversuchspläne
- > Versuchspläne mit Ausbalancierung

>> Versuchspläne Quasiexperiment

Pläne ohne Aussagekraft

> Ein-Gruppen-Nachtest-Design

- Experimentalgruppe nur nach Treatment gemessen

> Nur-Nachtest-Design mit nichtäquivalenten Versuchsgruppen

- Experimentalgruppe mit Treatment, Kontrollgruppe ohne: beide nur danach gemessen

Pläne mit Aussagekraft

> Ein-Gruppen-Vortest-Nachtest-Design

- Experimentalgruppe vor und nach Treatment gemessen

> Vortest-Nachtest-Design mit Kontrollgruppe

- Experimentalgruppe und Kontrollgruppe werden vorher und nachher gemessen

> Solomon-vier-Gruppen-Design

- wie Vortest-Nachtest-Design mit Kontrollgruppe verdoppelt, nur dass eine Kontrollgruppe und eine Experimentalgruppe nicht vorher gemessen werden

Feldstudie

- > weder randomisiert noch in Labor (keine Kontrolle der Störvariablen)
- > Bsp: 2 Schulklassen vergleichen
- > interne Validität = viel niedriger / externe Validität = hoch

→ nur das Laborexperiment hat vollen Anspruch auf Kausalität, die anderen nur bedingt

nichtexperimentelle Anordnungen

- > Ableitung von Kausalhypothesen, die durch Experimentvarianten geprüft werden

Ex-post-facto-Studie

- > aus vorgefundenen Daten wird auf Versuchsbedingungen geschlossen
- > keine zeitliche Sequenz der UV und AV mehr
- > keine Angaben zu Störfaktoren
- > niedrige interne Validität, kein Recht zur Kausalaussage
- > Längs- und Querschnittstudien möglich

vorexperimentelle Anordnung

- > kein echtes Experiment, denn UV wird nicht manipuliert (=Konstante)
- > für alle VP gleiche Situation (**Studie**)
- > Voruntersuchung für eigentliches Experiment > Aufbau austesten (Instruktion, Schwierigkeit..)

Erkundungsexperiment

- > kann alle oben genannten Formen annehmen
- > keine oder nur ungenaue Hypothesen
- > Ziel: Hypothesengenerierung, da nicht im gleichen Durchgang prüfbar

Versuchsplananlage (VPL – A)

- > Menge von bewährten Schemata zur Anordnung von UVn (**Faktoren**) und ihren Ausprägungen (**Faktorstufen**) =ANOVA?
- > Ziel: möglichst informationshaltige, ökonomische symbolische Repräsentation der Faktorstufen
- > Tabelle: Zeilenfaktor: A(hypothesenrelevant/Kontrollfaktor)/Spaltenfaktor: B(hypothesenrelevant)
- > **VPL-A1** Versuchsplananlage mit einem Faktor
- > **VPL-A2** Versuchsplananlage mit zwei Faktoren
 - hier sind Kombinationen möglich (**Zellen**)

Effekte

Haupteffekt (HE)

- > isolierte Wirkung eines Faktors
- > Mittelwerte der Faktorstufen miteinander vergleichen (Stufenvergleich in Anova)
 - über alle Stufen hinweg möglich

einfacher Haupteffekt (EHE)

- > kombinierte Wirkung
- > ein Faktor bleibt gleich, während anderer variiert (Zellvergleiche in Anova)

Interaktion

- > Wirkung, die nur bei bestimmten Kombinationen von A und B auftritt
 - grafisch: **Interaktionsplot**

Versuchsplan (VPL)

- > liefert fehlende Informationen zu Nachvollziehbarkeit des Experiments
- > Ziel: ökonomische, valide, präzise Hypothesenprüfung im Experiment

Merkmale

Anzahl realisierter Zellen

- > Zelle = Kombination zweier UVn
- > **vollständig gekreuzt** (alle realisiert)
- > unvollständig gekreuzt (nicht alle realisiert)

Anzahl Vpn pro Zelle

- > gleiche Anzahlen sind anzustreben (Auswertungsverfahren, Vergleichbarkeit)
- > Versuchspläne mit gleichen Zellhäufigkeiten = **ausbalanciert**

ein- oder mehrmalige Beobachtung einer VP

- > **interindividuelle** Bedingungssituation: jede VP nur in einer Zelle (1 Messwert pro Person)
- > **intraindividuelle** Bedingungssituation: jede VP in allen Kombinationen > Messwiederholung
 - präziser, da weniger Fehlerquellen

Art der Zuordnung (VP → Bedingung)

- > **randomisiert**: per Zufall → Bedingung für echtes Experiment
- > **nichtrandomisiert**: systematische Zuweisung

Arten von Versuchsplänen

1) echtes Experiment:

- > **VPL2RR** – 2UVn, beide randomisiert
- > **VPL1R** – 1UV, randomisiert

- > **VPLQ(W)R** – eine Uvn ist nur noch Qasiexperimntell, da Wiederholt gemessen (rmAnova)
→ trotzdem noch experimentell, wenn ausbalanciert wurde:
- **VPL2R(W)R** : 2UVn, eine randomisiert, eine randomisiert angenommen mit Messwiederholung

2) Quasiexperiment

- > **VPL2QR** – beide Uvn nicht randomisiert
 - > **organismische Uvn** können nicht randomisiert zugeordnet werden, da Vpn diese Merkmale bereits mitbringen (Geschlecht,..)
 - > interne Validität eingeschränkt
- es gibt **reine Pläne** (beide Uvn quasiexperimentell oder experimentell) und **Mischformen**
→ VPL-A zeigen prinzipiellen Aufbau, VPL ergänzen Anzahl und Zuordnung der Vpn

quantitative Einzelfallforschung

- > Analyse einer einzigen Person bezüglich einer oder mehrerer AV, um die Wirkung einer UV zu bewerten
- > möglichst kontrolliert
- > auch einzelne Gruppen möglich
- > Ziel: möglichst große **Verhaltensstichprobe**
 - möglichst große Zeitreihe (viele Messungen, zeitlich gleiche Abstände)
- > Ausgangspunkt: **Basisrate „base line“**
 - Bsp: Maßnahme X soll Ängstlichkeit verringern: Basisrate = vorher Ängstlichkeit messen
- > VP ist Versuchsperson und Kontrollperson gleichzeitig

Pläne

AB-Plan

- > Logik: wenn Therapie wirkt, wird sich Basisrate aufgrund der **Intervention** deutlich und stabil in ihrem Niveau verändern
- > Bsp: 10 Messungen für Baseline (A) verglichen mit 10 Messungen für Interventionszeitraum (B)
- > keine Kausalität nachweisbar (interne Validität - Störvariablen)

ABA-Plan

- > wie AB-Plan, nur noch eine unbehandelte Phase hintendran
- > interne Validität steigt (Störeffekte konfundieren relativ unwahrscheinlich in diesem Muster)

ABAB-Plan

- > noch eine Interventionsphase mehr
- > interne Validität sehr hoch
- > Übertragbarkeit: zusätzliche systematische Replikationen notwendig (mit anderen Patienten, Therapeuten, Indikatoren)

Kritik

- > Effektzurücknahme in 2. A-Phase (schlecht, wenn wieder Baseline → B verlängern)
- > unbehandelter Abbruch bei ABA (kann weitergemacht werden ohne Messung)
- > nur quasiexperimentell (Alternativerklärungen nicht ausgeschlossen → Replikationen)

Multiple Grundratenversuchspläne (MVG)

- > bei irreversiblen Lernprozessen sind diese Designs (außer AB) nicht möglich
 - > möglichst **zeitsynchron** werden **mind. 3 Interventionen** auf ein Individuum angewandt
 - > interne Validität = gesichert
 - > Rückfälle/Stabilität leicht erfassbar, keine Effektzurücknahme
- Bsp: Konfrontationstherapie nach traumatischem Erlebnis

- > VP soll sich 3 verwandte Szenen vorstellen und Grad der Störung wird gemessen
- > nacheinander werden alle 3 Szenen interveniert (alle 12 Wochen eine neue)
- > Vergleich mit nicht intervenierten Szenen: diese bleiben zunächst auf Basisratenniveau

Annahmen:

- > Unabhängigkeit: wenn Szene 1 interveniert wird, muss Szene 2 gleich gestört bleiben
- > Ähnlichkeit: gleiche Interventionen sollen bei allen Szenen zu gleichen Effekten führen
- wenn nicht: keine Kausalinterpretation

Generalisierbarkeit

- wie oben nur durch Replikationen erreichbar

Kriterien quantitativer Einzelfallforschung

statistische Signifikanz

- > Zustandekommen eines MW-Unterschieds/Zusammenhangs muss überzufällig sein
- > stichprobenabhängig

pädagogisch-klinische Signifikanz

- > Signifikanz = Bedeutsamkeit/Relevanz
- > 85% der Basisraten- und Interventionsphase sollen nicht überlappen (in Diagramm)
- > Unterschied der MW $\geq 2SD$

normative Validität

- > Übereinstimmung von Daten des Klienten nach Intervention mit Daten unauffälliger Personen
- > Ziel: Unauffälligkeit nach Therapie

soziale Validität

- > Meinungen von Vertrauten, Freunden.. zu Veränderung

- alle 4 Kriterien zusammen ergeben ausgewogenes Gesamturteil

nichtexperimentelle quantitative Forschungsmethoden

- > quantitative Forschungsmethoden ohne Kausalschlüsse: Beschreiben statt Erklären
- > Ziel: Zusammenhänge identifizieren und beschreiben = deskriptiver Ansatz

Korrelationsstudie

- > Zusammenhang zwischen 2 Variablen beschreiben (definieren, operationalisieren, messen)

Arten des Zusammenhangs:

- > positiv (gleichsinnig) vs negativ (gegensinnig)
- > kein Zusammenhang (0)
- > Abstufungen in Höhe

Vor-und Nachteile

- > keine Kausalität möglich
- > stabile Korrelationen erlauben gute Vorhersage des jeweils anderen Werts (für Durchschnittswerte, bei Individuen kann es Fehler geben)

mehr als 2 Variablen: Regression

- > Prädiktoreninterkorrelationen, Kriteriumskorrelation > Regression

Faktorenanalyse

- > aggregiert Items, die hoch korrelieren (Reduktion der Datenmenge), um begrenzte Anzahl an Faktoren für Vorhersage zu erhalten

DATENAUFBEREITUNG

- 1) Überführung in Zahlenform (zählen, kodieren)
- 2) Überführung in softwarekompatibles Datenformat
- 3) Organisation der Daten (Zusammenfassung einzelner Items durch Mittelwert mit Cronbachs Alpha; Behandlung von Ausreißern, Kodieren fehlender Werte, Transformationen von nicht normalverteilten Funktionen...)

Deskriptivstatistische Methoden

- > Stichprobe zusammenfassend beschreiben und darstellen
- > Aussagen über mehrere/alle Fälle hinweg treffen
- > nicht erklärend

Analyse von Häufigkeiten

- 1) Festlegung der **Kategorien** und Kategoriebreiten
 - > bei diskreten Variablen einfach (nur Breiten, wenn zu viele)
 - > bei stetigen Variablen komplexer (Breite berechnen, da sonst fast nie etwas doppelt vorkommt)
 - > alle Kategorien sollten gleich breit sein
- 2) **absolute, relative, prozentuale Häufigkeiten** berechnen
- 3) **kumulative Häufigkeiten** und **Prozentränge**
 - > ab Ordinalskala: kumulierte absolute/relative Häufigkeiten (kleiner/gleich)
 - > relative Häufigkeiten mit 100 multipliziert: Prozentränge (Quantile)
- 4) Veranschaulichung von Häufigkeitsverteilungen
 - > **Histogramm/Säulendiagramm**: Abszisse (x) = Kategoriegrenzen / Ordinate (y) = Häufigkeiten
 - > **Kreidiagramm**: Anteile im Verhältnis zur Gesamtheit darstellen

Univariate Deskriptivstatistik

Maße der zentralen Tendenz

- > Tendenz, die alle Messwerte mehr oder weniger gut beschreibt

Modalwert = häufigster Wert der Verteilung

- > unimodal/bimodal/multimodal
- > für alle Skalenniveaus

Median = Wert, der geordnete Rangfolge in zwei fast gleich große Hälften teilt

- > ab Ordinalskala

Arithmetisches Mittel = Mittelwert, Schwerpunkt der Verteilung

- > stärker ausreißerempfindlich, da Werte miteinbezogen
- > ab Intervallskala

Schiefe

- > wie weit weicht die Verteilung von symmetrischer Form ab?
- > links-/rechtssteil, symmetrisch

Exzess

- > ausgehend von Normalverteilung wird Verteilung als schmal- oder breitgipflig bezeichnet
- beides macht nur bei unimodalen Verteilungen Sinn

Streuungsmaße

- > wie stark unterscheiden sich verschiedene Werte in der Stichprobe voneinander?
- > Grad der Streuung bestimmt die Aussagekraft des Mittelwerts (viel Streuung = wenig)
- > neue nomothetische Forschungsfragen: warum ist Streuung so groß?...

Varianz (SD^2)

- > Summe aller quadrierten Abweichungen vom Mittelwert geteilt durch Anzahl
- > Summe der Abweichungsquadrate, relativiert an Stichprobengröße

Standardabweichung (SD)

- > Wurzel aus Varianz
- > besser vergleichbar und interpretierbar

weitere Streuungsmaße:

- > **Interquartilsabstand** (ab Ordinal)
- > **Spannbreite** (ab Intervall)

Multivariate Deskriptivstatistik

Korrelation

- > 2 Merkmale werden in Beziehung gesetzt
- > Korrelationskoeffizient r (+1 bis -1) – Richtung und Stärke des Zusammenhangs
- > Relation: Streuung von x und y Werten vom MW ausgehend, relativiert an den SDs
- > nie kausal bewerten, außer man weiß, dass Bedingungen erfüllt waren!

Regression

- > Instrument zur Prognose des Kriteriums (y) durch einen oder mehrere Prädiktoren ($x_1...x_k$) ohne Kausalität (außer Bedingungen waren im Versuch erfüllt)
- > Stärke des Zusammenhangs durch absoluten Regressionskoeffizienten b bzw relative Gewichtung durch Betakoeffizienten erfassbar
- > Beta ähnlich wie r : gibt Richtung an / anders: Höhe = wie viele SD verschiebt sich y wenn sich x_i um eine SD verschiebt
- > wenn nur ein Prädiktor: $\beta = r$
- > häufig mit mehreren Prädiktoren (**multiple Regression**), da Merkmale oft komplex
- > Korrelation unter den Prädiktoren wird rausgerechnet
- > R (multipler Korrelationskoeffizient) statt r : schwankt zwischen 0 und 1

Inferenzstatistische Methoden

- > Schluss von einer Stichprobe auf zugehörige Population
 - Unsicherheit: alles beruht auf Wahrscheinlichkeit (quasiuniverselle Hypothesen)
 - Parameter in Population werden durch Stichprobe geschätzt
 - Sicherheit der Vorhersage wird durch **Konfidenzintervall** ausgedrückt
- > griechische Buchstaben für Parameter

Signifikanztest

Test von H_0 und H_1

- > Instrument zur Prüfung von Hypothesen zu nicht beobachteten Populationen
 - H_0 = Null-Zusammenhang / Gleichheit der Parameter

- H_1 = Unterschied zwischen den Parametern, Vorliegen eines Zusammenhangs
- > Signifikanztest prüft, mit welcher WS man das beobachtete Ergebnis erhält, wenn H_0 in der Population gilt
- > ist die Überschreitungswahrscheinlichkeit sehr gering (5%; 1%), dann wird H_0 verworfen und H_1 angenommen
- > abhängig von Stichprobengröße und Streuung in der Population
- Fehlinterpretationen:** Signifikanz ist nicht gleich:
 - > WS des gefundenen Ergebnisses
 - > komplementäre WS zur Alternativhypothese
 - > WS der Nullhypothese
- Signifikanz = WS des gefundenen Ergebnisses bei unterstellter Gültigkeit von H_0
- Effektstärke**
 - > bestimmen, denn jede H_0 kann bei ausreichend großen Stichproben abgelehnt werden
 - > angegeben durch Cohen's d (Größe des Unterschieds standardisiert)

Testverfahren

t-Test

- > Wahrscheinlichkeit, dass bei unterstellter Gültigkeit von H_0 unter Kenntnis des Populations-MW der MW der Stichprobe zustande kommt; Vergleich mit Signifikanzniveau Alpha
- > entweder für eine Stichprobe oder für 2
- > Vergleich von 2 (un)abhängigen Stichproben
 - abhängige Stichprobe: Messwiederholung, Dyaden unter versch. Bedingungen vergleichen
 - unabhängige Stichproben: verschiedene VP unter verschiedenen Bedingungen verglichen

Varianzanalyse (ANOVA)

- > Tests zu Unterschieden zwischen mehr als 2 Stichproben
- > Zerlegung der Streuung: Gesamtvarianz = Primärvarianz (aufgeklärt) + Sekundärvarianz (Fehler)
- > Streuung einer AV wird auf Faktorstufen einer/mehrerer UV und Fehlerstreuung zurückgeführt
- > abhängig bei rmANOVA (Messwiederholung)
- > einfaktoruell/mehrfaktoruell (bezogen auf UV)
- > univariat/multivariat (bezogen auf AV)
- > Haupteffekte, einfache Haupteffekte, Interaktionseffekte (siehe VPL-A)
- > Überschreitungs-WS $p < \alpha$ = Effekt signifikant

WISSENSCHAFTLICHE QUELLEN & ZITIEREN

Arten wissenschaftliche Quellen

- > Enzyklopädien
- > Handbücher
- > Lehrbücher
- > Monografien (1 Autor)
- > Herausgeberbände (1 oder mehrere Herausgeber)
- zusammengefasste Darstellung eines oder mehrerer Inhaltsgebiete → Orientierung
- > Dissertatation (herkömmlich/kumulativ)
- > Diplomarbeit (herkömmlich/im Stil wiss. Artikel)
- > Bachelor-/Masterarbeiten (herkömmlich/im Stil wiss. Artikel)
- Darstellung von Studien, Zusammenfassung aktueller Literatur, Vorgehensweise, form. Struktur

> Artikel in wissenschaftlicher Zeitschrift

- Studie

- Metaanalyse und Reviews

> Metaanalyse = Analyse von mehreren schon vorhandenen Analysen

> Review?

→ Darstellung von Studien, form. Struktur, Qualitätssicherung durch peer-review

Schema wissenschaftlicher Artikel

1) Titel

> alleine verständlich, erwähnt Variablen/Frage

2) Abstract

> alle wichtigen Informationen – verständlich, knapp, objektiv, genau, (max. 10 Sätze)

3) Introduction

> Einführung in Problem, Überblick über Forschungsstand, Hypothesen

> Literaturreview: Sichtweisen darstellen, Vorwissen, Wesentliches aus Studien, Vergleiche..

4) Methods

> Stichprobe, Material, Versuchsablauf → Replizierbarkeit gewährleisten: Genauigkeit

5) Results

> erhobene Daten, statistische Vorgehensweise, keine inhaltliche Bewertung – nur Beschreiben, (un)beständige Hypothesen nennen

6) Discussion

> Bewertung und Interpretation der Ergebnisse, Vergleiche, Kritik, Ausblick

7) References (Literaturverzeichnis)

> alle im Text verwendete Quellen