

# Einführung in die Forschungsmethoden der Psychologie und Psychotherapie

## Einheit 7: Kausalität, Experiment und Studiendesigns

12.06.2024 | Dr. Caroline Zygar-Hoffmann

# Heutige Themen

Ziele von statistischen Analysen und der Versuchsplanung

Kausalität

Störvariablen und deren Kontrolle

Versuchspläne

Zusammenfassung: Forschungsstrategien

**Take-Aways und Schlüssel-/Fachbegriffe**

- Take-Aways
- Schlüssel-/Fachbegriffe

# Ziele von statistischen Analysen (Auswahl)

## Ausprägungen deskriptiv beschreiben oder statistisch schätzen:

- Forschungsfragen zur Ausprägung einzelner Variablen (z.B. Häufigkeiten, Mittelwerte)
- Wenn komplette Population erhoben wird (-> Deskriptivstatistik), wenn von Stichprobe auf Population geschlossen wird (-> Parameterschätzung)
- Bsp.: Wie viel % aller Jugendlichen in Deutschland spielen täglich Computerspiele?

## Korrelationsfragestellungen statistisch testen:

- Zusammenhänge zwischen Merkmalen anhand korrelativer Versuchspläne aufdecken
- Bsp.: Steht die Häufigkeit des Computerspielens bei Jugendlichen in einem positiven Zusammenhang mit ihrer Gewaltbereitschaft?

## Kausalfragestellungen statistisch testen:

- Ursache-Wirkungs-Beziehungen (meist) anhand experimenteller Versuchspläne aufdecken
- Bsp.: Erhöht Computerspielen die Gewaltbereitschaft von Jugendlichen?

# Ziel der Versuchsplanung (Studiendesign)

Das Ziel der Versuchsplanung ist eine empirische Untersuchung zu planen, sodass eine Hypothese valide & ökonomisch überprüft werden kann

- Validität im Kontext der Versuchsplanung: Gültigkeit oder allgemeine Güte der Hypothesenprüfung
- Ökonomie im Kontext der Versuchsplanung: mit möglichst geringem Aufwand

## **Leitfrage:**

Wie muss ich eine Untersuchung gestalten, damit sie mir neue Erkenntnisse über die Gültigkeit meiner Hypothese liefern kann?

→ Besonders knifflig, wenn man an Kausalfragestellungen interessiert ist

Voraussetzungen für kausale Schlussfolgerungen:

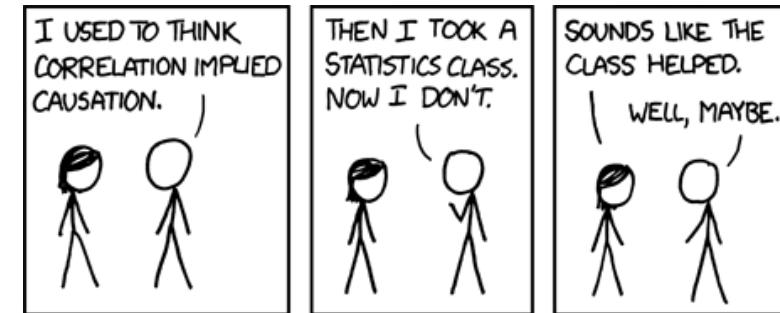
- 1) **Kovariation/Korrelation:** Variation in UV muss mit Variation in AV einhergehen
  - 2) **Zeitliche Vorgeordnetheit der UV:** Veränderungen auf UV zeitlich vor Veränderungen der AV
  - 3) **Ausschluss von Alternativerklärungen / Kontrolle von Störvariablen**
- **Korrelation alleine  $\neq$  Kausalität** z.B. Die Sonne geht nicht auf, weil der Hahn kräht.

## Korrelation =/= Kausalität

- Sachverhalt: Zusammenhang zwischen A & B
- Anders ausgedrückt: A wird häufig zusammen mit B beobachtet

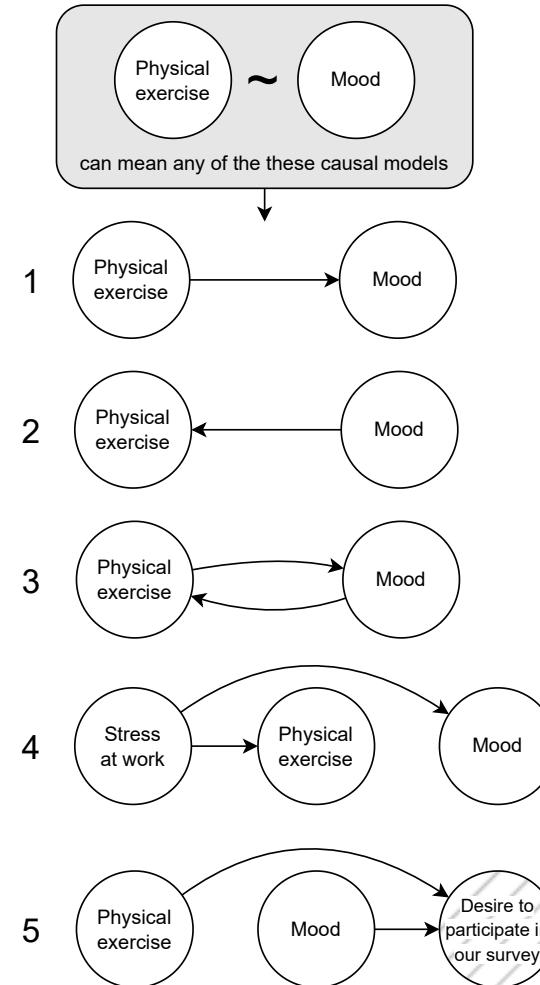
### Mögliche Kausalbeziehungen:

1. A verursacht B
2. B verursacht A
3. Eine Drittvariable C (bzw. eine Menge von zusammenhängenden Variablen X<sub>1</sub>, ..., X<sub>n</sub>) verursacht A & B



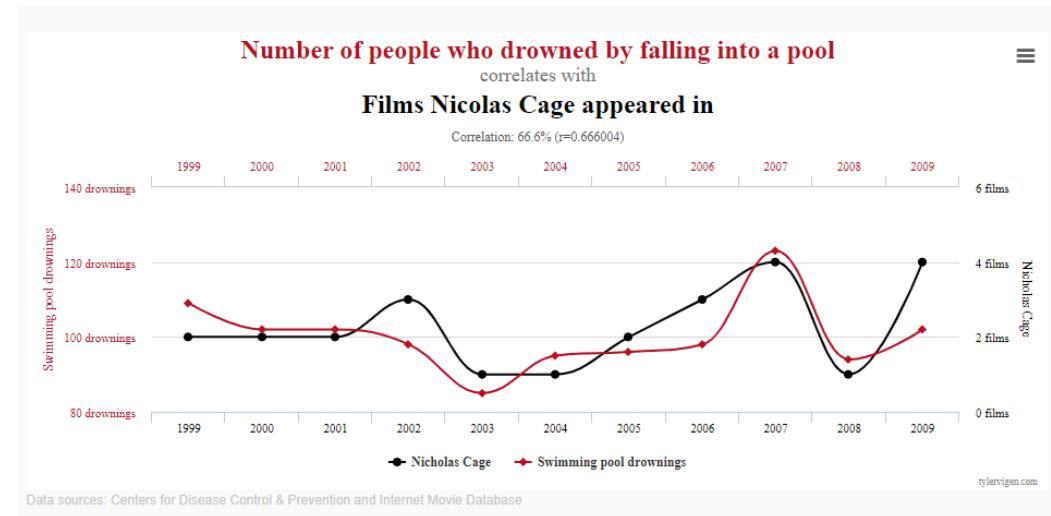
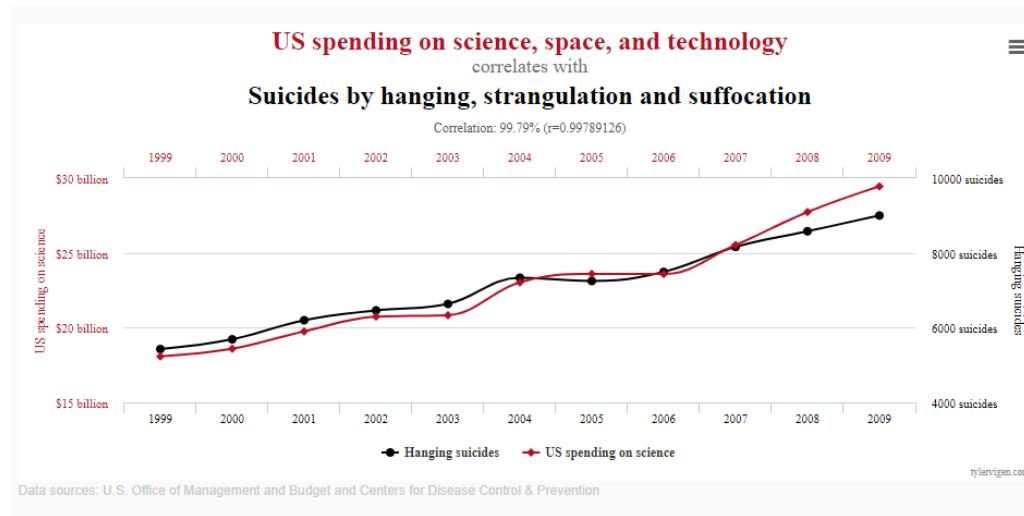
**Folge:** Korrelation impliziert nicht Kausalität

**Kausale Modelle, die eine Korrelation erzeugen können:**



## "Spurious Correlations" = "Scheinkorrelationen"

zufällige Korrelationen, die nicht auf eine Kausalität zwischen zwei Variablen zurückzuführen sind



## Interventionistische Auffassung von Kausalität

*"...to think of a relation between events as causal is to think of it under the aspect of (possible) action. [...] that p is the cause of q [...] means that I could bring about q if I could do (so that) p.*

von Wright, 1971

*The paradigmatic assertion in causal relationships is that manipulation of a cause will result in the manipulation of an effect. [...] Causation implies that by varying one factor I can make another factor vary.*

Cook & Campbell, 1979

Methodologische Implikation: **aktive Manipulation** der interessierenden Variablen

## X verursacht Y (INUS-Bedingungen)

Ursachen sind sogenannte INUS-Bedingungen:

- I = Insufficient, but
- N = Necessary parts of an
- U = Unnecessary, but
- S = Sufficient condition

Bsp.: Waldbrand verursacht durch Streichholz.

- Insufficient: Es braucht auch (z.B.) Sauerstoff
- Necessary: Eine Flammenquelle wird benötigt
- Unnecessary: Feuerzeug geht auch
- Sufficient: Sauerstoff, Streichholz, Trockenheit zusammen hinreichend

## X verursacht Y (INUS-Bedingungen)

Ursachen sind sogenannte INUS-Bedingungen (Insufficient, but necessary parts of an Unnecessary, but Sufficient Condition):

### Konzeptuelle Implikationen:

- Ursachen sind nur als Teil einer Menge von ermögliche Randbedingungen wirksam → Ursachen können auch nur unter bestimmten Bedingungen wirksam sein
- Konkret bedeutet dies trotzdem: Hätte es die Ursache *nicht* gegeben (und alles andere wäre gleich geblieben), wäre es nicht zur Wirkung gekommen
- Theorien können sich auf unterschiedliche Teile der Bedingungsmenge beziehen
- Eine Wirkung kann auch durch andere Ursachen entstehen

**Methodologische Implikationen:** Manipulation der interessierenden Variablen (= UVs), Konstanthaltung aller anderen Bedingungen/Faktoren → Experiment als Idealmodell für die Überprüfung von Kausalhypothesen

## Experiment - Kernelemente

### 1) Kausalhypothese

- Vermutung über den Einfluss einer Variable (UV) auf eine andere Variable (AV)
- z.B. „Ego\*Shooter verursachen erhöhte Gewaltbereitschaft“

### 2) Manipulation einer UV

- 2 (oder mehr) Bedingungen herstellen, die sich nur bzgl. einer einzigen Variable (UV) unterscheiden
- z.B. Ego-Shooter spielen vs. Flugsimulation spielen

### 3) Randomisierung

- jede Person (Untersuchungseinheit) wird per Zufall einer der Bedingungen zugewiesen
- z.B. per Münzwurf entscheiden, was eine Person spielen muss

### 4) Beobachtung (Messung) einer AV

- Ausprägung der interessierenden AV wird gemessen
- z.B. das Aggressivitätsniveau der Personen

## Störvariablen und Konfundierung

### Störvariablen (SV)

- Definition: alle Variablen (außer der UVs), die potentiell Einfluss auf das Ergebnis haben können → Variablen die einen Zusammenhang mit der AV aufweisen
- Störvariablen sind besonders problematisch, wenn sie mit UVs assoziiert (konfundiert) = systematisch sind

### Konfundierung

- gemeinsame Variation einer vermuteten Ursache (UV) mit (mindestens) einer anderen Störvariable
- Störvariable kann als Ursache für den beobachteten Sachverhalt nicht ausgeschlossen werden → UV kann nicht als Ursache interpretiert werden!

### wichtiges Ziel der Versuchsplanung (nach Hager, 1987)

- gemeinsame systematische Variation von möglichen Störfaktoren mit hypothesenrelevanter UV verhindern
- bzw. statistische Assoziation zwischen den potentiellen Störfaktoren & der UV auf den Wert 0 bringen
- in dem Ausmaß, in dem dies für einen der möglichen Störfaktoren gelingt, nennen wir diesen kontrolliert

## Experiment als Versuch zur Kontrolle von Störvariablen

### Methode: Experiment

- Vergleich von Bedingungskonstellationen, die sich nur im Hinblick auf das Vorhandensein der vermuteten Ursache (UV = unabhängige Variable) unterscheiden → Aktive Manipulation der UV
- Konstanthaltung anderer Faktoren (Kontrolle von Störvariablen), u.a. durch Randomisierung
- Beobachtung, ob der zu erklärende Sachverhalt (AV = abhängige Variable) eintritt oder nicht

### Kausale Interpretation

- Ergebnis: (k)ein Effekt in der AV (Unterschied zwischen 2 oder mehr Bedingungen)
- Kausale Interpretation: UV ist (k)eine Ursache für den Effekt auf die AV

## Systematische Störvariablen

- kovariieren mit UV (d.h. in verschiedenen Versuchsbedingungen unterschiedlich stark ausgeprägt) → führen zu systematischen Unterschieden zwischen den Bedingungen
- können fälschlicherweise einen Effekt der UV auf die AV suggerieren → bieten Alternativerklärungen warum man einen vermeintlichen Effekt von der UV auf die AV beobachtet:

→ Beispiel: Trainierte haben eine bessere Leistung (AV), aber nicht wegen Training (UV), sondern wegen höherer Motivation (SV)

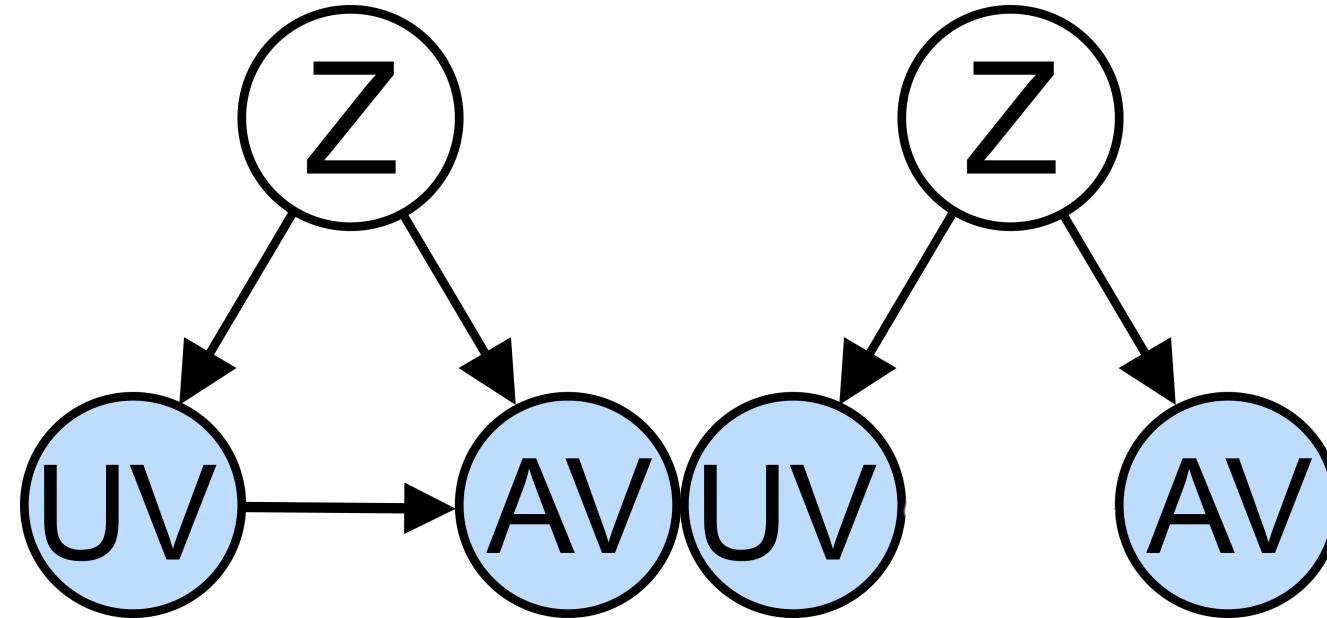
- können einen tatsächlich vorhandenen Effekt verschleiern:

→ Beispiel: überwiegend unmotivierte Versuchspersonen in der Trainingsbedingung: positiver Effekt des Trainings (UV) wird durch negativen Effekt der Motivation (SV) verdeckt

## Konsequenz systematischer Störvariablen

Falsche Schlussfolgerungen hinsichtlich des Effekts der UV auf die AV sind möglich, wenn Störvariable nicht berücksichtigt wird.

## Systematische Störvariablen



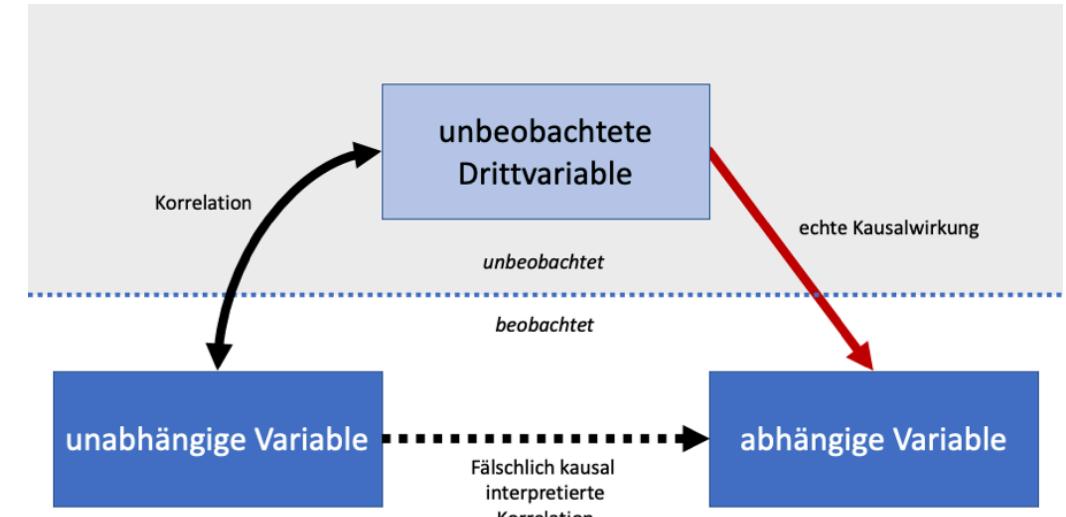
→ Werden systematische Störvariablen nicht kontrolliert spricht man von Konfundierung, die Störvariable ist dann eine konfundierende Variable

# Störvariablen und deren Kontrolle

## Systematische Störvariablen

Beispiel:

- Untersuchung von Mathenachhilfe oder keine Mathenachhilfe (UV) auf Matheleistung (AV)
- Systematische Störvariable "Interesse an Mathe" einer Person hängt damit zusammen, ob Personen sich für die Nachhilfe anmelden (d.h. in welcher Bedingung sie landen -> Zusammenhang mit UV) als auch welche Leistung sie zeigen (d.h. Zusammenhang mit AV)



- Wenden wir keine Randomisierung an, dann ist "Interesse an Mathe" eine konfundierende Variable
- Wenden wir Randomisierung an, ist die systematische Störvariable kontrolliert und keine konfundierende Variable mehr

## Unsystematische Störvariablen

- kovariieren nicht mit der UV (d.h. in allen Versuchsbedingungen ungefähr gleich stark ausgeprägt) → führen zu *keinen* systematischen Unterschieden *zwischen* Bedingungen, aber erklären generelle Unterschiede auf der AV *innerhalb* von Bedingungen
- erklären unter anderem warum nicht alle Personen denselben Wert auf der AV haben → gibt es in jeder Untersuchung; vergrößern die Varianz in der AV (Fehlervarianz; „Rauschen“; geringe Reliabilität der Messung)
- können systematische Effekte der UV durch hohe Fehlervarianz überdecken
- können das Auftreten von Effekten nicht erklären

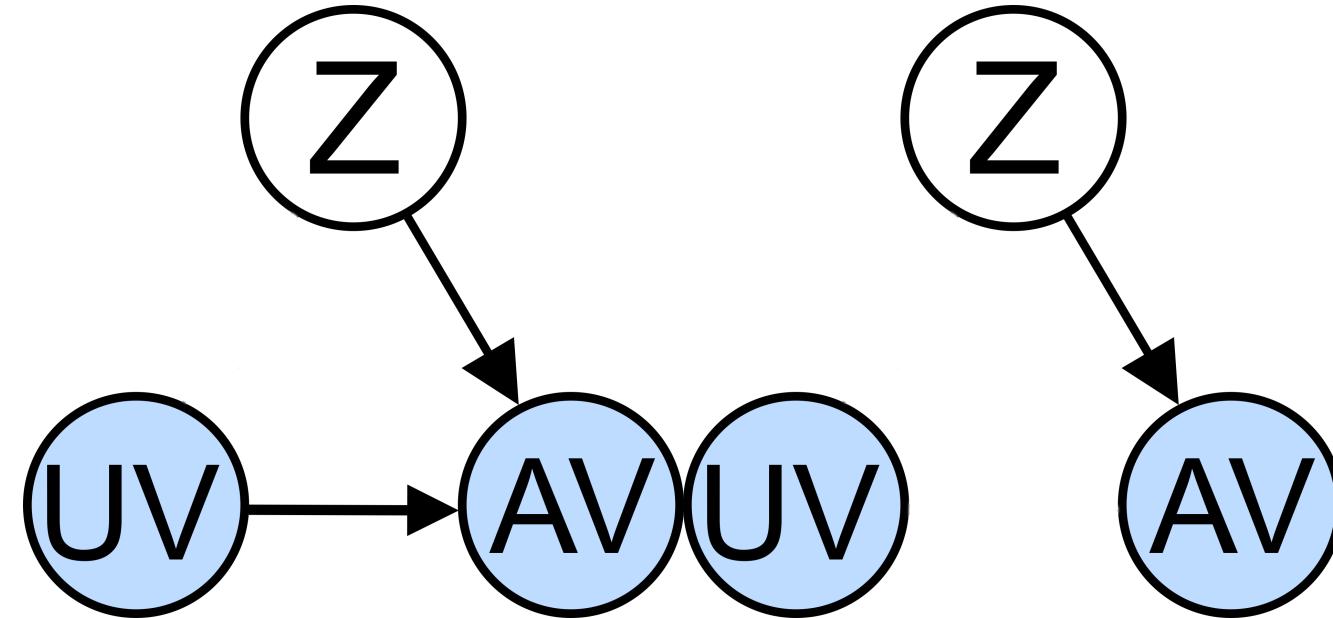
→ Beispiel: in beiden Veruchsbedingungen gibt es interindividuelle Unterschiede in der Sorgfalt der Beantwortung des Fragebogens mit dem die AV gemessen wird

## Konsequenz unsystematischer Störvariablen:

Wenn kein Effekt beobachtet wird, kann dies an der erhöhten Fehlervarianz liegen.

Aber: Wenn ein Effekt beobachtet wird, kann die unsystematische SV **nicht** dafür verantwortlich sein.

## Unsystematische Störvariablen



## Interne und statistische Validität

### Ziele in der Versuchsplanung:

#### 1. Kontrolle systematischer Störvariablen

- die gemeinsame systematische Variation von möglichen Störfaktoren mit der hypothesenrelevanten UV zu verhindern
- die statistische Assoziation zwischen den potentiellen Störfaktoren und der UV auf den Wert Null zu bringen  
→ erhöht sogenannte "interne Validität"

#### 2. Reduktion unsystematischer/zufälliger Störvariablen

- Reduktion der Fehlervarianz erhöht den Anteil der Effektvarianz an der Gesamtvarianz  
→ erhöht sogenannte "statistische Validität"

## Quellen von Störvariablen

- **Teilnehmer:innen:**
  - Geschlecht, Intelligenz, Einkommen, Ängstlichkeit, Sucht, Haustier,...,
  - alle Eigenschaften mit potentiellem (direktem oder indirektem) Einfluss auf die AV
- **Versuchsleiter:innen:**
  - Geschlecht, Alter, Autorität, Status, Attraktivität, Strenge,...
  - Erwartungen bzgl. der Untersuchungsergebnisse...
- **Situation:**
  - Lärm, Beleuchtung, Tageszeit, Jahreszeit, Konjunktur, Publikum
  - Messinstrument, Reaktivität der Messung, demand characteristics (Aufforderungscharakter der Situation)
- **Reihenfolge/Messwiederholung**
  - Positionseffekte
  - Übungs-, Erinnerungs-, Ermüdungseffekte
  - Sensibilisierungseffekte

# Störvariablen und deren Kontrolle

## Kontrolltechniken

Quelle	Technik	Systematische Fehler	Unsystematische Fehler
Teilnehmer:innen	Randomisierung	++	
	Parallelisierung	+	
Versuchsleiter:innen	Standardisierung	+	+
	Versuchleiter-Training	+	+
	Automatisierung	+	+
	Verblindung	+	
Situation	Konstanthaltung	+	+
	Elimination	+	+
	Kontrollfaktoren	+	
	Täuschung	+	
Messwiederholung	Ausbalancieren	+	

## Randomisierung

### Definition:

- zufällige Zuweisung der Untersuchungseinheiten zu Bedingungen
- Technik zur Kontrolle in der Person liegender, unveränderlicher Störvariablen
- vorhandene Störvariablen werden nach Zufall (gleichmäßig) auf die Gruppen verteilt

### Ergebnis der Randomisierung:

- personenbezogene Störvariablen sind nicht mit Bedingungszugehörigkeit konfundiert (bis auf Zufallsschwankungen)
- Definition Experiment: Wenn bzgl. einer UV randomisiert wird, dann ist Untersuchung bzgl. dieser UV ein Experiment
- Bedrohung der internen Validität durch Störeinflüsse wird verringert
- Es gibt verschiedene Varianten der Randomisierung (vom einfachen "Münzwurf", über "Block-Randomisierung" zu Erreichung gleicher Gruppengrößen bis zu "stratifizierter Randomisierung" zu Erreichung einer möglichst großen Ähnlichkeit der Bedingungen hinsichtlich einer Drittvariable)

## Randomisierung

**Randomisierung  $\neq$  Zufallsstichprobe**

### Zufallsstichprobe:

- zufällige Auswahl aus der Population
- Jeder hat die gleiche Chance, ausgewählt zu werden
- erhöht Ähnlichkeit von Stichprobe und Population
- führt zu sogenannter "externer Validität"

### Randomisierung:

- zufällige Zuweisung bereits ausgewählter Probanden zu Bedingungen/Gruppen
- Jeder hat die gleiche Chance, in jede Bedingung zu gelangen
- erhöht Ähnlichkeit der verglichenen Bedingungen/Gruppen
- führt zu interner Validität

## Parallelisierung bzw. Matching

- Vergleichbarkeit der Gruppen bzgl. einer bekannten Störvariable herstellen
- Bei kleinen Stichproben zuverlässiger als Randomisierung – aber nur bzgl. einer SV!

### Umsetzung:

- SV bei allen Teilnehmern des Experiments erfassen
- Rangreihe bzgl. der SV bilden
- jeweils benachbarte Rangplätze werden per Zufall auf die Bedingungen aufgeteilt

### Voraussetzungen:

- reliable & valide Messbarkeit der zu kontrollierenden SV
- Verfügbarkeit der gesamten Stichprobe zur Erfassung der SV vor der eigentlichen Untersuchung
- theoretische und/oder empirische Begründung der Bedeutsamkeit der SV

## Standardisierung, Versuchsleiter-Training, Automatisierung

### Versuchsablauf standardisieren:

- Präzises und detailliertes Ablaufprotokoll festlegen
- Systematische & unsystematische Störeinflüsse reduzieren, die durch Unterschiede im Ablauf entstehen können (z.B. unterschiedliche Erläuterungen/Instruktionen des Versuchsleiters)

### Versuchsleiter trainieren:

- Einüben des Ablaufs (idealerweise anhand eines Ablaufprotokolls)
- Systematische & unsystematische Störeinflüsse reduzieren, die durch fehlerhafte Durchführung des Versuchs entstehen können

### Versuchsablauf automatisieren:

- Ersetzen des Versuchsleiters durch z.B. Computer

## Verblindung

- Verblindung: Information über Versuchsbedingung vorenthalten
- Verhindert systematische Effekte dieser Information (z.B. über Erwartungen der Versuchsperson oder des Versuchsleiters)

## 3 Varianten:

- Einfache Verblindung: Versuchsperson hat keine Kenntnis über die Versuchsbedingung, der sie zugeordnet ist
- Doppelte Verblindung: Versuchsperson & Versuchsleiter haben keine Kenntnis über Versuchsbedingung
- Dreifache Verblindung: Versuchsperson, Versuchsleiter & Auswerter haben keine Kenntnis über Versuchsbedingung

## Konstanthaltung

- Ausprägung der Störvariable in allen Bedingungen gleichhalten
- Verhindert systematische & unsystematische Störeinflüsse
  - z.B. Kontext: alle Versuchsbedingungen in gleicher Umgebung durchführen
  - z.B. Temperatur: identisch klimatisierte Laborräume
  - z.B. Instruktion: identische Wortwahl in allen Versuchsbedingungen

## Elimination

- Einfluss der Störvariablen komplett verhindern
- Verhindert systematische & unsystematische Störeinflüsse
  - z.B. Lärm: Schallisolierung
  - z.B. Licht: Fenster abdunkeln
  - z.B. Anwesenheit anderer Personen: individuelle Datenerhebung
  - z.B. Versuchsleiter: Automatisierung

## Kontrollfaktoren

- Einbeziehen der Störvariable in das Untersuchungsdesign als Kontrollfaktor
- Effekte der UV & der SV können analysiert werden
  - z.B. UV Frustration: 2 Stufen (frustriert, nicht frustriert)
  - z.B. SV Tageszeit als Kontrollfaktor: 2 Stufen (vor vs. nach Mittagessen)
- Untersuchung mit 4 Gruppen:
  1. frustriert & vor dem Mittagessen
  2. nicht frustriert & vor dem Mittagessen
  3. frustriert & nach dem Mittagessen
  4. nicht frustriert & nach dem Mittagessen
- Unterschied 1 vs. 2 → Frustration
- Unterschied 1 vs. 3 → Tageszeit

## Täuschung

- Versuchspersonen wollen Versuchsleitern „einen Gefallen tun“, indem sie der (wahren/vermuteten) Hypothese entsprechend reagieren → AV nicht valide, evtl. systematisch verfälscht
- Fehlinformation über einzelne Aspekte des Versuchs
- Verhindert systematische Effekte einzelner Aspekte der Situation
- z.B. UV: Geschlecht des Versuchsleiters, Täuschung: UV verschweigen; „wir messen Kreativität“ → lenkt von UV ab
- z.B. AV: Leistungstest, Täuschung: „Pilotversuch, Daten werden nicht gespeichert“ → reduziert Einfluss der Prüfungsängstlichkeit
- besonders hilfreich bei demand characteristics: Aspekte der Situation, die (a) die wahren Hypothesen der Untersuchung verraten oder (b) falsche Hypothesen nahelegen
- Problem: Frage der ethischen Vertretbarkeit von Täuschung

## Ausbalancieren

- Reihenfolgeeffekte kontrollieren: alle möglichen Reihenfolgen realisieren; auf Effekte der Reihenfolge prüfen
- Vermeiden von Positions- & Sequenzeffekten bei Messwiederholung (Lernen, Ermüdung, Carry-Over-Effekte)
- Beispiel: Effekte von Lärm (UV) auf kognitive Leistungen (AV: Konzentrationstest)
  - Versuchsablauf A: (1) Test mit Lärm; (2) Test ohne Lärm
  - Ergebnis: Leistung (2) > Leistung (1) → Interpretation? Lerneffekt?

→ Ausbalancieren: zusätzlicher Versuchsablauf

- Versuchsablauf B: (1) Test ohne Lärm; (2) Test mit Lärm

Nachteil:

- Aufwand
- Bei  $k$  Versuchsbedingungen sind  $k!$  Reihenfolgen notwendig ( $2! = 2$ ;  $3! = 6$ ;  $4! = 24$ ; ...;  $10! = 3.628.800$ )

## Notation von Versuchsplänen

- O (Observation): Beobachtung, Messung einer oder mehrerer Maße; durchnummiert bei mehrfachen Messungen (z.B. Messzeitpunkte, Gruppen)
- X (Treatment, Intervention): kontrollierte Manipulation der UV
- R (Randomisierung): zufällige Zuweisung der Untersuchungseinheiten zu Bedingungen (vor Untersuchung)
- Zeitverlauf von links nach rechts
- verschiedene Gruppen sind zeilenweise untereinander notiert, z.B. oben EG & unten KG

## Nicht-Experimentelle Versuchspläne

„nur“ systematische Beobachtung einer/mehrerer Variablen

- keine Manipulation
- keine randomisierte Zuweisung zu Bedingungen

## Typische Versuchspläne:

- One-shot-design
- Korrelationsstudien
- Einfache Gruppenvergleiche
- Prä-Post-Vergleich

→ z.B. deskriptive Untersuchungen, Umfrageforschung (univariat); Korrelationsforschung (bi\* oder multivariat)

- Interpretationsmöglichkeiten: Prüfung von Kausalhypothesen nicht (oder nur unter speziellen Annahmen) möglich!

# Versuchspläne

## Nicht-Experimentelle Versuchspläne

### One-shot Design



## Nicht-Experimentelle Versuchspläne

### One-shot Design

- Beobachtung einer AV an einer Stichprobe (mit/ohne Intervention)
- „nur“ systematische Beobachtung einer Variablen; d.h. rein deskriptive Erhebung oder Schätzung des Ist-Zustands
- nur 1 Variable (keine Unterscheidung UV/AV, keine Manipulation)
- nur 1 Gruppe (keine Kontrollgruppe, keine Randomisierung)
- z.B. Umfrageforschung, Studienreform (X) & Messung Studierendenzufriedenheit (O); Spendenkampagne (X) & Messung Spendenaufkommen (O)

### Interpretationsmöglichkeiten:

- beschreibende Aussagen über Häufigkeiten oder Merkmalsverteilungen zum Zeitpunkt der Messung
- Effekt des Treatments nicht quantifizierbar (kein Vergleichswert)
- Zusammenhang X & O kann nicht untersucht werden

## Nicht-Experimentelle Versuchspläne

### Korrelationsstudien

- Beobachtung von 2 Variablen in einer Stichprobe
- keine Kontrollgruppe, keine Randomisierung, keine Trennung von UV-AV, Simultane Erhebung der Variablen, keine Manipulation (z.B. Geschäftserfolg & Extraversion)

Bei theoretischer Trennung von UV-AV:

- Simultane Erhebung der Variablen; keine Manipulation
- aber: theoretisch angenommene Kausalrichtung (z.B. je extrovertierter, desto erfolgreicher)

Bei „UV-AV-Sequenz“ (Ex-post-facto-Studien):

- Simultane Erhebung der Variablen; keine Manipulation
- Kausalitätsheuristik: Erhebung zeitlich zurückliegender „UV“ (AV folgt UV)

### Interpretationsmöglichkeiten:

- Aussagen über Zusammenhang möglich
- Aussagen über Kausalität nicht (oder nur unter speziellen Annahmen) möglich

# Versuchspläne

## Nicht-Experimentelle Versuchspläne

### Einfacher Gruppenvergleich



## Nicht-Experimentelle Versuchspläne

### Einfacher Gruppenvergleich

- Beobachtung einer AV in 2 (durch kategoriale UV definierten) Gruppen
- keine Manipulation, keine Randomisierung, UV-AV-Sequenz
- z.B. UV Geschlecht, AV Aggression; Ergebnis: Geschlechtsunterschied

### Problem:

Was ist für den Unterschied verantwortlich?

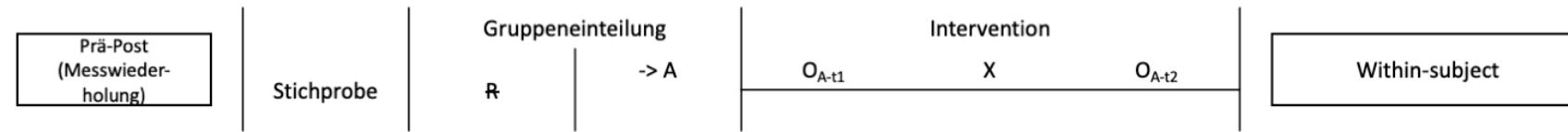
### Interpretationsmöglichkeiten:

- Aussagen über Zusammenhang Gruppe (UV) & AV
- Aussagen über Kausalität nicht (oder nur unter speziellen Annahmen) möglich

# Versuchspläne

## Nicht-Experimentelle Versuchspläne

### Prä-Post-Vergleich („vorexperimentelle Anordnung“)



# Versuchspläne

## Nicht-Experimentelle Versuchspläne

### Prä-Post-Vergleich („vorexperimentelle Anordnung“)

- Beobachtung einer AV in einer Stichprobe vor und nach einer Intervention
- keine Kontrollgruppe, keine zufällige Zuordnung
- z.B. Studierendenbefragung vor und nach einer Studienreform

### Probleme:

- alle personengebundenen & zeitgebundenen Störvariablen

### Interpretationsmöglichkeiten:

- Aussagen über Zusammenhang Intervention & AV
- Aussagen über Kausalität nicht (oder nur unter speziellen Annahmen) möglich

## Quasi-experimentelle Versuchspläne

### Charakteristika:

- Trennung UV/AV
- systematische Beobachtung der AV
- gezielte Manipulation der UV
- keine randomisierte Zuweisung der Versuchspersonen zu den Bedingungen, stattdessen "natürlich vorgefundene Gruppen"

### Interpretationsmöglichkeiten:

- Aussagen über Zusammenhang UV und AV
- Aussagen über Kausalität nur eingeschränkt möglich
- nur verwenden, wenn Experiment nicht durchführbar

## Quasi-experimentelle Versuchspläne

**Problem von Quasi-Experimenten:**

**Experiment:**

- alle möglichen/denkbar den **systematischen personenbezogenen** Störvariablen sind durch Randomisierung kontrolliert
- UV ist Ursache

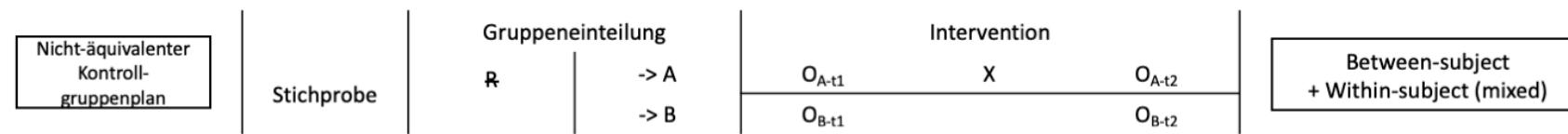
**Quasi-Experiment:**

- alle möglichen/denkbar den **systematischen personenbezogenen** Störvariablen können systematisch mit Bedingung (d.h. mit UV) konfundiert sein
- UV und SV als mögliche Ursache

# Versuchspläne

## Quasi-experimentelle Versuchspläne

### Nicht-äquivalenter Kontrollgruppenplan



## Quasi-experimentelle Versuchspläne

### Nicht-äquivalenter Kontrollgruppenplan

- Vorher-Nachher-Messung in 2 Bedingungen (mit Intervention)
- Trennung & Sequenz UV-AV, Kontrollgruppe, Manipulation
- keine zufällige Zuordnung
  - Vortest notwendig zur Korrektur der Vorher-Unterschiede
  - weit verbreiteter Versuchsplan, wenn Randomisierung nicht möglich

### Interpretationsmöglichkeiten

- Aussagen über Zusammenhang UV & AV
- Aussagen über Kausalität eingeschränkt möglich
- Zentrales Problem bei Veränderungsmessung (Vorher-Nachher-Messung): Unterschiede im Ausgangsniveau

→ Veränderungen im Bezug zum Ausgangsniveau beurteilen, z.B. Größe der Veränderung an Ausgangsniveau relativieren: prozentuale Veränderung statt Differenz

# Versuchspläne

## Experimentelle Versuchspläne

### Charakteristika:

- Trennung & Sequenz UV-AV
- systematische Beobachtung der AV
- gezielte Manipulation der UVs
- randomisierte Zuweisung zu den Bedingungen

### Interpretationsmöglichkeiten:

- Aussagen über Zusammenhang UV & AV
- Aussagen über Kausalität, sofern alle nicht-personenbezogenen, systematischen Störvariablen kontrolliert wurden

### Typische Versuchspläne:

- Labor vs. Feldexperiment
- Randomisierter Kontrollgruppenplan mit/ohne Vortest
- 1 oder 2 Treatments & eine Kontrollgruppe
  - z.B. 1 oder 2 inhaltliche Varianten des neuen Treatments vs. KG
  - z.B. unterschiedlich starke Ausprägungen des Treatments vs. KG
- 2 Treatments, keine Kontrollgruppe
  - z.B. neues Treatment im Vergleich zu Standardtreatment, das gut gegen KG abgesichert ist
  - z.B. Vergleich unterschiedlicher Ausprägungen des Treatments

## Experimentelle Versuchspläne

### Labor- & Feldexperiment

- Bsp: Steigt die Hilfsbereitschaft unter positiver Stimmung?

### Laborexperiment:

- Beobachtung einer AV bei randomisierter Zuweisung zu Bedingungen der UV in „künstlicher“ Umgebung
- Sind Probanden eher bereit, noch einen zweiten FB auszufüllen, wenn man ihnen zuvor fröhliche Musik vorgespielt hat?
- Umgebung kontrollierbar → höhere interne Validität

### Feldexperiment:

- Beobachtung einer AV bei randomisierter Zuweisung zu Bedingungen der UV in der „natürlichen“ Umgebung
- Sind Leute eher bereit, jemanden am Kopierer vorzulassen, wenn sie zuvor dort 2 Euro gefunden haben?
- Umgebung „natürlicher“ → u.U. höhere ökologische Validität

## Experimentelle Versuchspläne

### Arten von Kontrollgruppen

- keine Behandlung
- Placebo-Behandlung
- etablierte Standard-Behandlung (Treatment as usual, TAU)
- Wartelistenplatz: Behandlung erfolgt nach der Studie

→ je nach Fragestellung andere Typen von Kontrollgruppen sinnvoll

# Versuchspläne

## Experimentelle Versuchspläne

### Kontrollgruppenplan ohne Vortest



## Experimentelle Versuchspläne

### Kontrollgruppenplan ohne Vortest

- Beobachtung der AV nach Randomisierung & Manipulation der UV
- Trennung und Sequenz UV-AV
- gezielte Manipulation der UV
- randomisierte Zuweisung zu den Bedingungen
- z.B. Wirksamkeit des Marathon-Trainings nach zufälliger Zuweisung

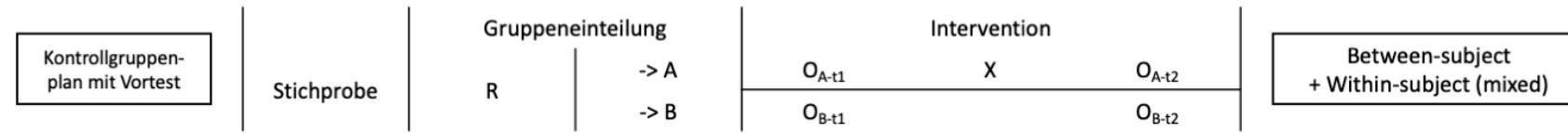
### Interpretationsmöglichkeiten

- Aussagen über Zusammenhang UV und AV
- Aussagen über Kausalität, sofern alle nicht-personenbezogenen, systematischen Störvariablen kontrolliert wurden

# Versuchspläne

## Experimentelle Versuchspläne

### Kontrollgruppenplan mit Vortest



## Experimentelle Versuchspläne

### Kontrollgruppenplan mit Vortest

- Beobachtung der AV vor und nach Manipulation der UV
- vorher bestehende Unterschiede zwischen den Gruppen prüfbar
- reduziert Fehlervarianz (erhöht statistische Validität)

### Problem:

- Sensitivierung und Übungseffekte durch Vortest
- potentielle Gefährdung interne bzw. Konstruktvalidität, abhängig von Messmethode → Kontrolle des Vortest-Effekts wünschenswert

### Interpretationsmöglichkeiten:

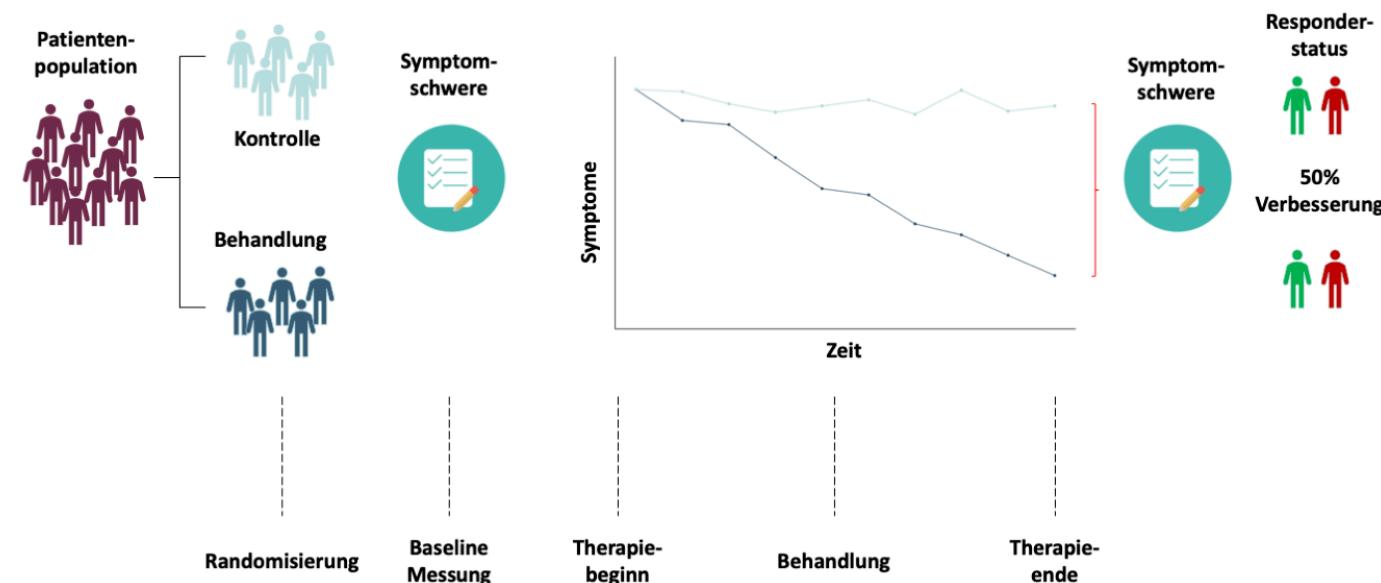
- Aussagen über Zusammenhang UV und AV
- Aussagen über Kausalität, sofern alle nicht-personenbezogenen, systematischen Störvariablen kontrolliert wurden

# Versuchspläne

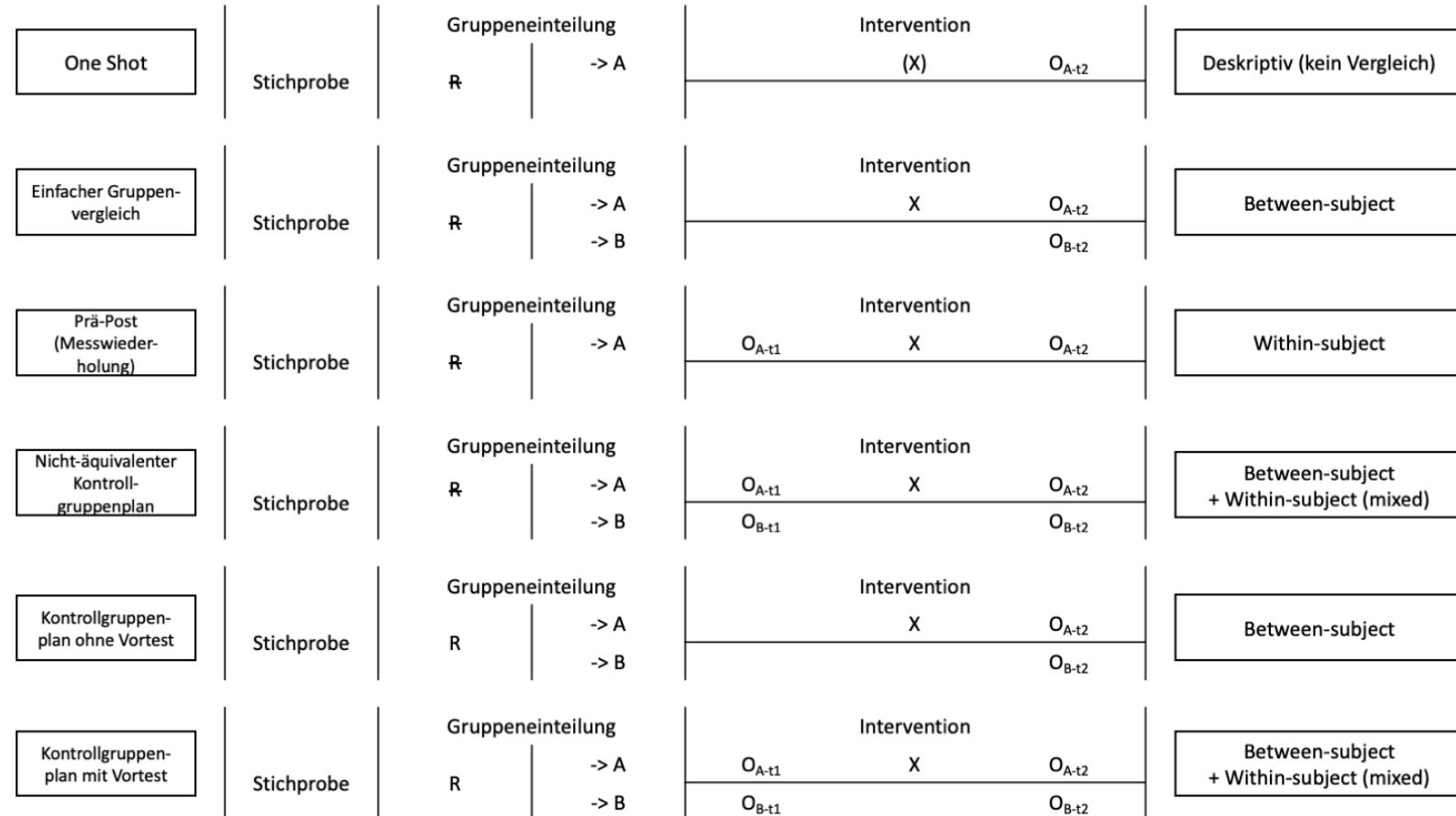
## Experimentelle Versuchspläne

### Das randomisiert-kontrollierte Trial (randomized-controlled trial, RCT)

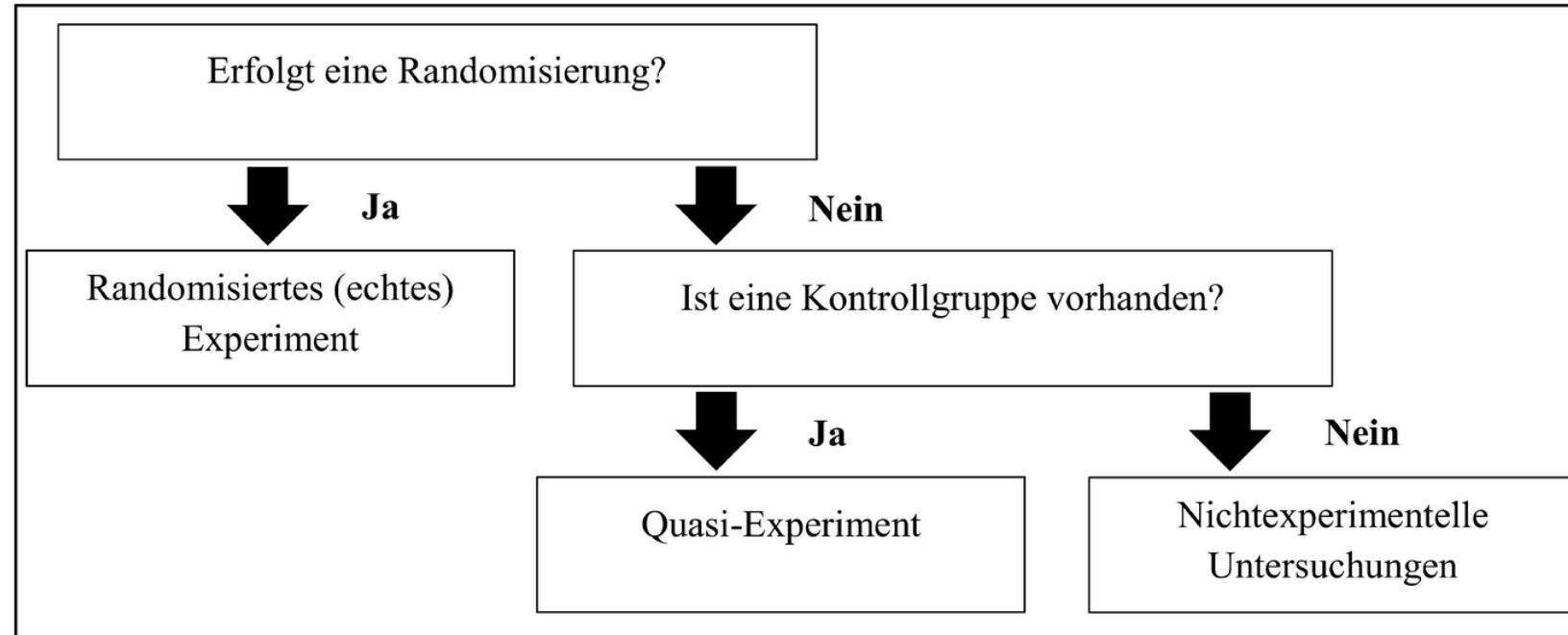
Standardparadigma der klinischen Wirksamkeitsforschung: Kontrollgruppenplan (meist mit Vortest)



# Versuchspläne



Nicht abgebildet: Korrelationsstudien



## Ausblick: Mehrfaktorielle Versuchspläne

Faktorielles Design: Alle möglichen Kombinationen der Faktoren (=Gruppenausprägungen) sind realisiert

Beispiel: Pharmakologische Behandlung einer Depression

- UV A: Medikament (A1: Verum vs. A2: Placebo)
- UV B: Dosis (B1: niedrig vs. B2: hoch)

2x2-Design: 4 Versuchsbedingungen als Kombination von A und B:

- A1-B1: Verum, niedrig
- A1-B2: Verum, hoch
- A2-B1: Placebo, niedrig
- A2-B2: Placebo, hoch

Vorteile:

- Effizienter als eine Durchführung jeweils einzelner Experimente für jeden der Faktoren
- Untersuchung von Interaktionseffekten möglich

<!--

--> <!-- -->

# Zusammenfassung: Forschungsstrategien

## Validität im Kontext der Versuchsplanung & Güte der Hypothesenprüfung

- **Interne Validität**

- eindeutige Schlussfolgerung bezüglich der Wirkbeziehung zwischen AV und UV
- alle relevanten systematischen Störvariablen ausgeschlossen
- (+) Experiment

- **Externe Validität**

- Erkenntnisse können auf Population generalisiert werden

- **Ökologische Validität**

- Erkenntnisse können auf Wirklichkeit generalisiert werden
- Problem zu starker Kontrollbedingungen
- (-) Experiment

- **Statistische Validität**

- Reduktion der Fehlervarianz in der AV
- alle relevanten unsystematischen Störvariablen ausgeschlossen

# Zusammenfassung: Forschungsstrategien

- Qualitative vs. quantitative Herangehensweise
- Experimenteller vs. Assoziativer/korrelativer Ansatz
- Laborstudie vs. Feldstudie (Unterscheidung oft nur bei experimentellen Ansätzen)
- Quasi-experimenteller Ansatz
- Längsschnittlicher vs. Querschnittlicher Ansatz

# Zusammenfassung: Forschungsstrategien

## Qualitative vs. quantitative Herangehensweise

### Quantitativ:

- **Vorrangiges Ziel:** soziale Phänomene messbar machen und statistisch auswerten
- **Voraussetzungen:** Vorliegen von Hypothesen und Theorien, die statistisch/mathematisch überprüft werden können
- **Methode:** eher standardisiert
- **Typische Fragestellungen:**
  - A mehr als B?
  - Wenn viel X, dann viel/wenig Y?
- **Typische Verfahren der Datenerhebung:**
  - standardisierter Fragebogen
  - Experiment

### Qualitativ:

- **Vorrangiges Ziel:** soziale Phänomene rekonstruieren; Hypothesen und Theorien generieren
- **Voraussetzungen:** offener, explorativer Zugriff auf das soziale Phänomen
- **Methode:** keine Standardisierung
- **Typische Fragestellungen:**
  - Welche Einstellungen liegen bei ... vor?
  - Gibt es Gemeinsamkeiten zwischen..
  - Was heißt es, ... zu sein?
- **Typische Verfahren der Datenerhebung:**
  - narratives Interview
  - Gruppendiskussion
  - Beobachtung

# Zusammenfassung: Forschungsstrategien

## Experimenteller vs. Assoziativer Ansatz

### Experimenteller Ansatz:

- „Königsweg“, um kausale Zusammenhänge zu prüfen
- Vergleich von Bedingungen (z.B. Experimentalgruppe vs. Kontrollgruppe)
- Kontrollierter Aufbau:
  - Kovariation von AV und UV systematisch herbeiführen
  - Sicherstellung zeitliche Vorgeordnetheit der UV
- Randomisierung: Störeinflüsse durch zufälliges Zuteilen zu Gruppen ausschließen

### Assoziativer Ansatz

- pragmatischer/retrospektiver Weg zu Prüfung von Zusammenhängen
- Prüfen gemeinsamen variieren (Korrelation)
- Störeinflüsse: Hoffen auf Wegmittelung durch Gruppengröße
- Möglichkeit: Matching Verfahren (z.B. Propensity Score Matching)

# Zusammenfassung: Forschungsstrategien

## Experimenteller vs. Assoziativer Ansatz

Beispiel: Koffein und Intelligenz

### Experimenteller Ansatz:

1. Intelligenzmessung (AV = Leistung)
2. Zufällige Gruppenzuweisung
3. Intervention (UV = Gruppe)
  - Experimentalgruppe: Kaffee
  - Kontrollgruppe: Placebo koffeinfreier Kaffee
4. Intelligenzmessung

### Assoziativer Ansatz

- Erfragen des täglichen Koffeeinkonsums (UV)
- Intelligenztest durchführen (AV = Leistung)
- Prüfen gemeinsamen variieren (Korrelation)
- Störeinflüsse: Hoffen auf Wegmittelung durch Gruppengröße
- Gefahr von Scheinzusammenhängen

# Zusammenfassung: Forschungsstrategien

## Labor vs. Feld (Untersuchungs-Setting)

### Laborstudie:

- **Künstliche Situation**
- Bedingungen/Störvariablen gut kontrollierbar
- Bedingungen können gut variiert, repliziert und manipuliert werden.
- gezielten Variation einer einzigen Variable (UV)
- Unterschiede in AV auf diese Veränderungen zurückführbar

**Beispiel:** Lernprozesse am Computer mit Eye-Tracker untersuchen

### Feldstudie:

- Datenerhebung im **natürlichen Kontext**
- Bedingungen/ Störvariablen kaum kontrollierbar
- Weniger systematische Variation möglich
- **ökologische Validität:** mehr Nähe zum Verhalten in Alltagssituationen

**Beispiel:** Spielverhalten von Kindern am Spielplatz beobachten

# Zusammenfassung: Forschungsstrategien

## Quasi-Experimenteller Ansatz

- Zentraler Unterschied zum Experiment: keine zufällige Zuteilung zu den Untersuchungsbedingungen
- Auch hier Manipulation der UV und Beobachtung der Folgen für AV
- Zuteilung erfolgt nach anderem Prinzip (z.B. Wunsch der TN, natürliche Gruppen)
- Reduziert die interne Validität einer Untersuchung, ist aber in manchen Forschungskontexten nicht zu ändern
- Wichtige Folge: Maximale Erfassung von möglichen Störvariablen und explizites Berücksichtigen von Alternativerklärungen

# Zusammenfassung: Forschungsstrategien

## Quasi-Experimenteller Ansatz

### Beispiel

Evaluation eines zusätzlichen Therapiemoduls in einer Rehaklinik

- **Hypothese:** Das neue Therapiemodul steigert den Therapieerfolg.
- Man kann nicht zur Therapie gezwungen werden
- **Selbstselektion** → „Verweigerer“ sind Kontrollgruppe
- **Alternativerklärung** für mehr Erfolg in Experimentalgruppe: Therapiemotivation

Wichtig: Ausgangsunterschiede **statistisch kontrollieren**

- **Problem:** kausale Effekte nicht belegbar

# Zusammenfassung: Forschungsstrategien

## Quasi-Experimenteller Ansatz

### Möglichkeiten der Kontrolle von Störvariablen:

- Einbezug eines Vortests – Messwiederholung
- Aggregation über mehrere Beobachtungseinheiten
- Statistische Kontrolle (Auspartialisieren)
- Parallelisierung/ Matching

# Zusammenfassung: Forschungsstrategien

## Längsschnittlicher vs. querschnittlicher Ansatz

### Längsschnittstudien

- aka longitudinal; Messwiederholung
- Personen werden **mehrfach/zu unterschiedlichen Zeitpunkten** gemessen
- **Panelstudien** (dieselbe Stichprobe wird über längere Zeit hinweg beobachtet)
- **Veränderungen** können abgebildet werden
- "**Pre-Post**" **Vergleich** nach Intervention (z.B. Therapie) / Ereignis (z.B. Pandemie) möglich
- Meist sehr aufwändig (Nachverfolgung)
- Bei Leistungstests: Vorsicht **Trainingseffekte!**
- Häufiges Problem: Dropout
- Zeitreihen sind empfehlenswert bei Variablen mit Schwankungen (z.B. Stimmung)

### Querschnittstudien:

- aka cross-sectional
- Eine oder mehrere Stichproben werden **zum gleichen Zeitpunkt** untersucht
- Oft geringere **interne Validität** (z.B. Kohorteneffekte konfundieren u.U. Alterseffekte)
- gut um große Stichproben zu realisieren (z.B. Umfragen)

# Zusammenfassung: Forschungsstrategien

## Beispiel Längsschnittstudie:

### Minnesota Twin Study (Thomas J. Bouchard)

- Untersuchung getrennt aufgewachsener eineiiger Zwillinge (EZ)
- Nature vs. Nurture Debatte
- unterschiedliche Biographien, oft verblüffende Ähnlichkeiten
  - gleichen Wertvorstellungen
  - gleicher Beruf
  - ähnliche Hobbies
- Derzeit in Deutschland: „TwinLife“-Studie (2013 - 2025)



# Take-Aways

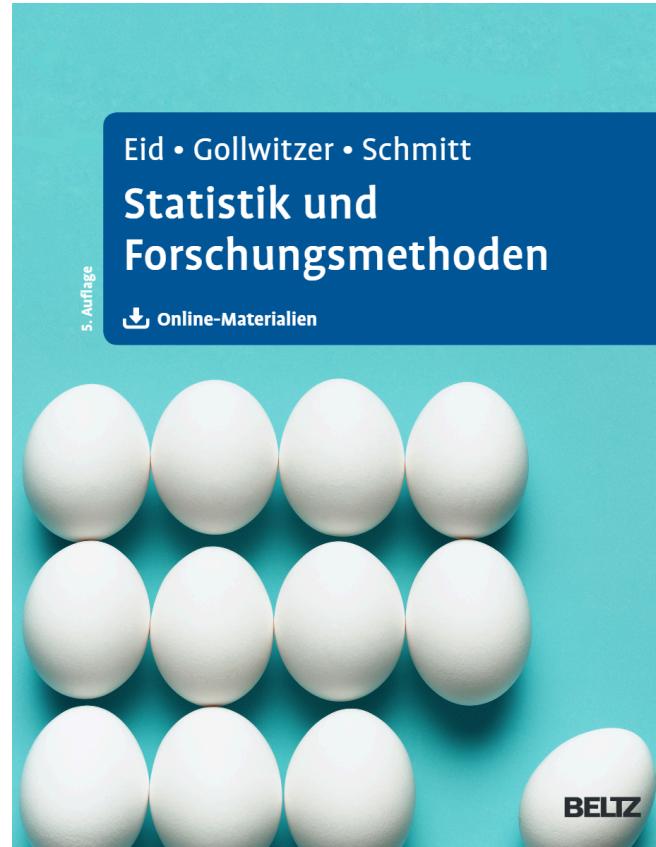
- Experiment gilt als Königsweg zur Kausalität
- Systematische Konfundierungen nach Möglichkeit durch Studiendesign verhindern
- Randomisierung ist ein effizienter Weg, systematische personenbezogene Konfundierungen aus dem Unterschied zwischen experimentellen Bedingungen herauszuhalten
- Durch Verblindung lassen sich Konfundierungen verursacht durch Versuchsteilnehmer und -leiter verringern
- Problem Quasi-Experiment: keine Randomisierung, nur gezielte Manipulation der UV
- Nich-experimentelle Versuchspläne ermöglichen Schlüsse über Zusammenhänge, Kausalitätsschlüsse nicht (oder nur unter speziellen Annahmen) möglich
- Klassischer Experimentaufbau: Kontrollgruppenplan (mit/ohne Pretest) mit randomisierter Gruppenzuweisung

# Schlüssel-/Fachbegriffe der heutigen Vorlesung

Kausalität	Kontrolle von Störvariablen	Parallelisierung / Matching	Prä-Post-Vergleich / vorexperimentelle Anordnung
Scheinkorrelationen	Systematische Störvariablen	Verblindung	Quasi-Experimente
INUS-Bedingungen	Unsystematische Störvariablen	Konstanthaltung	Nicht-äquivalenter Kontrollgruppenplan
Experiment	interne Validität	Elimination	Randomisiert-kontrolliertes Trial (RCT)
Kausalhypothese	statistische Validität	Kontrollfaktoren	Laborstudie
Unabhängige Variable	Zufallsstichprobe	Ausbalancieren	Feldstudie
Abhängige Variable	externe Validität	One-Shot Design	Längsschnittstudie
Randomisierung	ökologische Validität	Korrelationsstudien	Querschnittstudie
Störvariablen		Einfacher Gruppenvergleich	
Konfundierung			

[zurück zur heutigen Übersicht der Vorlesung → zum Quiz zur Wissensprüfung →](#)

# Literatur für die heutige Sitzung



Kapitel 4.3. in Eid, M., Gollwitzer, M., & Schmitt M. (2017). Statistik und Forschungsmethoden. Beltz