



**Universität
Zürich^{UZH}**

Psychologisches Institut

Vorlesung Forschungsmethoden

22.11.2018

Urte Scholz



Lernziele der heutigen Veranstaltung

Am Ende der Veranstaltung ...

... können Sie verschiedene Formen der Stichprobenziehung definieren, voneinander abgrenzen und die jeweiligen Vor- und Nachteile benennen.

... können Sie den Unterschied zwischen Labor- und Feldstudien einem Laien erklären sowie auch, welche Vor- und Nachteile jeweils verbunden sind.

... sind Sie in der Lage, interne und externe Validität zu definieren und können erklären, in welchem Verhältnis sie zueinander stehen.

... können Sie Gefährdungen der internen und der externen Validität und entsprechende Umgangs- / Lösungsmöglichkeiten erklären und Beispiele dafür generieren.

... wissen Sie, was unter einem quer- und einem längsschnittlichen deskriptiven Forschungsdesign zu verstehen ist. Sie können einem Laien erklären, welche Fragestellungen Sie mit diesen verschiedenen Designs beantworten können und welche nicht sowie welche Vor- und Nachteile mit den jeweiligen Designs verbunden sind.

Nicht-Zufallsstichproben: Gelegenheitsstichprobe (convenience sampling)

(Döring & Bortz, 2016; Gravetter & Forzano, 2018)

Nicht-zufälliges Auswahlverfahren = **aktuell leicht verfügbare Teilnehmende**

- häufigste Methode der Stichprobenziehung in der psychologischen Forschung

→ **Vorteil:** Einfache Methode zur Stichprobenziehung

→ **Nachteil:** Stichprobe höchstwahrscheinlich verzerrt, nicht repräsentativ für Zielpopulation
(Gravetter & Forzano, 2018, S.122ff.)



<http://asherald.files.wordpress.com/2011/02/calvinsnowball.gif>

Snowball sampling



Nicht-Zufallsstichproben: Quotenstichprobe (quota sampling)

(Döring & Bortz, 2016; Gravetter & Forzano, 2018)

Gelegenheitsstichprobe mit vorgegebenen Quoten hinsichtlich bestimmter Merkmale

→ **Vorteil**: Forschende können Zusammensetzung der Stichprobe kontrollieren

→ **Nachteil**: Stichprobe enthält wahrscheinlich bias

- bei nichtzufällig ausgewählten Stichproben
 - deutliche **Einschränkungen** der Übertragbarkeit / Generalisierbarkeit der Untersuchungsergebnisse auf Zielpopulation



Themenblock II: Quantitative *Erhebungsmethoden*

Ablauf des Forschungsprozess

- ✓ Forschungsidee / Forschungsfrage finden (z.B. Literatursuche, Ethik)
- ✓ Hypothesen formulieren
- ✓ Definition und Messung der Variablen (z.B. Besonderheiten psychologischer Erhebungen; Gütekriterien; Beobachten, Zählen, Messen, Befragen, Testen)
- ✓ Identifizierung und Auswahl der Studienteilnehmenden (Stichprobenziehung)



Themenblock III: Quantitative *Forschungsmethoden*

Ablauf des Forschungsprozess

5. Forschungsdesign wählen

- deskriptives Design
- korrelatives Design
- Experimente
- Quasiexperimente, nicht-experimentelle Forschungsdesigns
- Meta-Analyse



Forschungsdesign wählen (Gravetter & Forzano, 2018)

- Hängt vom Stand der Forschung und von Fragestellung ab
→ Basisziele der Psychologie

Forschungsdesigns - Arten:

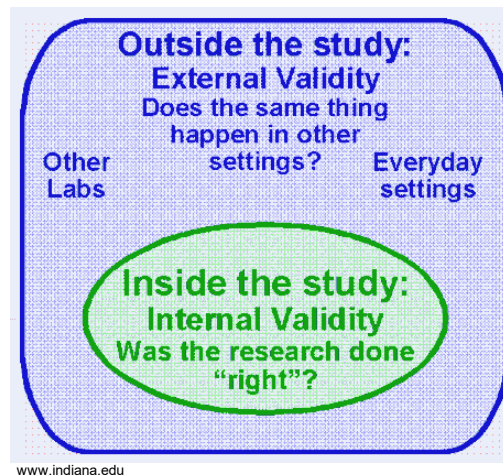
- **Deskriptiv** → reine Beschreibung einzelner Merkmale
- **Korrelativ** → Zusammenhänge zwischen zwei Variablen, keine Erklärung
- **Nicht-experimentell** → Zusammenhänge zwischen zwei Variablen (durch Gruppenunterschiede), keine Erklärung
- **Quasi-experimentell** → Versuch einer Annäherung an Ursache-Wirkungs-Zusammenhänge (Versuch der Erklärung); Problem der natürlichen Gruppen und Konfundierung von Alternativerklärungen mit dem Design
- **Experimentell** → Ursache-Wirkungs-Zusammenhänge (Erklärung) zwischen Variablen

Wichtig: Nur Experimente erlauben Aussagen über Ursache-Wirkungs-Zusammenhänge

Interne und externe Validität

- Validität (Gültigkeit) einer Untersuchung
→ Sind Ergebnisse aussagekräftig, d.h. eindeutig und generalisierbar?

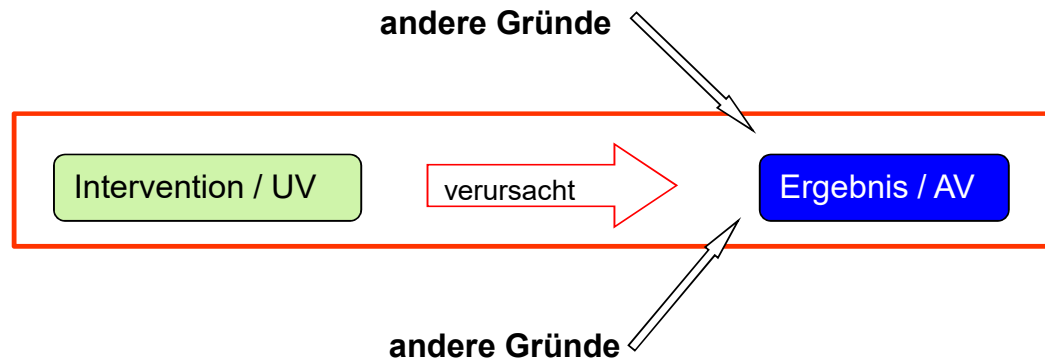
Zur Erinnerung: **interne / externe Validität** \neq **Validität eines Messinstruments / Tests**



Interne Validität

«Eine Untersuchung ist **intern valide**, wenn ihre Ergebnisse **kausal eindeutig interpretierbar** sind in dem Sinne, dass Effekte in den abhängigen Variablen zweifelsfrei auf die Wirkung der unabhängigen Variablen zurückzuführen sind.» (Döring & Bortz, 2016, S. 195)

💣 Sinkt mit Anzahl plausibler Alternativerklärungen (Döring & Bortz, 2016, S.99)



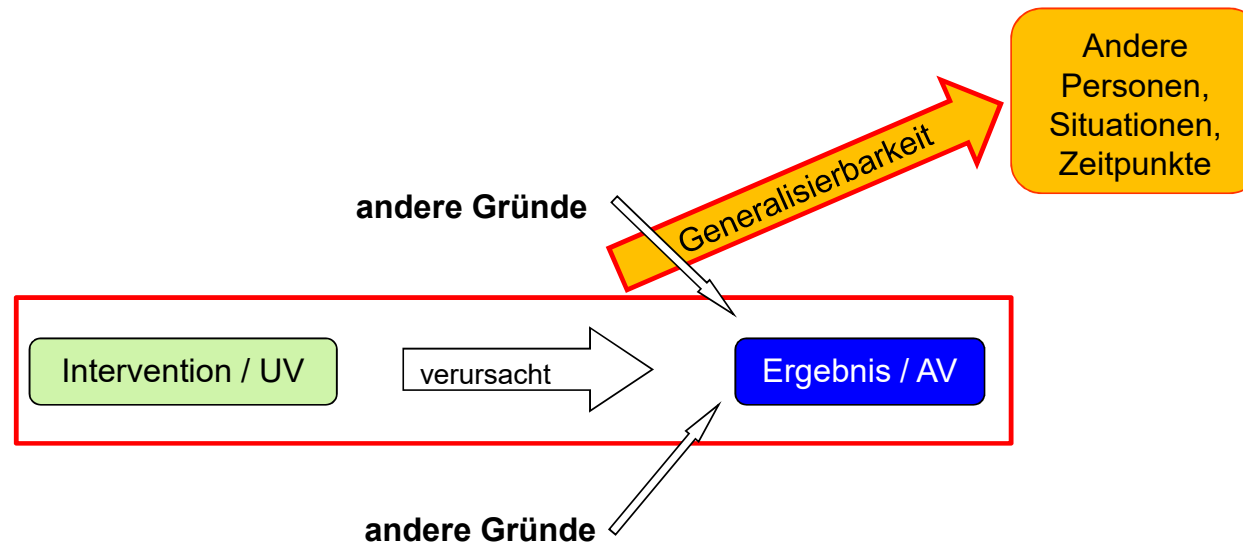
Externe Validität

→ Generalisierung der Ergebnisse zulässig

«Eine Untersuchung ist **extern valide**, wenn ihre **Ergebnisse** über die Bedingungen der Untersuchungssituation und über die untersuchten Personen hinausgehend **generalisierbar** sind.

.» (Döring & Bortz, 2016, S. 195)

☀ Sinkt mit wachsender Unnatürlichkeit der Untersuchungsbedingungen (Döring & Bortz, 2016)





Intern oder extern valide?

Sie möchten prüfen, ob regelmässige sportliche Aktivität positive Auswirkungen auf den Cholesterinspiegel (HDL, LDL) hat.

Sie haben zwei Möglichkeiten, das zu untersuchen:

- A) Im Labor, d.h. die Teilnehmenden leben 3 Monate im Labor, sind täglich 30 Minuten unter Aufsicht sportlich aktiv, erhalten alle die gleiche Ernährung, etc.
- B) Im Feld, d.h. die Teilnehmenden werden instruiert 3 Monate lang täglich 30 Minuten sportlich aktiv zu sein und sich alle nach der gleichen Vorgabe zu ernähren.

Bitte diskutieren Sie die Vor- und Nachteile der beiden Möglichkeiten in Hinblick auf die interne und externe Validität.



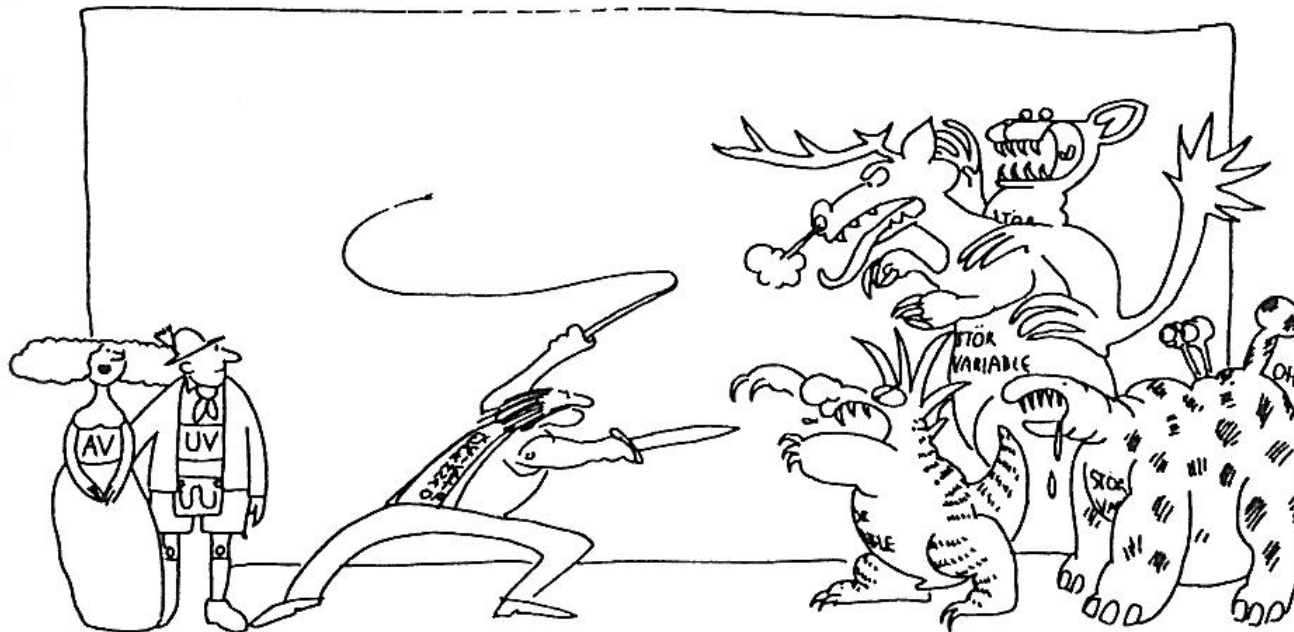
Universität
Zürich^{UZH}

Psychologisches Institut

Feld- und Laboruntersuchungen (Döring, & Bortz, 2016; Gravetter & Forzano, 2018)



Was könnte die interne Validität beeinflussen oder gefährden?



Oswald Huber, 2013



Gefährdungen

Gefährdungen der internen Validität

(Döring & Bortz, 2016)



Umgang / Lösungsmöglichkeiten

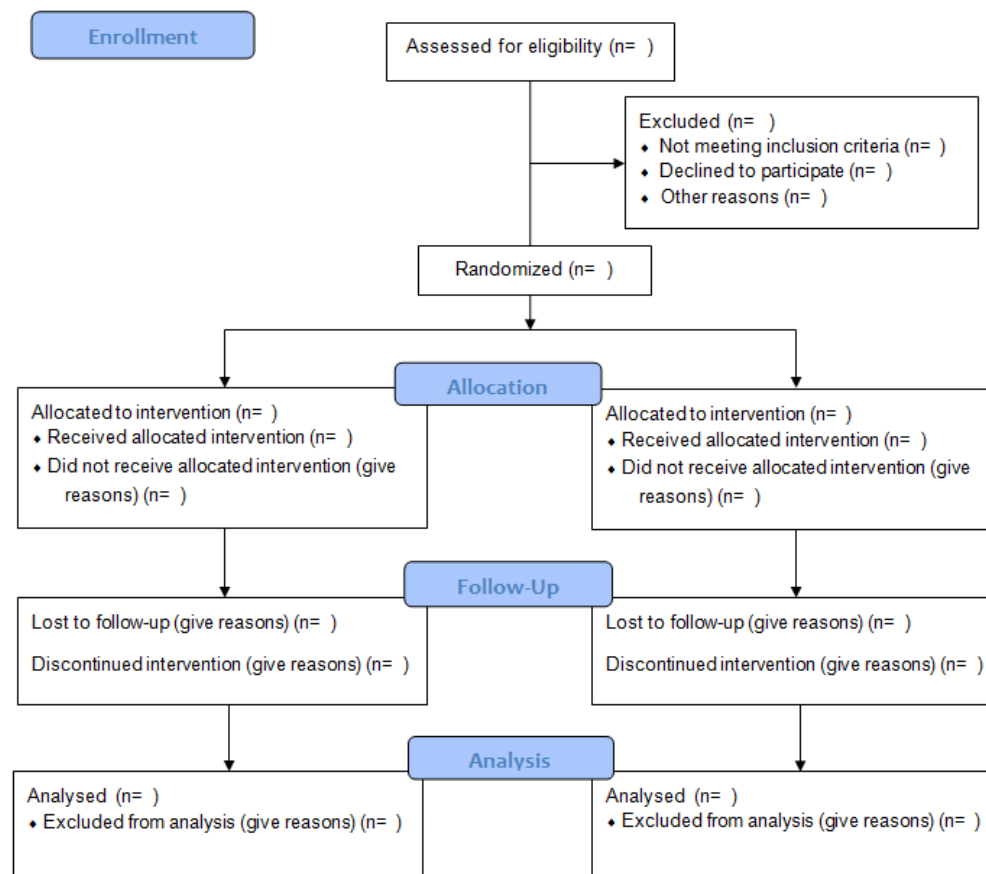
1. Unklare zeitliche Abfolge (ambiguous temporal precedence)
2. Selektionseffekte (selection)
3. Externe zeitliche Einflüsse (history)
4. Reifungsprozesse (maturation)
5. Statistische Regressionseffekte (regression)
6. Experimentelle Mortalität (mortality, subject attrition)
7. Testübung (testing)
8. Mangelnde instrumentelle Reliabilität (instrumentation)
9. Kombination der genannten Bedrohungen (additive and interactive effects of threats to internal validity)

1. Keine Kausalschlüsse
2. Randomisierung, Parallelisierung, Konstanthalten
3. Kontrollgruppe
4. Kontrollgruppe
5. Vermeidung von Extremgruppen (s. nächste Folien)
6. Genaue Dokumentation; statistischer Umgang mit fehlenden Werten
7. Verschiedene Testversionen
8. Verwendung standardisierter Messinstrumente mit hoher Reliabilität
9. Siehe Umgang mit einzelnen Gefährdungen



Experimentelle Mortalität

CONSORT 2010 Flow Diagram





Gefährdungen der internen Validität: **Statistische Regressionseffekte - Regression zur Mitte**

(Döring & Bortz, 2016, S.103)

Extremwerte haben die Tendenz, sich bei wiederholter Messung zur Mitte der Verteilung hin zu verändern:

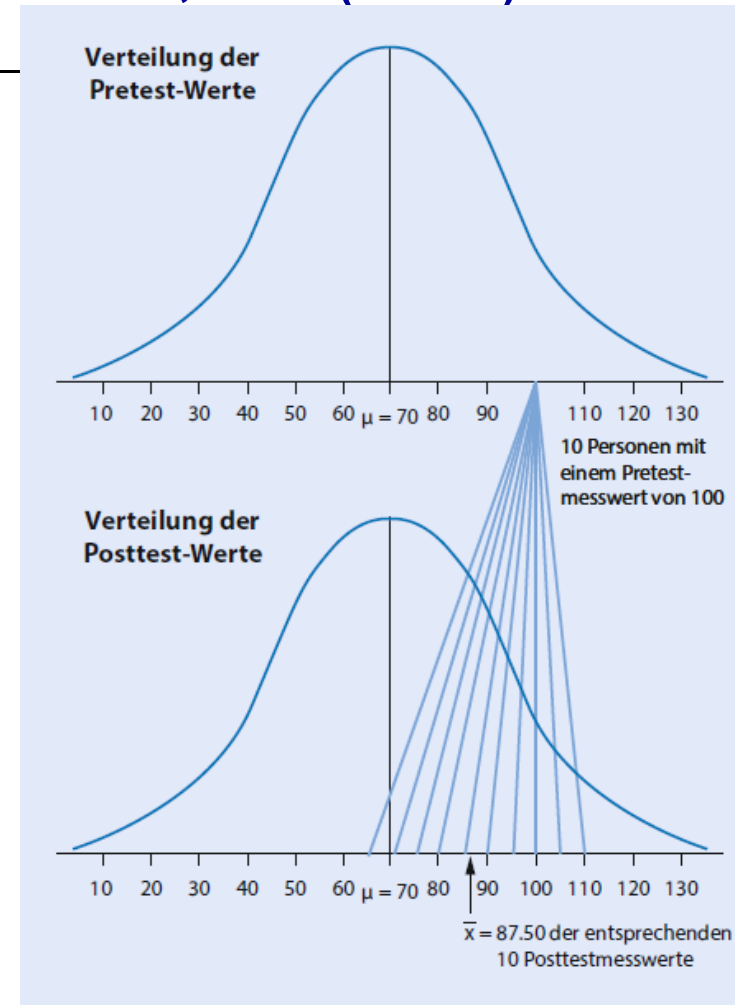
Regression zur Mitte (Galton, 1886)
auch Regressionseffekt / statistische Regression

**“Regression toward the mean is as inevitable as death and taxes”
(Campbell & Kenny, 1999)**

- Immer vorhanden bei verbundenen Messungen (Nachtigall & Suhl, 2002)
- Bei linearer Regression und nicht perfekter Korrelation der beiden verbundenen Messungen Regressionseffekt immer vorhanden
- Bei Veränderungsmessung: je schlechter die Retestreliabilität (Stabilität) eines Tests, desto ausgeprägter die Regression zur Mitte von extremen Werten

Regression zur Mitte Beispiel 1 aus Döring & Bortz, 2016 (S. 736)

- Pretest:
10 Personen mit Messwert = 100; Mittelwert = 100
 - Posttest:
gleiche 10 Personen Werte zwischen 65 und 100;
Mittelwert = 87,5
- weniger weit entfernt vom Mittelwert der Verteilung ($M = 70$) als vorher





Regression zur Mitte Beispiel 2 aus Döring & Bortz, 2016 (S. 737)

Simulationsstudie von Preacher et al. (2005):

Normalverteilte Population ($N = 1000$), pro Person zwei Messungen, (Retestreliabilität=0.8)

1. Messung → Einteilung in unteres, mittleres und oberes Drittel

2. Messung → Einteilung in unteres, mittleres und oberes Drittel

Zahlen in der Tabelle = Fallzahlen (n)

	2. Messung	Unteres Drittel	Mittleres Drittel	Oberes Drittel	Zeilen- summe
1. Messung	Unteres Drittel	241	76	16	333
	Mittleres Drittel	83	183	68	334
	Oberes Drittel	9	75	249	333
	Spalten- summe	333	334	333	1000

Insgesamt nur 74% der Fälle halten Extremgruppenstatus aufrecht

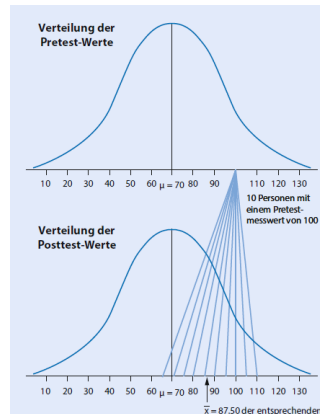
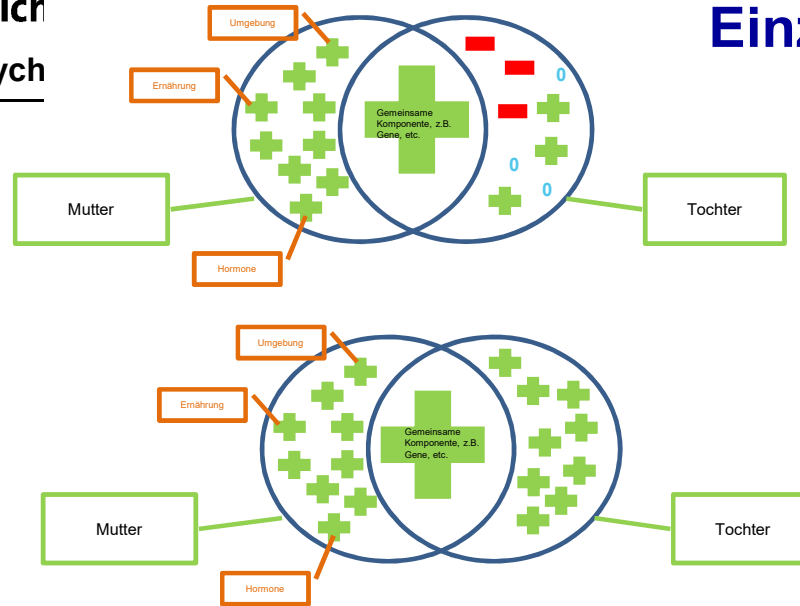


Regression zur Mitte: Erklärung (Nachtigall & Suhl, 2002)

Mittelwert von Merkmal x und Mittelwert von Merkmal y bei verbundenen Messungen:

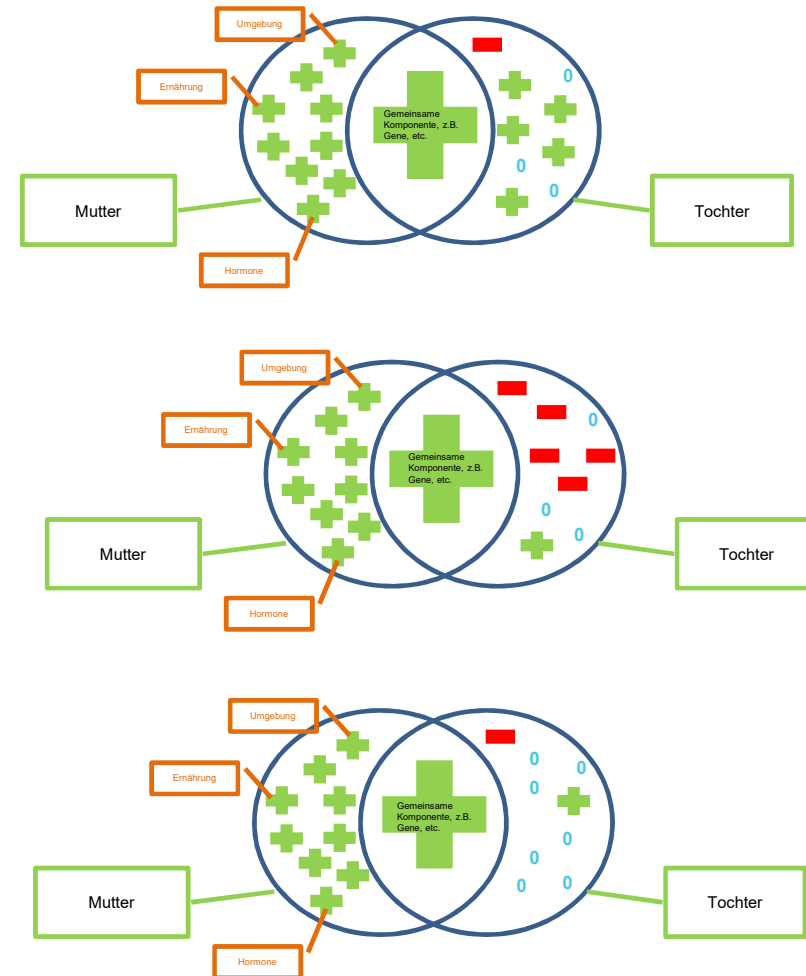
- Gemeinsame Komponente plus weitere Einflussfaktoren
 - Bei extrem hohen Werten im Merkmal x sind z.B. sowohl die gemeinsame Komponente als auch die weiteren Einflussfaktoren hoch ausgeprägt
 - Die gemeinsame Komponente ist entsprechend auch bei y hoch ausgeprägt, aber die weiteren Einflussfaktoren können hoch, durchschnittlich oder niedrig ausgeprägt sein, so dass der Mittelwert von Merkmal y $<$ als der Mittelwert von Merkmal x

Regression zur Mitte bezieht sich auf Mittelwerte, nicht auf Einzelfälle



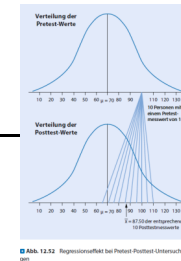
Beispiel 1 aus Döring & Bortz, 2016 (S. 736)

PMH 14.34 Regressionskoeffizienten und Fregressionskoeffizienten



Fazit Regression zur Mitte (Döring & Bortz, 2016)

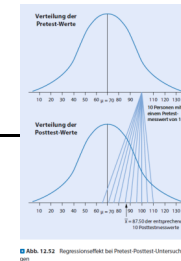
- Regressionseffekt bezieht sich auf Mittelwerte, nicht auf Einzelfälle
- Nicht gerichtet, d.h. auch keine zeitliche Reihenfolge
- Überprüfung einer Veränderungshypothese an einer Extremgruppe
- Veränderung aufgrund des Regression zur Mitte-Effekts zu erwarten
(→ Bedrohung der internen Validität)
- Schlussfolgerung, dass Veränderung ausschliesslich auf die experimentelle Manipulation zurückzuführen ist, wäre nicht haltbar
- bei allen Untersuchungen mit Extremgruppen sollte Regressionseffekt in Betracht gezogen werden
- Nur echte Zufallsstichproben nicht vom Regression zur Mitte-Effekt betroffen



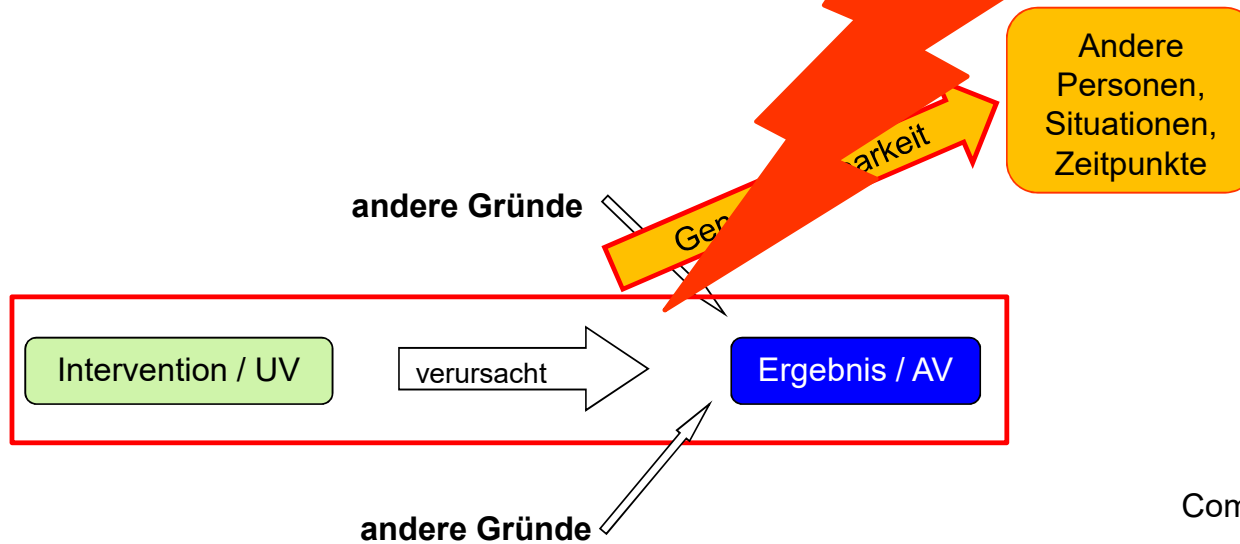
Implikationen Regression zur Mitte (Döring & Bortz, 2016)

Implikationen:

- Vergleich von Veränderungen in Extremgruppen problematisch
- Zufallsstichproben ziehen
- Kontrollgruppen verwenden
- Weniger geeignet: Parallelisierung
- Untersuchungen mit >3 Messzeitpunkten vorteilhafter



Was könnte die externe Validität beeinflussen oder gefährden?



Comic: Oswald Huber



Gefährdungen

1. Merkmale der Teilnehmenden
(Selektionseffekte, Studierende, Freiwilligkeit, weitere Charakteristika der Teilnehmenden)
2. Besonderheiten der Studie (Neuheitseffekt, gegenseitige Beeinflussung durch mehrere Treatments; Charakteristiken der Studienleitung)
3. Besonderheiten der Messung
(Sensibilisierung; spezielle Operationalisierung; Zeitpunkt der Messung)

Gefährdungen der externen Validität

(Gravetter & Forzano, 2018)



Umgang / Lösungsmöglichkeiten

1. Nachweis des Effekts an Stichproben aus unterschiedlichen Populationen
2. Nachweis des Effekts über unterschiedliche Studien / Settings hinweg
3. Berücksichtigung mehrerer Operationalisierungen / Zeitpunkte der abhängigen Variablen

Weißkittelhypertonie


< zurück Seite 4 / 5 weiter >

Wenn der Blutdruck steigt

Gefällt mir Teilen 1 Twittern G+1 Pin.it ★★★★★ 2 Fehler melden



Steigt der Blutdruck nur in der Arztpraxis, sprechen Mediziner von „Weißkittelhypertonie“ Colourbox.de

Erhöhter Blutdruck ist eine typische Angstreaktion: Der Körper wappnet sich für einen Kampf oder Flucht. Wer Angst vor dem Arzt hat, dessen **Blutdruck** steigt im Wartezimmer. Doch manche Menschen fühlen sich in der Sprechstunde nicht anders als sonst. Trotzdem schnell ihr Blutdruck in die Höhe, wenn ein Arzt in der Praxis ihre Werte misst. Die promovierte Psychologin Franziska Einsle von der TU **Dresden**  hat dieses Phänomen an über 400 Patienten untersucht. Sobald die Werte mindestens 20 Millimeter Hg höher

ZUM THEMA



liegen als bei einer Messung zu Hause, sprechen Experten von „Weißkittelhypertonie“. Davon betroffen sind etwa zehn Prozent der Bevölkerung.

http://www.focus.de/gesundheit/arzt-klinik/mein-arzt/tid-12614/iatrophobie-wenn-der-blutdruck-steigt_aid_350044.html

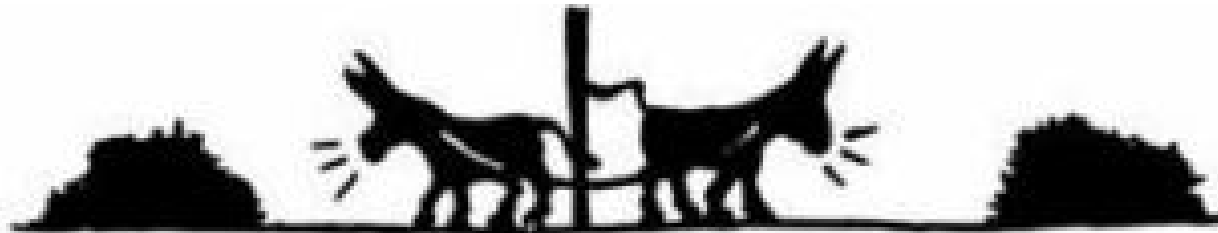


Das Zusammenspiel interner und externer Validität erfordert meistens Kompromisse

- **Interne Validität:** steigt, wenn Alternativerklärungen ausgeschlossen, Störquellen kontrolliert werden
- **Externe Validität:** steigt, wenn Setting natürlich, repräsentative Stichprobe

→ selten beide Gütekriterien in einer Untersuchung voll erfüllt

→ Kompromisslösung





Das Nonplusultra-Design (?): Randomisiertes Kontrollgruppenexperiment im Feld mit grossen Stichproben

■ **Tabelle 7.5** Interne und externe Validität für (quasi-)experimentelle Labor- und Feldstudien

	Externe Validität: gering	Externe Validität: hoch
Interne Validität: hoch	Laborexperiment	Feldexperiment
Interne Validität: gering	Quasi-experimentelle Laborstudie	Quasi-experimentelle Feldstudie

Döring & Bortz,
2016, S. 208



Forschungsdesign wählen (Gravetter & Forzano, 2018)

- Hängt vom Stand der Forschung und von Fragestellung ab
- Basisziele der Psychologie

Forschungsdesigns - Arten:

- **Deskriptiv** → reine Beschreibung einzelner Merkmale
- **Korrelativ** → Zusammenhänge zwischen zwei Variablen, keine Erklärung
- **Nicht-experimentell** → Zusammenhänge zwischen zwei Variablen (i.d.R. Gruppenunterschiede), keine Erklärung
- **Quasi-experimentell** → Versuch einer Annäherung an Ursache-Wirkungs-Zusammenhänge (Versuch der Erklärung); Problem der natürlichen Gruppen und Konfundierung von Alternativerklärungen mit dem Design
- **Experimentell** → Ursache-Wirkungs-Zusammenhänge (Erklärung) zwischen Variablen



Weiterhin Unterscheidung von **Quer- und Längsschnittdesigns**



Deskriptives Forschungsdesign: Querschnittstudie

Prinzip der deskriptiven Querschnittsmethode:

Beschreibung der Ausprägung eines Merkmals anhand einer (oder mehrerer) möglichst repräsentativen Stichprobe(n) **zu einem Messzeitpunkt**

→ Umfrage- / Survey - Forschung

Beispiel:

Repräs.
Stichprobe A

Variable A

Zeitpunkt t1

Anmerkung: Repräs. = Repräsentativ

Deskriptives Forschungsdesign: Querschnittstudie

Beschreibung der Ausprägung eines Merkmals anhand einer (oder mehrerer) möglichst repräsentativen Stichprobe(n) **zu einem Messzeitpunkt**

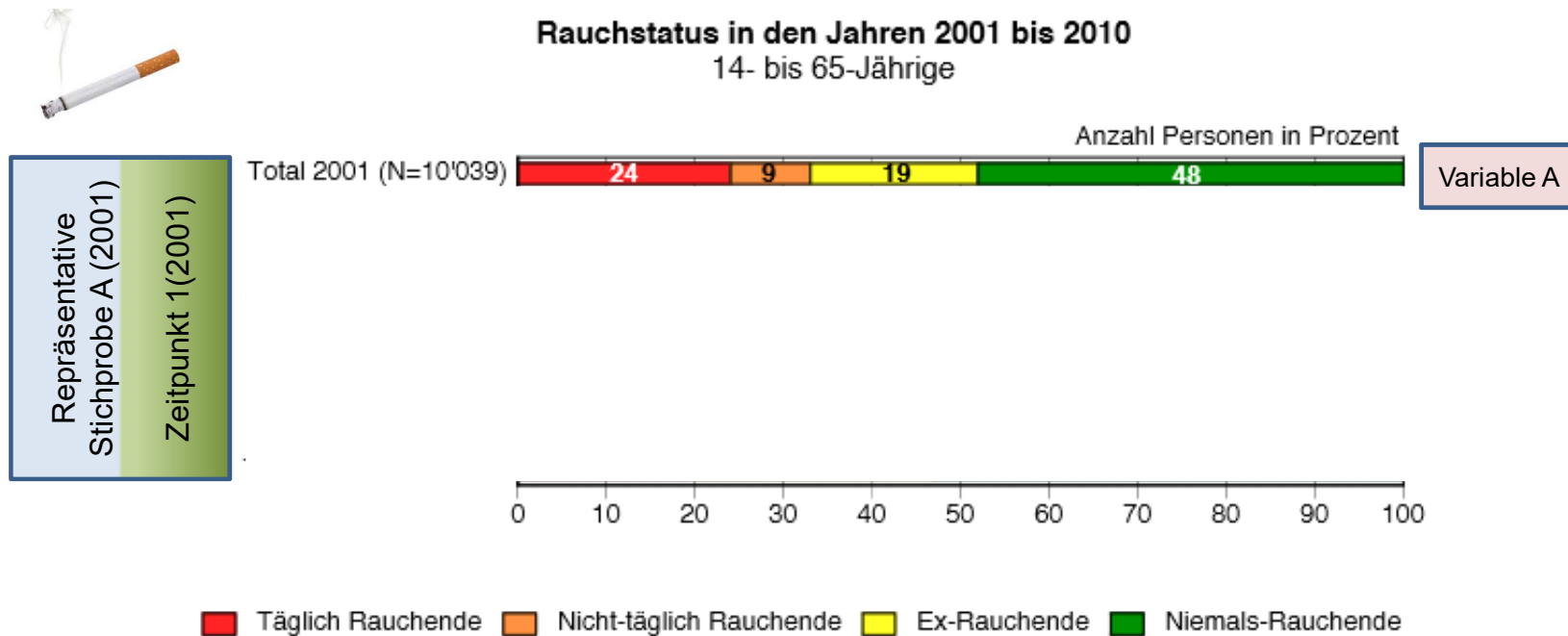


Abbildung 5.2: Rauchstatus der 14- bis 65-jährigen Wohnbevölkerung in den Jahren 2001-2010

Daten aus dem Schweizerischen Tabakmonitoring

http://www.tabakmonitoring.ch/Berichte/Tabakkonsum_Schweiz/Forschungsbericht/Gesamt_Tabakkonsum_10_dt.pdf



Deskriptives Forschungsdesign: Längsschnittstudie

Prinzip der deskriptiven Längsschnittmethode:

Beschreibung der Ausprägung eines Merkmals anhand einer oder verschiedener möglichst repräsentativer Stichprobe(n) **zu verschiedenen Messzeitpunkten**

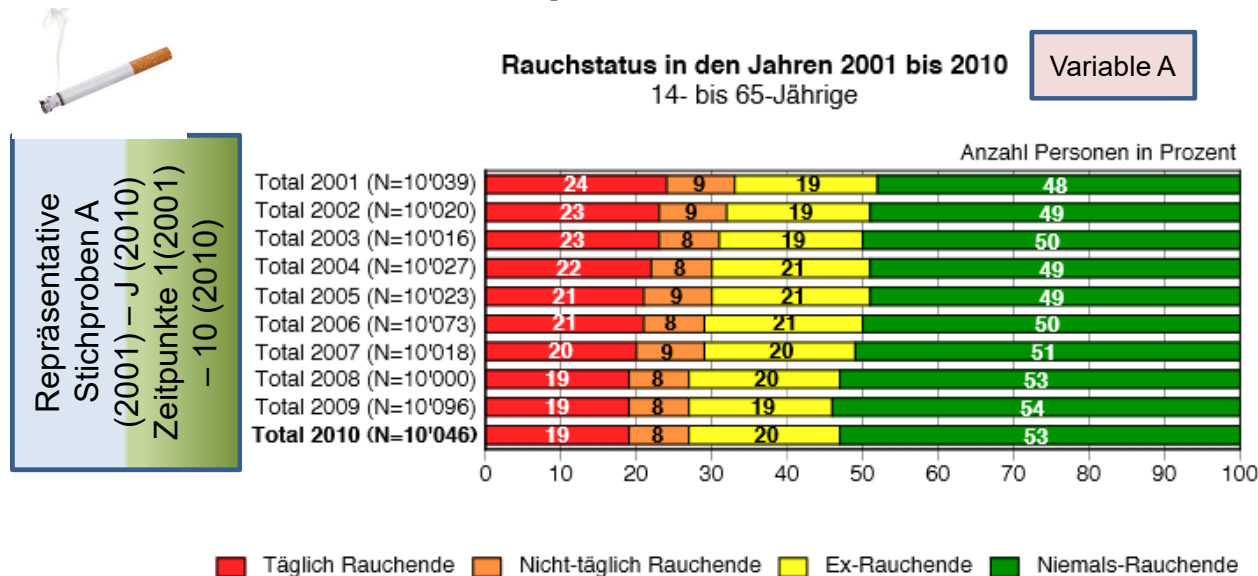
Beispiel:

Repräs. Stichprobe A	Repräs. Stichprobe B	Repräs. Stichprobe C	Repräs. Stichprobe D
Variable A	Variable A	Variable A	Variable A
Zeitpunkt t1	Zeitpunkt t2	Zeitpunkt t3	Zeitpunkt t4

Anmerkung: Repräs. = Repräsentativ

Deskriptives Forschungsdesign: Längsschnittstudie

Beschreibung der Ausprägung eines Merkmals anhand einer oder verschiedener möglichst repräsentativer Stichprobe(n) **zu verschiedenen Messzeitpunkten**



Daten aus dem
Schweizerischen
Tabakmonitoring

http://www.tabakmonitoring.ch/Berichte/Tabakkonsum_Schweiz/Forschungsbericht/Gesamt_Tabakkonsum_10_dt.pdf

Abbildung 5.2: Rauchstatus der 14- bis 65-jährigen Wohnbevölkerung in den Jahren 2001-2010



Lernziele erreicht?

Am Ende der Veranstaltung ...

... können Sie verschiedene Formen der Stichprobenziehung definieren, voneinander abgrenzen und die jeweiligen Vor- und Nachteile benennen.

... können Sie den Unterschied zwischen Labor- und Feldstudien einem Laien erklären sowie auch, welche Vor- und Nachteile jeweils verbunden sind.

... sind Sie in der Lage, interne und externe Validität zu definieren und können erklären, in welchem Verhältnis sie zueinander stehen.

... können Sie Gefährdungen der internen und der externen Validität und entsprechende Umgangs- / Lösungsmöglichkeiten erklären und Beispiele dafür generieren.

... wissen Sie, was unter einem quer- und einem längsschnittlichen deskriptiven Forschungsdesign zu verstehen ist. Sie können einem Laien erklären, welche Fragestellungen Sie mit diesen verschiedenen Designs beantworten können und welche nicht sowie welche Vor- und Nachteile mit den jeweiligen Designs verbunden sind.