



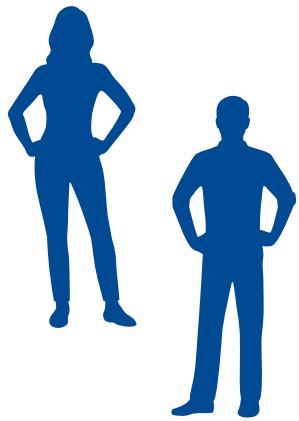
angewandte **w**irtschafts- und **m**edienpsychologie

# Interventionelle Versuchspläne

Thema 04

# Effekte finden

# Intervention



Untersuchungsobjekte

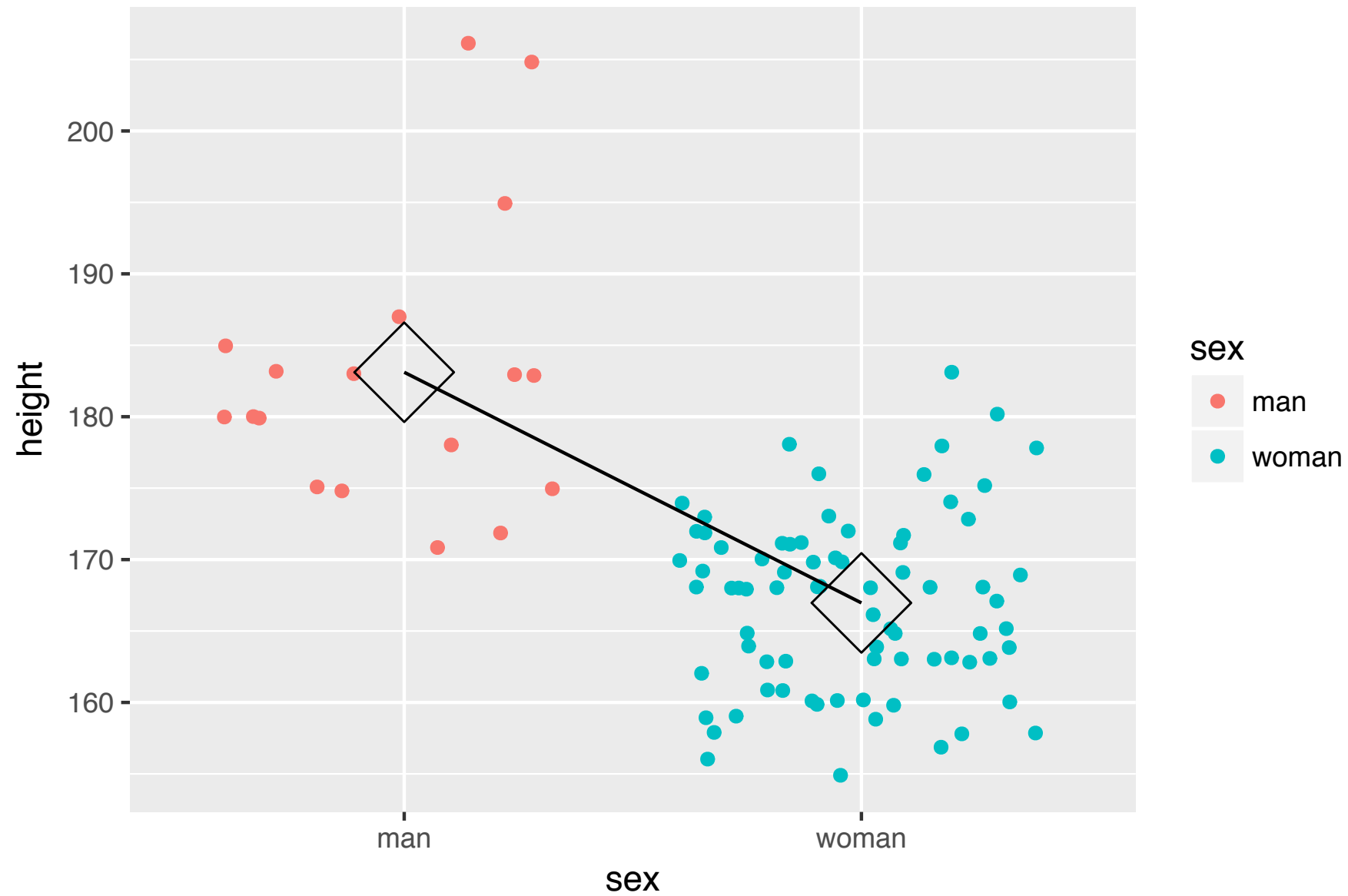


Intervention  
(Treatment)



Messung der Effekte

# Wir wollen Signal, nicht Rauschen



# Wir wollen viel „Signal“, wenig „Rauschen“



Kevin Doncaster, Flickr, 2016, [URL](#)

- ▶ Experimentelle Kontrolle des „Rauschens“
- ▶ Das zentrale Bemühen des Untersuchers gilt der experimentellen Kontrolle der sog. Datenfluktuation (Datenvarianz; Rauschen).
- ▶ Variabilität in Daten ist natürlich und allgegenwärtig.
- ▶ Der Grundgedanke der experimentellen Versuchsplanung besteht darin, unerwünschte Variabilität (Rauschen) von erwünschter (Signal) zu trennen.
- ▶ Rauschen ist Variabilität in den Messergebnissen, die nicht durch die Theorie erklärt werden, sondern z. B. von Messfehlern stammen.
- ▶ Signal ist die Variabilität, die durch die Theorie erklärt wird. So sollte z. B. in einer Gruppe, die das neue Medikament bekommt, eine höhere Heilungsrate bestehen, als in der Kontrollgruppe. Es gibt also Variabilität zwischen den Gruppen, die erwünscht ist.

# Drei Arten epistemologischer Studien

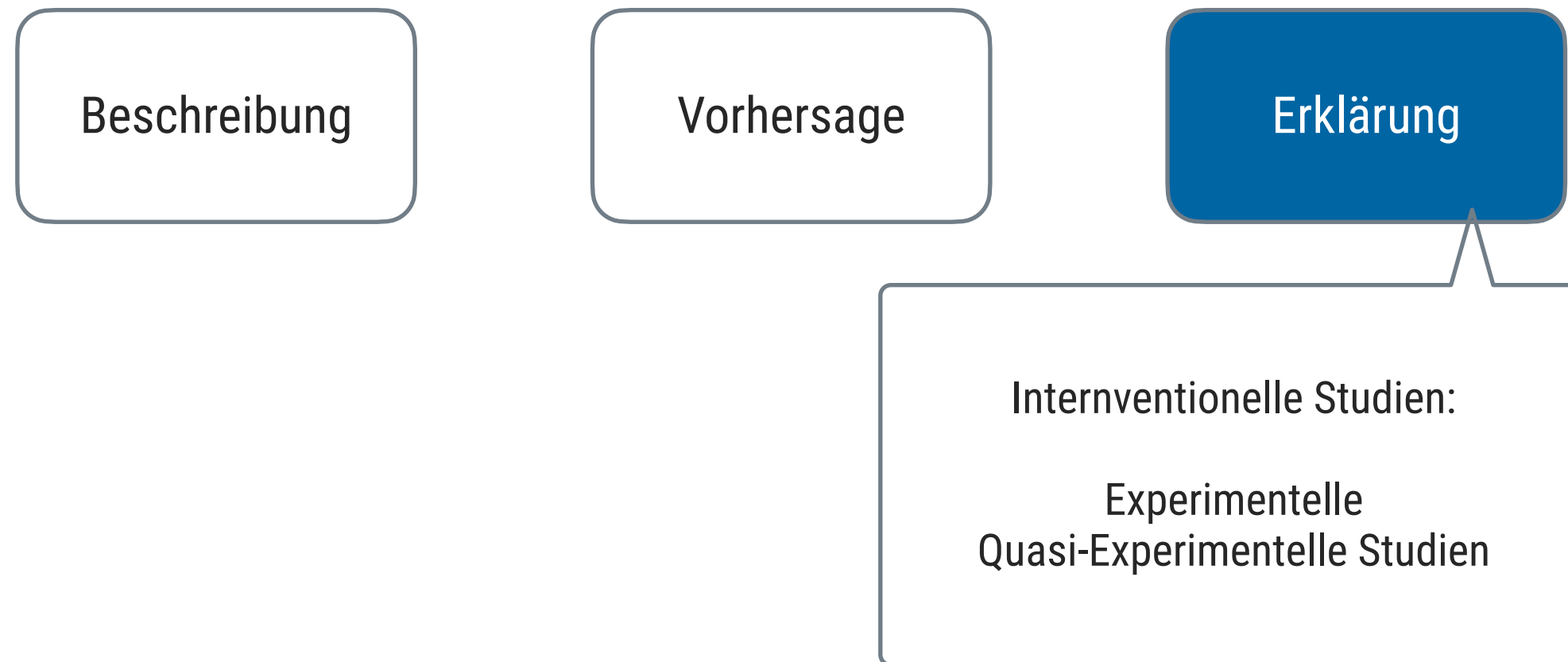
---

Beschreibung

Vorhersage

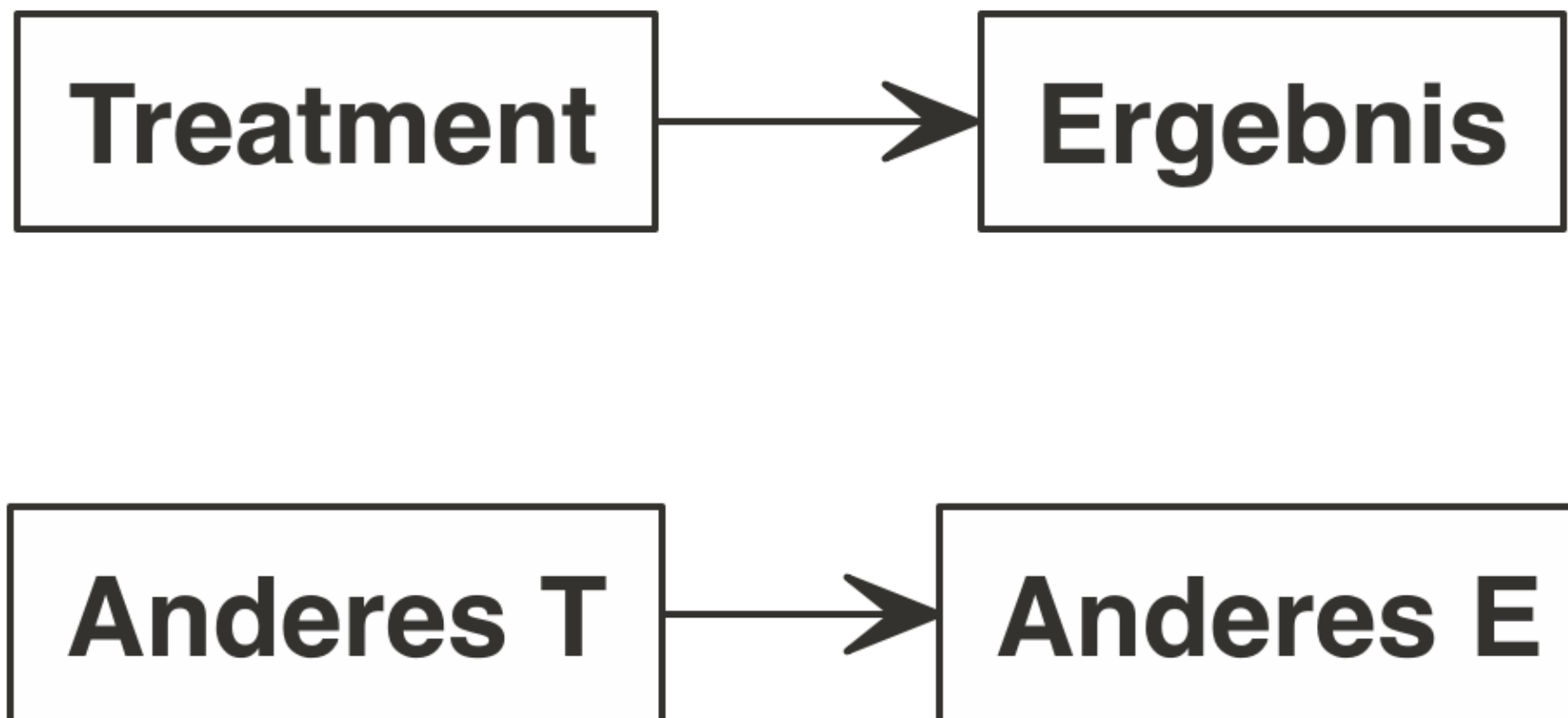
Erklärung

# Drei Arten epistemologischer Studien (2)



# Was ist ein Effekt?

- ▶ Ein Effekt (eines Treatments T/einer Intervention) ist hier definiert als der Unterschied im Ergebnis zwischen was passiert ist nachdem das Treatment verabreicht wurde und was passiert wäre, wenn stattdessen eine andere Intervention verabreicht worden wäre, unter der Annahme, dass sonst alles andere gleich gewesen wäre.





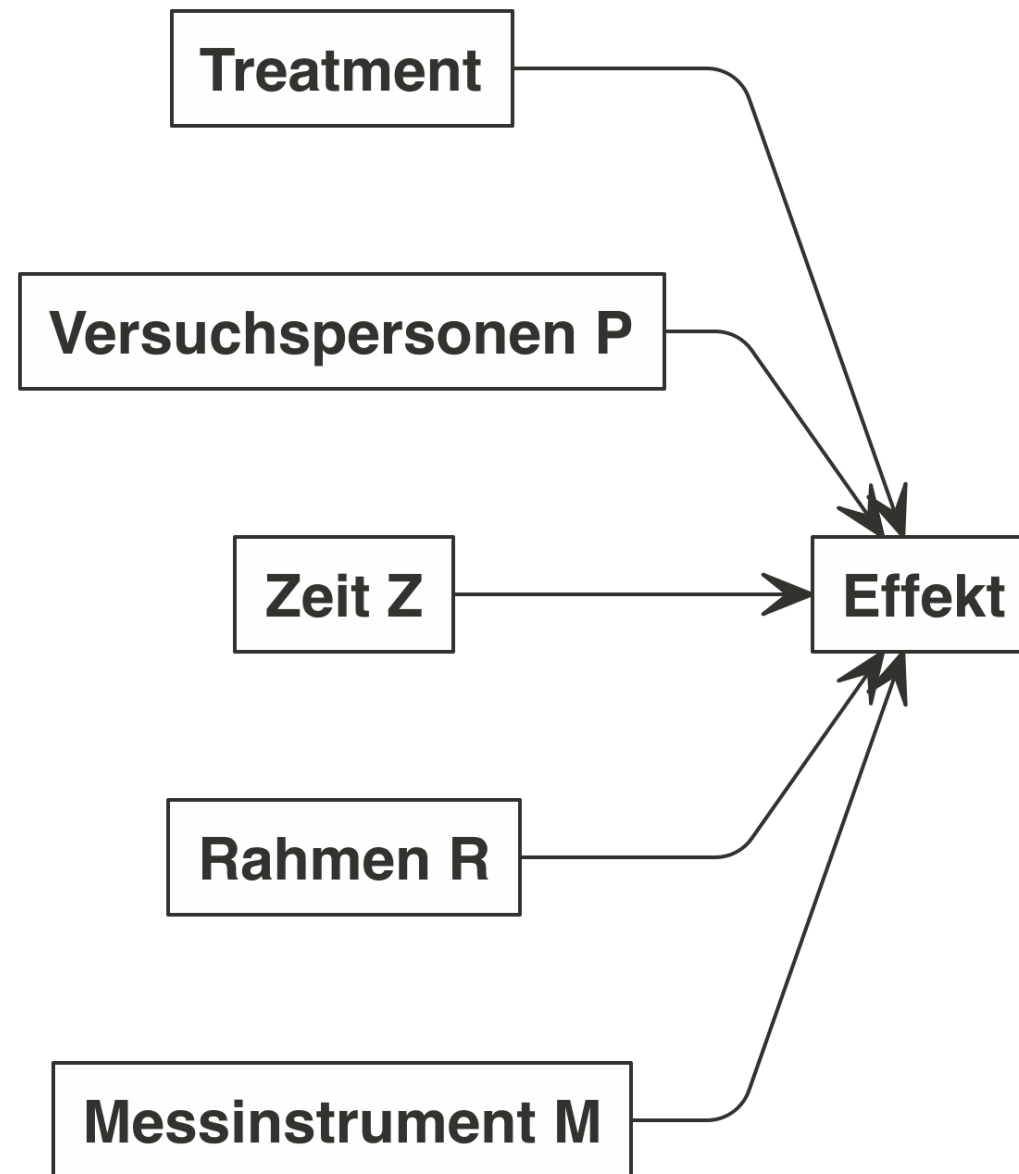
# Leider ist es unmöglich, einen Effekt zu bestimmen

- ▶ Leider ist es per definitionem unmöglich, zu messen, was der Fall wäre, wenn die Welt anders wäre bzw. wir eine andere Intervention angewendet hätten.
- ▶ Man spricht daher von einer *kontrafaktischen* Definition.
- ▶ Auf dieser Basis definiert man den mittleren Treatmenteffekt (average treatment effect, AVE) wie folgt (bei einer Stichprobe oder Population der Größe  $N$  und für die AV  $Y$ ).

$$AVE = \sum_i^N [Y_i(1) - Y_i(0)]$$

- ▶ Nota bene: Man kann bei einer Person ( $i$ ) nur entweder  $Y(0)$  oder  $Y(1)$  beobachten.

# Der Effekt ist eine Funktion von fünf Größen



# Die fünf Determinanten eines Effekts

- ▶ **Ursache (U)** – wie?
  - ▶ Damit ist das Treatment gemeint, soweit es die Ursache für den Effekt ist.
  - ▶ Art und Menge des Treatment spielt natürlich eine Rolle (50 mg Ibuprofen wirken anders als 500 mg).
- ▶ **Versuchsobjekte (O)** – wer?
  - ▶ Häufig Personen in der Psychologie, aber auch Firmen, Teams etc.
  - ▶ Auf dem Versuchsobjekt werden das Treatment angewendet und die Effekte gemessen.
  - ▶ Die Effekte können als zwischen den Versuchsobjekten variieren.
- ▶ **Zeit (Z)** – wann?
  - ▶ Der Effekt hängt vom Zeit des Treatments ab und von der Periode zwischen Applikation und Messung des Effekts.
- ▶ **Rahmen (R)** – wo?
  - ▶ Eine Kopfschmerztablette wirkt vielleicht besser, wenn sie während eines entspannenden Bades an einem ruhigen Abend zuhause eingenommen wird (als während eines stressigen Projektmeetings).
- ▶ **Messinstrument (M)** – was?
  - ▶ Je nach verwendeten Messinstrument oder Zielvariablen können die Effekte des Treatments deutlicher oder weniger deutlich zum Tragen kommen.

# Effektstärke als Funktion der fünf Determinanten

$$ES = f(UOZRM)$$

# Validität

# Konstruktvalidität als Funktion der fünf Determinanten

- ▶ Fehlzuordnungen in den fünf Determinanten begrenzen die Gültigkeit eines Effekts.
- ▶ Die Gültigkeit eines Effekts kann man als Konstruktvalidität bezeichnen.
- ▶ **Ursache (U)** – wie?
  - ▶ Eine neue Coachingmethode (T) hat keinen Effekt, aber die Freundlichkeit des Coaches wurde fälschlich als Effekt des Treatments T interpretiert.
- ▶ **Versuchsobjekte (O)** – wer?
  - ▶ Die Forscher gaben an, die Teilnehmer:innen waren „Health care professionals“, dabei handelte es sich um ungelernte Hilfskräfte.
  - ▶ Die Teilnehmer waren nicht kooperativ und haben sich nicht an die Instruktionen gehalten.
- ▶ **Zeit (Z)** – wann?
  - ▶ Die Forscherin nahm an, das Treatment habe keinen Effekt, aber sie hat nur zu früh gemessen.
- ▶ **Rahmen (R)** – wo?
  - ▶ Vielleicht funktioniert Coaching nur im reichen, demokratischen Westen bei gebildeten Menschen, hängt also vom Rahmen (Ort, Kontext) ab? Die Forscher sind sich über diese Beschränkung aber nicht im klaren.
- ▶ **Messinstrument (M)** – was?
  - ▶ Der Statistiktest des Professors hat nicht das Statistikwissen gemessen, sondern die Spick-Kompetenz (aber die dafür sehr genau). Leider fiel das dem Prof nicht auf (leider für den Prof, die Studis fanden es super).

# Interne Validität als Spezialfall der Konstruktvalidität

- ▶ Die interne Validität kann als Spezialfall der Konstruktvalidität gesehen werden.
- ▶ Die interne Validität fragt, inwieweit der Determinant der Ursache korrekt bekannt ist: Ist die angenommene Ursache tatsächlich auch die Ursache der beobachteten Effekte?
- ▶ Bedrohungen bzw. Einschränkungen der internen Validität beziehen sich nur bestimmte Fehlbeschreibungen beim Determinanten der Ursache (U): *Störvariablen die auch ohne das Treatment T zu einen Effekt führen würden.*
- ▶ **Versuchsobjekte (O)** – wer?
  - ▶ In einer medizinischen Studie werden die gesünderen Patienten in die Gruppe mit dem neuen Medikament gebracht, aber die kränkeren in die Gruppe mit dem bisherigen Standard-Medikament.
- ▶ **Zeit (Z)** – wann?
  - ▶ Zur Messung von Stress wird Speichel-Kortisol entnommen. Die Proben der Experimentalgruppe bleiben aber zulange der Raumtemperatur und Sauerstoff ausgesetzt (im Gegensatz zu den Proben der Kontrollgruppe).
- ▶ **Rahmen (R)** – wo?
  - ▶ Die Verabreichung des neuen Medikaments fand in einer Wellness-Klinik statt. Das alte Medikament im Keller des alten Krankenhauses.
- ▶ **Messinstrument (M)** – was?
  - ▶ In der Experimentalgruppe wurde ein erfahrener Beobachter zur Analyse des Assessment-Centers eingesetzt, in der Kontrollgruppe der Praktikant, der von Tuten und Blasen keine Ahnung hatte.

# Externe Validität

- ▶ Die externe Validität fragt, wie gut ein Effekt verallgemeinert werden kann.
- ▶ Ursache (U) – wie?
  - ▶ Eine Forscherin hat nur einen Teil eines neuen Coachingskonzepts in einem Treatment umgesetzt. Sind die Ergebnisse jetzt auf das ganze Coachingkonzept verallgemeinerter?
  - ▶ Gibt die Forscherin korrekt an, dass nur ein Teil im Treatment umgesetzt wird, liegt hingegen keine Schwäche in der Konstruktvalidität vor.
- ▶ Versuchsobjekte (O) – wer?
  - ▶ Ein Krebsmedikament wurde nur an älteren Frauen evaluiert (und dies hat das Forscherteam korrekt angegeben). Ob der Effekt wohl auch für andere Altersgruppen und Geschlechter gilt?
- ▶ Zeit (Z) – wann?
  - ▶ Ein Allergiemittel reduziert kurzfristig die relevanten Symptome. Aber ob es wohl auch langfristig Effekte hat?
- ▶ Rahmen (R) – wo?
  - ▶ Ob psychologische Erkenntnisse auch in nicht-studentischen Populationen gelten?
- ▶ Messinstrument (M) – was?
  - ▶ Finden sich die Effekte des Vorschultrainings wohl nicht nur in einem bestimmten Intelligenztest, sondern auch in anderen? Haben sie vielleicht auch Effekte auf die emotionale (nicht nur kognitive) Entwicklung des Kindes?



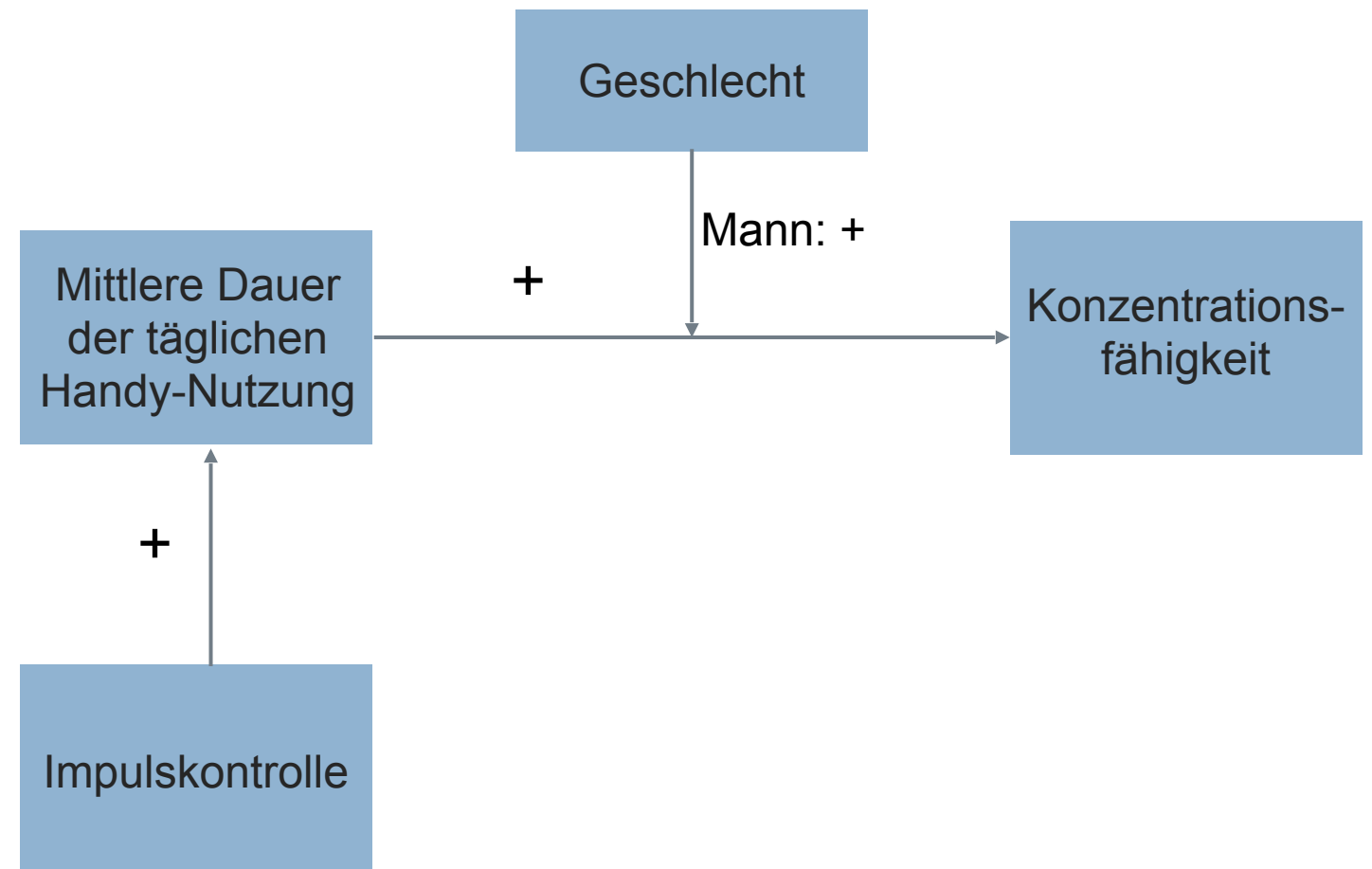
# Übersetzen Sie Ihre Forschungsfrage in ein Modell

- ▶ Suchen Sie in der Literatur nach theoretischen Modellen bzw. Theorien, die die von Ihnen untersuchten Konstrukte *gemeinsam* betrachten.
- ▶ Man kann sich auch selber Modelle ausdenken; diese sollten aber möglichst stark in existierenden Theorien abgebildet sein und durch empirische Belege gestützt sein.

## Laienformulierung

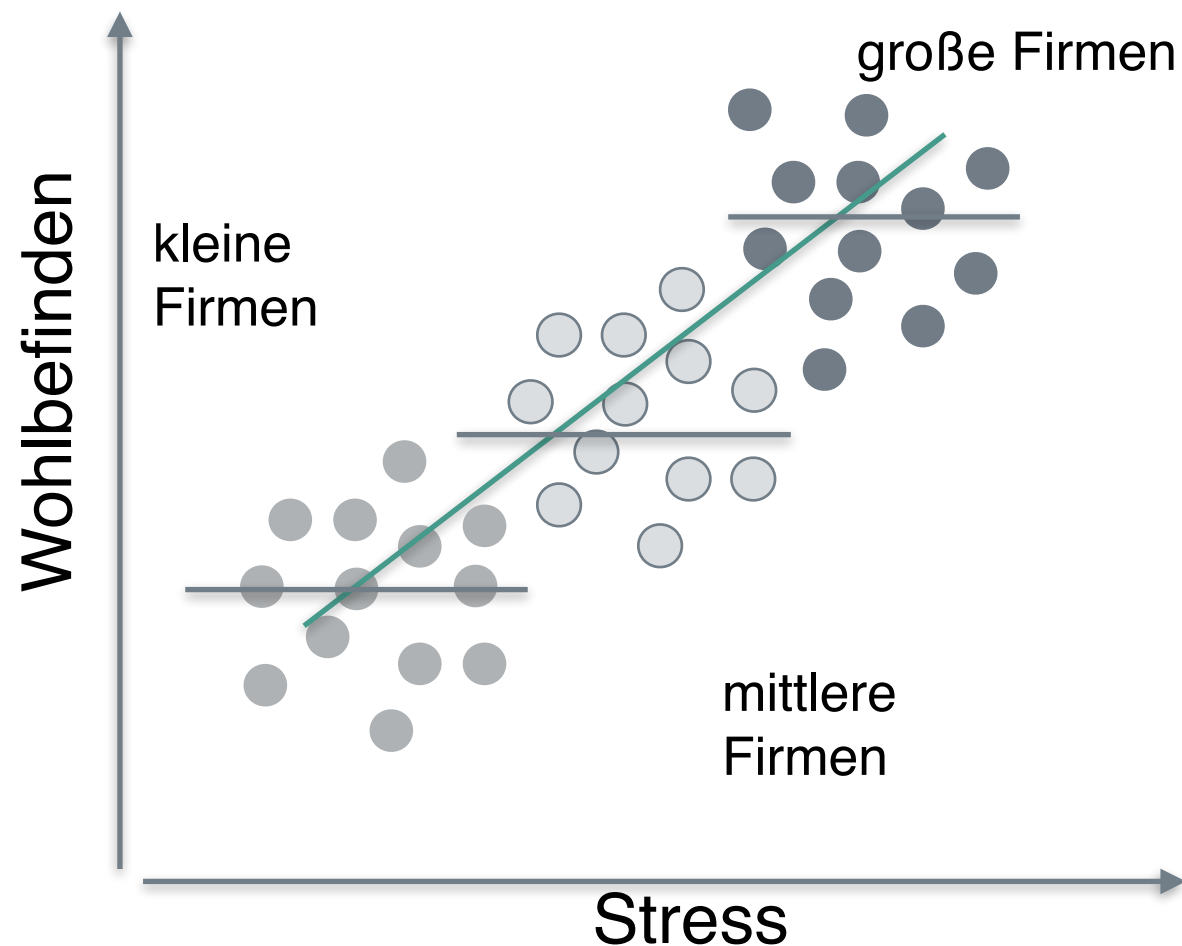
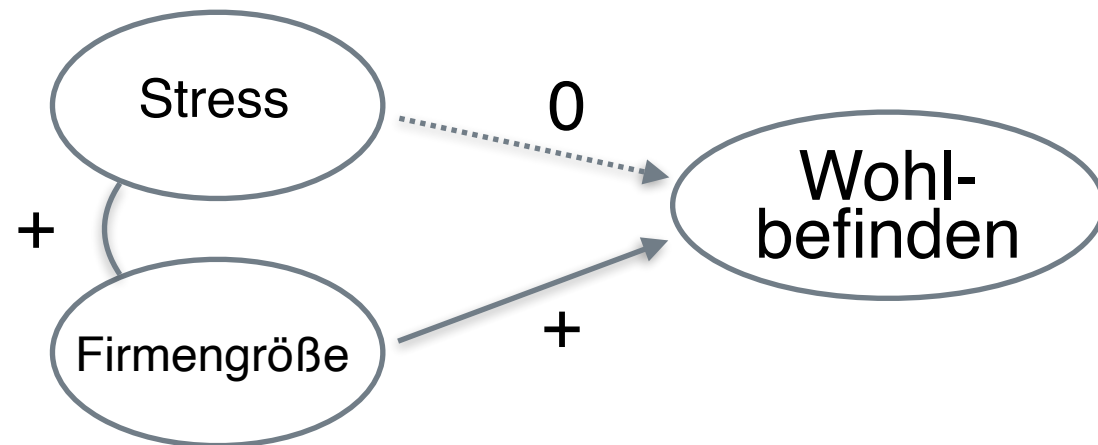
„Wer viel am Handy rumdaddelt, der ist halt nicht so auf Zack, im Hirn, und so, vor allem bei Jungs, übrigens, liegt an den Genen. Ach ja, wer sich halt nicht im Griff hat, der daddelt halt mehr.“

## Wissenschaftliches Modell

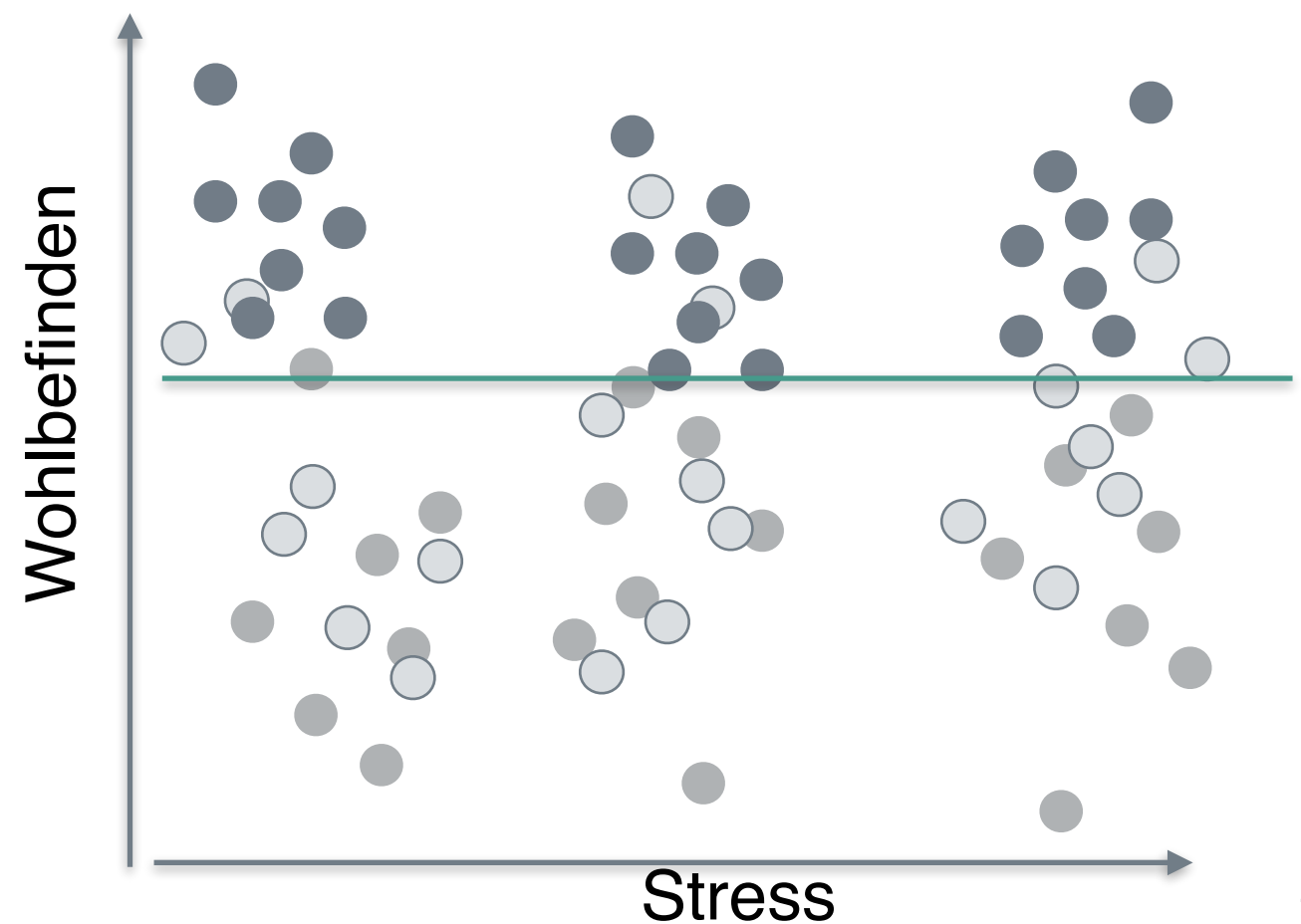
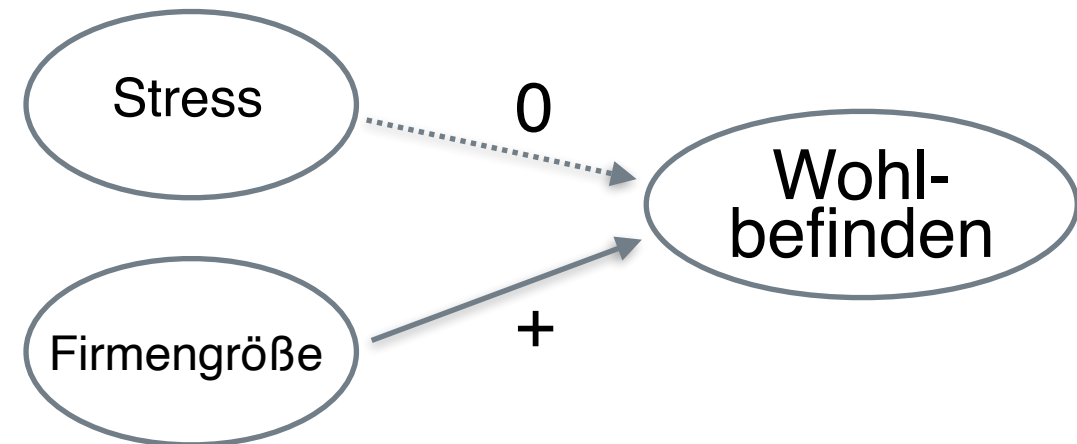


# Aufdecken verborgener Zusammenhänge

## Multivariate Beobachtungsstudie



## Randomisiertes Experiment



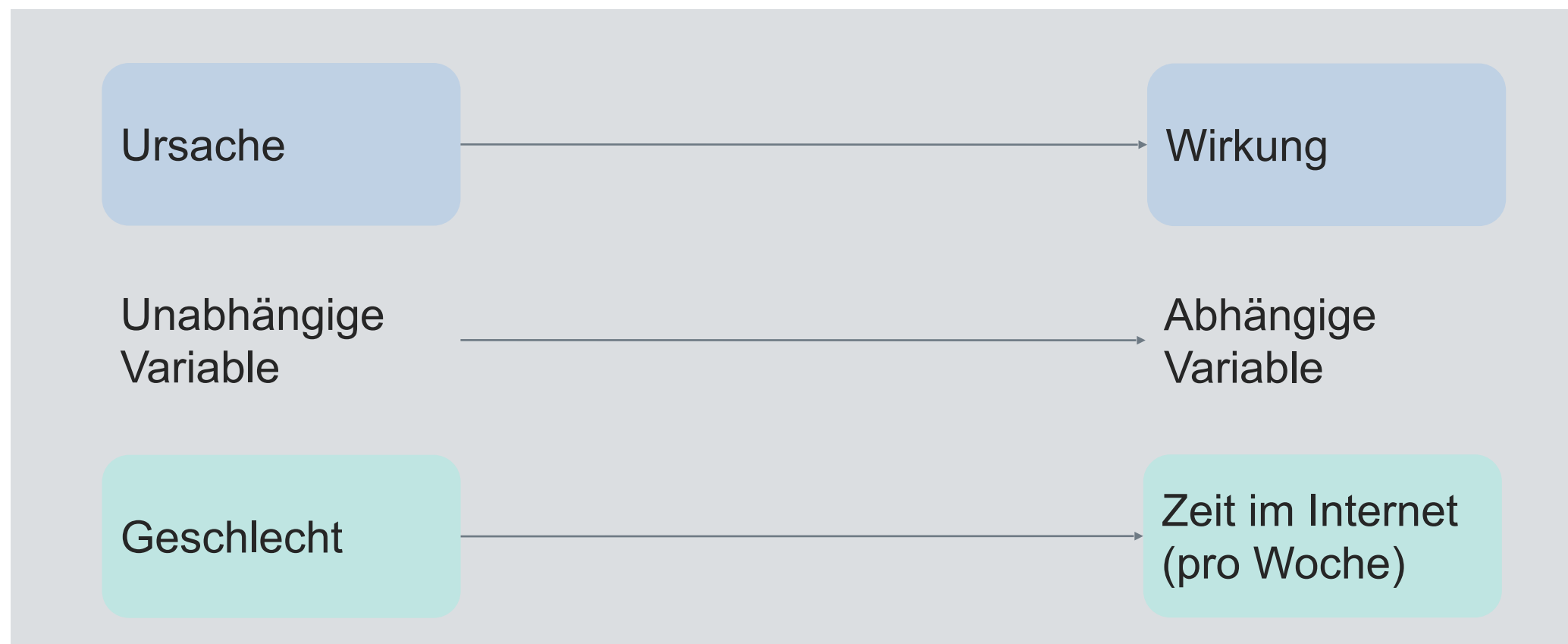
# Variablen in experimentellen Studien (Kausalhypothesen)

## Unabhängige Variable (UV)

- unabhängige Variable; „Ursache“
- Ihr Wert hängt von **keiner anderen Variable** (in der Studie) ab

## Abhängige Variable (AV)

- abhängige Variable; „Wirkung“
- Ihr Wert hängt von der **Variation der unabhängigen Variable** ab



Die Begriffe UV und AV machen nur Sinn, wenn es eine Intervention in der Studie gibt.

# Beeinträchtigung der Validität

# Störfaktoren der internen Validität 1/2

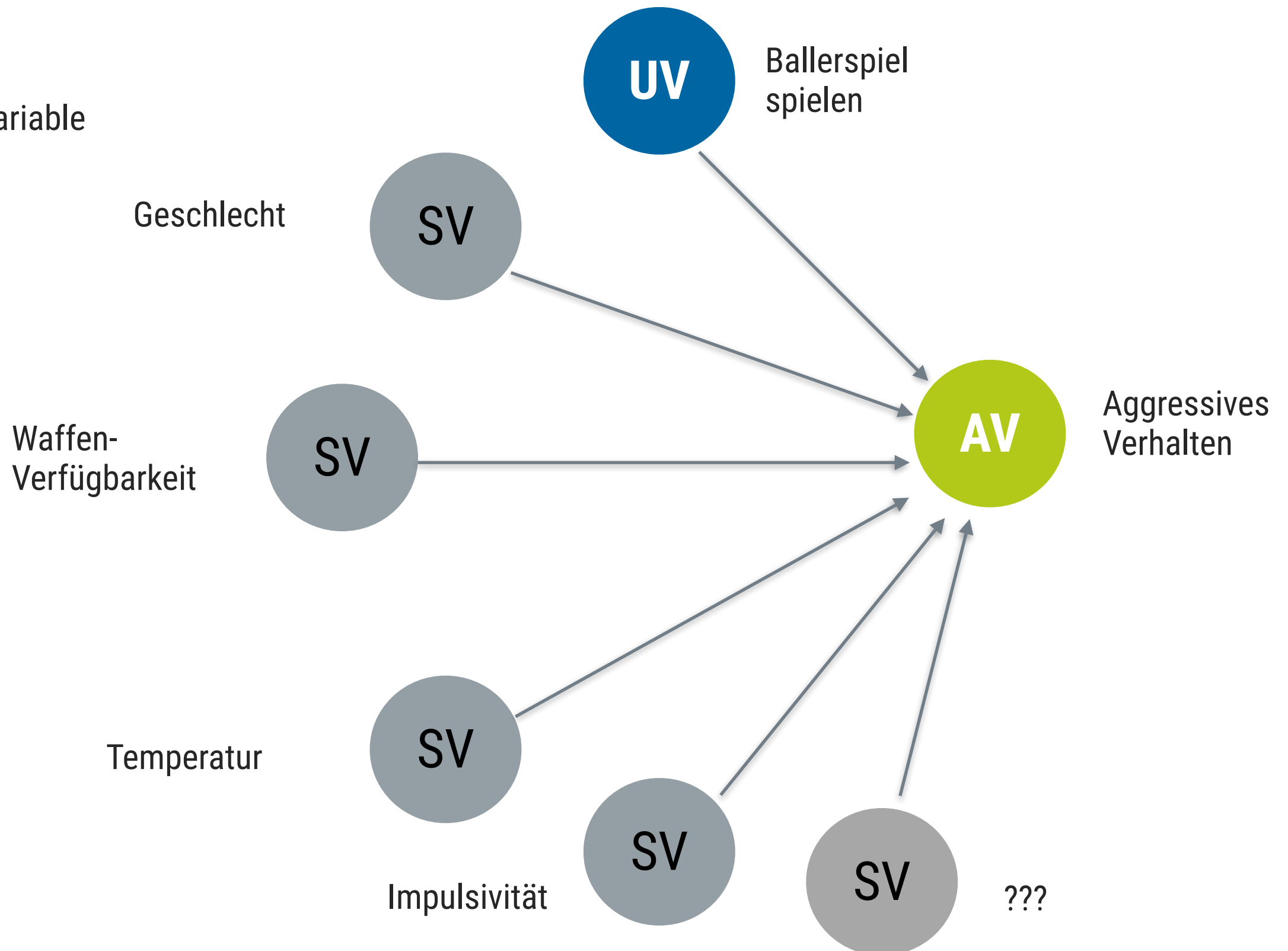
Störfaktor	Beeinträchtigung
Zeitgeschehen (history)	Die beobachteten Effekte gehen nicht allein auf die Untersuchungsbedingungen, sondern auf unkontrollierte zwischenzeitliche Ereignisse zurück.
Reifung (maturation)	Wenn sich der zu untersuchende Sachverhalt bezüglich biologischer (und/oder psychosozialer Reifungsmerkmale verändert, ist mit reifungsbedingten Effekten zu rechnen, die den eigentlichen Befund überlagern.
Mehrfache Testung (test sophistication)	Bei mehrfacher Erhebung derselben Messdaten an ein und demselben Individuum können die während des zweite (dritten ...) Messzeitpunkts erhobenen Daten aufgrund vorangegangener Testung beeinflusst sein.
Instrumentierung (instrumentation)	Die gemessenen Werte gehen z. T. auf die (zwischenzeitlich erfolgte) Veränderung der Messinstrumente zurück (z. B. aufgrund mangelnder Objektivität und Reliabilität eines Tests.)
Statistische Regression (regression)	Werden mehr oder weniger extrem verschiedene Leistungsgruppen z. B. mit Hilfe eines Vortests gebildet, dann kann die mangelnde Reliabilität des Instruments zu einer statistischen Regression zur Mitte bei der zweiten Testung führen.

# Störfaktoren der internen Validität 2/2

Störfaktor	Beeinträchtigung
Auswahlverzerrung (selection)	Bei nicht-zufälliger Bildung von Versuchsgruppen können die damit von Anfang an bestehenden systematischen Ausgangsunterschiede zwischen den Gruppen den eigentlichen Effekt überlagern.
Ausfalleffekte (experimental mortality)	Fallen im Untersuchungsverlauf Versuchspersonen von verschiedenen Versuchsgruppen aus, so kann das die eigentlichen Effekte beeinflussen, wenn die Ausfallquote für die Gruppen systematisch verschieden ist (z. B. mehr „Dropouts“ in Kontrollgruppe).
Versuchsleitereffekte (experimenter-bias effects)	Bleiben die Eigenschaften, Verhaltensweisen und/oder Versuchserwartungen des Untersuchers unkontrolliert, kann das eine systematische Beeinträchtigung der eigentlichen Befunde auslösen.
Interaktive Effekte (carry-over effects)	Wird ein Individuum unter verschiedenen Untersuchungsbedingungen untersucht und bleiben dabei Übertragungseffekte unkontrolliert, können dadurch die Untersuchungsbefunde verfälscht werden.

# Beispiel für eine Kausalhypothese: „Ballerspiele machen aggressiv“

\*SV: Störvariable



# Wann ist eine Ursache wirklich eine Ursache?

Theorie nach John Stuart Mill:

Die Stufen der UV gehen systematisch mit den Stufen der AV einher (Korrelation)

KOVARIATION

**Kausal-  
schluss**  
  
(Ursache  
als Ursache bestätigt)

Veränderungen in der UV  
müssen zeitlich vor den  
Änderungen der AV  
auftreten

Alternativerklärungen für  
den Zusammenhang von UV  
und AV müssen  
ausgeschlossen sein

UV ZEITLICH VOR AV

KEINE ANDEREN URSACHEN  
MÖGLICH



# Störvariablen im Experiment

- ▶ Meist gibt es viele kausale Variablen, die eine AV beeinflussen (können)
- ▶ In einer Studie oder einem Experiment interessiert häufig nur der Einfluss einer Variablen, um eine spezifische Hypothese zu prüfen
- ▶ Alle in dieser Untersuchung nicht interessierenden Variablen, die die AV beeinflussen, heißen Störvariablen
- ▶ Störvariablen mögen selbst theoretisch interessant sein, in einem Experiment versucht man jedoch, sie unter Kontrolle zu bringen, um den Einfluss der interessierenden UV untersuchen zu können
- ▶ Kontrolliert man Störvariablen nicht, kann eine Veränderung in AV nicht eindeutig auf UV zurückgeführt werden, ein Kausalschluss ist dann nicht möglich.
- ▶ Beobachtungsstudien haben meist weniger Kontrolle über Störvariablen als Studien mit Intervention.
- ▶ Studien in kontrolliertem Umfeld sind meist in geringerem Maße Störvariablen ausgesetzt als Studien in „freier Wildbahn“.
- ▶ Studien mit Randomisierung umgehen das Problem der personengebundenen Störvariablen (wenn die Fallzahl groß genug ist).

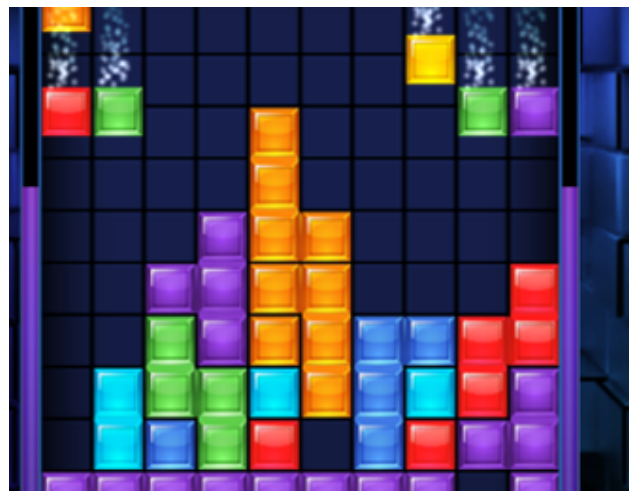
# Systematische Störvariable

UV

Art des Computerspiel



**Gruppe 1:**  
Computerspiel  
mit **viel** Gewalt  
Temperatur:  
heißer Raum



**Gruppe 2:**  
Computerspiel  
mit **wenig** Gewalt  
Temperatur:  
kalter Raum

AV

Aggressives  
Verhalten

- ▶ Hier lässt sich der Einfluss der Raumtemperatur **nicht** vom Einfluss der Computerspiels trennen.
- ▶ Der Einfluss der Raumtemperatur variiert systematisch mit der UV.
- ▶ Man spricht von einer systematischen Störvariablen, wenn die Variable auch einen Einfluss auf die AV hat.
- ▶ Die beiden Variablen (Art des Computerspiels und Raumtemperatur) sind dann konfundiert.
- ▶ Die interne Validität ist hier gefährdet!

# UN-Systematische Störvariable

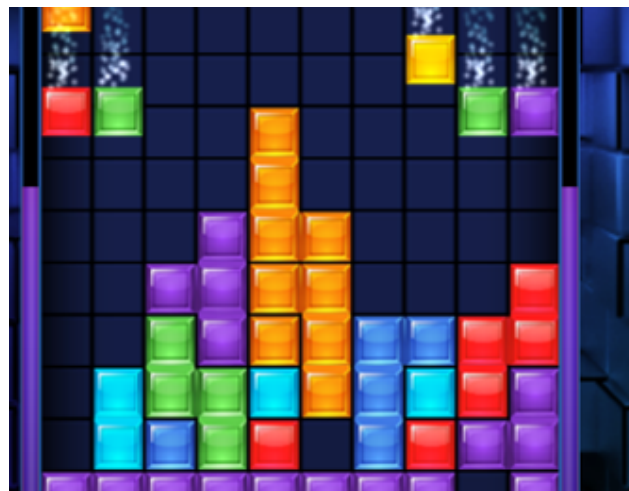
UV

Art des Computerspiel



**Gruppe 1:**  
Computerspiel  
mit **viel** Gewalt

Die Heizung war kaputt!  
Es war mal kalt und mal  
warm, ganz gemischt und  
ohne System!



**Gruppe 2:**  
Computerspiel  
mit **wenig** Gewalt

AV

Aggressives  
Verhalten

- ▶ Schwankt eine Störvariable nicht systematisch mit der UV, so liegt keine (massive) Einschränkung der internen Validität vor.
- ▶ Unter einer nicht-systematischen Störvariablen versteht man eine Variable, die einen Einfluss auf die AV hat (mit ihr kovariiert), aber nicht systematisch mit der UV kovariiert
- ▶ Durch das nicht-systematische Variieren der Raumtemperatur „mittelt“ sich der Einfluss der Raumtemperatur wieder heraus.
- ▶ Es liegt dann keine Konfundierung vor.

# Abschluss

# Hinweise

- ▶ Dieses Dokument steht unter der Lizenz CC-BY 3.0.
- ▶ Autor: Sebastian Sauer
- ▶ Für externe Links kann keine Haftung übernommen werden.
- ▶ Dieses Dokument entstand mit reichlicher Unterstützung vieler Kolleginnen und Kollegen aus der FOM. Vielen Dank!