



**Universität  
Zürich** <sup>UZH</sup>

**Psychologisches Institut**

---

# **Vorlesung Forschungsmethoden**

**13.12.2018**

Urte Scholz

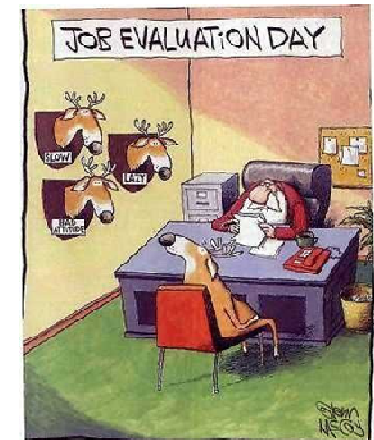
## Evaluation der Veranstaltung

Ich freue mich über Ihre Rückmeldung zur Veranstaltung.

<https://qmsl.uzh.ch/de/A7PXR>

Die Evaluation läuft bis zum 17.12.

Vielen Dank für Ihre Teilnahme.





## Lernziele der heutigen Veranstaltung

Am Ende der Veranstaltung ...

... können Sie die zentralen Unterschiede sowie die jeweiligen Vor- und Nachteile von within- und between-subjects designs einem Laien erklären.

... sind Sie in der Lage, mögliche Störeffekte bei within-subjects designs zu definieren und entsprechende Kontroll- bzw. Umgangsmöglichkeiten zu erklären.

... können Sie non-experimentelle, quasiexperimentelle und experimentelle Designs voneinander unterscheiden und Beispiele für verschiedene Arten dieser Designs herleiten.

... haben Sie einen umfassenden Überblick über die verschiedenen Forschungsdesigns und können entscheiden, welches Design Sie bei welcher Fragestellung anwenden.



## Versuchsplan

= logischer Aufbau einer empirischen Untersuchung im Hinblick auf Hypothesenprüfung.  
(Huber, 2013)

vier Entscheidungen (Hussy et al., 2013):

1. vollständige oder unvollständige Pläne
2. Bestimmung der Anzahl der Beobachtungen pro Zelle
3. interindividuelle oder intraindividuelle Bedingungsvariation
4. randomisierte oder nichtrandomisierte Zuordnung der Vpn zu den Zellen

Interindividuell = Zwischensubjekt- /between-subjects designs

→ Intraindividuell = Innersubjekt- / within-subjects / Messwiederholungs- / repeated measures designs



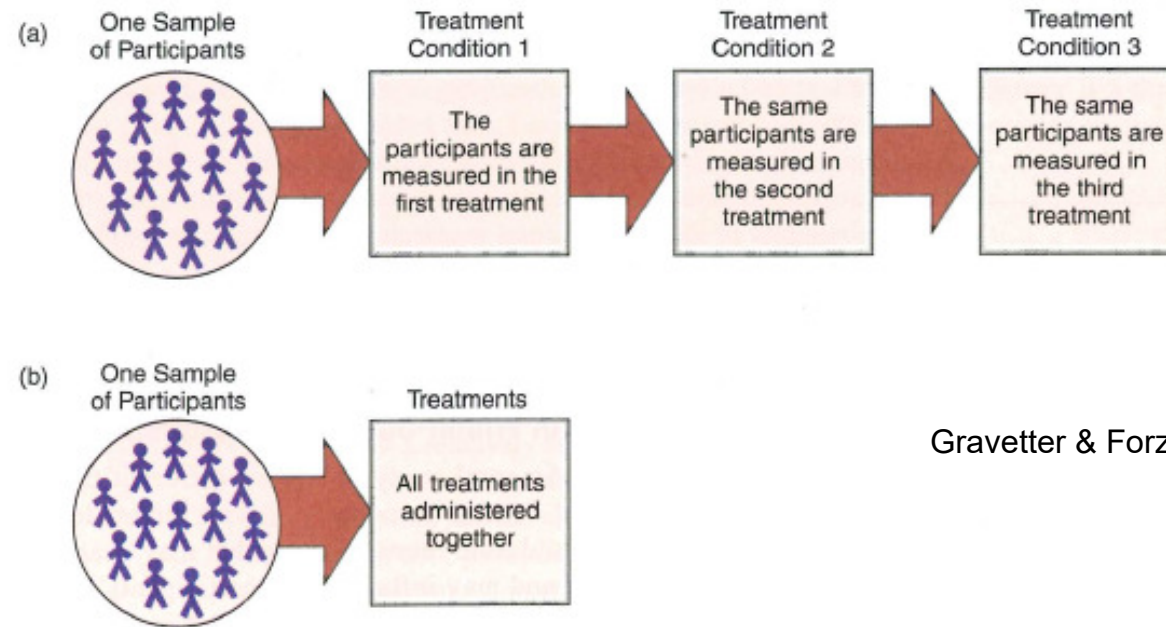
## Vor- und Nachteile von within- und between-subject designs

angelehnt an Martin, 2008

	Within-Subjects	Between-Subjects
Vorteile		
Nachteile		<ul style="list-style-type: none"><li>• Immer <b>potentielle</b> Konfundierung von Personvariablen mit experimenteller Bedingung</li></ul>

## Within-Subjects Designs

In einem within-subjects design werden zwei oder mehrere experimentelle Treatments innerhalb der gleichen Individuen verglichen (nach Gravetter & Forzano, 2018)



Gravetter & Forzano, 2018, S. 213

**FIGURE 9.1**

### Two Possible Structures for a Within-Subjects Design

The same group of individuals participates in all of the treatment conditions. Because each participant is measured in each treatment, this design is sometimes called a repeated-measures design. The different treatments can be administered sequentially, with participants receiving one treatment condition followed, at a later time, by the next treatment (a). It also is possible that the different treatment conditions are administered all together in one experimental session (b). Note: All participants go through the entire series of treatments but not necessarily in the same order.



## Beispiel within-subjects Design



<https://www.youtube.com/watch?v=I589B5dIPn4>



**TABLE 9.3**

**Hypothetical Data Showing the Results from a Between-Subjects Experiment and a Within-Subjects Experiment**

The two sets of data use exactly the same numerical scores.

**(a) Between-Subjects Experiment—Three Separate Groups**

Treatment I		Treatment II		Treatment III	
(John)	20	(Sue)	25	(Beth)	30
(Mary)	31	(Tom)	36	(Bob)	38
(Bill)	51	(Dave)	55	(Don)	59
(Kate)	62	(Ann)	64	(Zoe)	69
Mean =	41	Mean =	45	Mean =	49

**(b) Within-Subjects Experiment—One Group in All Three Treatments**

Treatment I		Treatment II		Treatment III	
(John)	20	(John)	25	(John)	30
(Mary)	31	(Mary)	36	(Mary)	38
(Bill)	51	(Bill)	55	(Bill)	59
(Kate)	62	(Kate)	64	(Kate)	69
Mean =	41	Mean =	45	Mean =	49

**TABLE 9.4**

**Removing Individual Differences from Within-Subjects Data**

This table shows the same data from Table 9.3b, except that we have eliminated the individual differences from the data. For example, we subtracted 20 points from each of Kate's scores to make her more "average," and we added 20 points to each of John's scores to make him more "average." This process of eliminating individual differences makes the treatment effects much easier to see.

Treatment I		Treatment II		Treatment III	
(John)	40	(John)	45	(John)	50
(Mary)	41	(Mary)	46	(Mary)	48
(Bill)	41	(Bill)	45	(Bill)	49
(Kate)	42	(Kate)	44	(Kate)	49
Mean =	41	Mean =	45	Mean =	49

Gravetter &  
Forzano, 2018,  
S. 227, 228



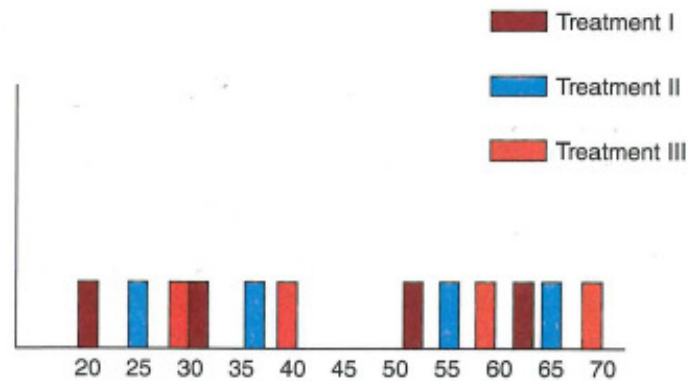
**FIGURE 9.2**

**Removing Individual Differences from Within-Subjects Data**

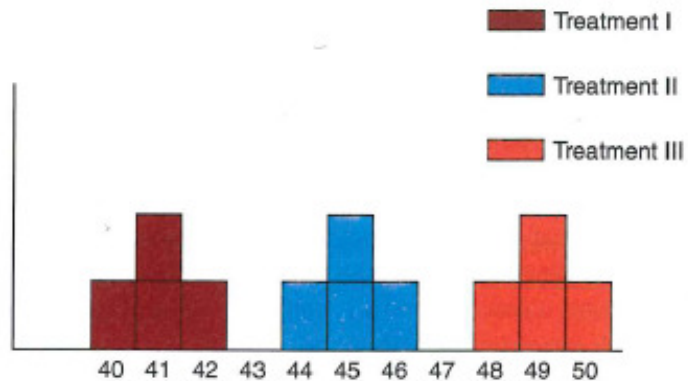
(a) The original data, which include the individual differences among the four participants.

(b) The individual differences have been removed by adjusting each participant's scores. When the individual differences are removed, it is much easier to see the differences between treatments.

(a) Data Including Individual Differences (from Table 9.3b)



(b) Data with Individual Differences Removed (from Table 9.4)



Gravetter &  
Forzano,  
2018, S. 229



## Vor- und Nachteile von between- und within-subjects Designs

Stellen Sie sich vor, Sie möchten prüfen, ob eine bestimmte Strategie zur Emotionsregulation (Perspektivenübernahme) effektiv ist, um negative Gefühle zu reduzieren. In der Experimentalbedingung 1 trainieren Sie diese Emotionsregulationsstrategie. In der Experimentalbedingung 2 lassen Sie die Teilnehmenden Zeitung lesen (Kontrollbedingung). Ihre AV sind die berichteten emotionalen Reaktionen auf Bilder, die neutrale, positive und negative Gefühle hervorrufen.

Sie wählen ein within-subjects design zur Überprüfung dieser Fragestellung. Stellen Sie die Vor- und Nachteile dieser Wahl dar.



## Vor- und Nachteile von within- und between-subject designs angelehnt an Martin, 2008

	Within-Subjects	Between-Subjects
Vorteile	<ul style="list-style-type: none"><li>• kein Problem mit personenspezifischer Konfundierung von Störvariablen</li><li>• Weniger Teilnehmende notwendig</li><li>• Effekte der UV werden eher entdeckt als in between-subjects designs (grössere «Power»)</li></ul>	<ul style="list-style-type: none"><li>• die Teilnahme an einer experimentellen Bedingung hat keinen Einfluss auf die Teilnahme an einer anderen experimentellen Bedingung</li><li>• Innerhalb einer Bedingung können mehr Daten gesammelt werden, als wenn jemand an mehreren Bedingungen teilnimmt</li></ul>
Nachteile	<ul style="list-style-type: none"><li>• Dropout bei mehreren Messzeitpunkten</li></ul> Gefährdung der internen Validität durch <ul style="list-style-type: none"><li>• zeitliche Veränderung</li><li>• Positionseffekte</li><li>• Übertragungseffekte (carry over effects)</li></ul>	<ul style="list-style-type: none"><li>• Immer potentielle Konfundierung von Personvariablen mit experimenteller Bedingung</li></ul>

## Spezielle Störeffekte bei wiederholter Messung

(Huber, 2013)



### Begriffsklärung Messwiederholung:

- generell: mehrere Messungen pro Person zu unterschiedlichen Zeitpunkten
- bezogen auf within-subjects Experimente: Personen nehmen mehrfach innerhalb des gleichen Experiments in verschiedenen Bedingungen teil (Hussy et al., 2013)

### Spezielle Störeffekte

- Zeitliche Veränderung ausserhalb der Untersuchungssituation (s. auch Gefährdungen der internen Validität)
- Positionseffekte (Position der experimentellen Bedingung)
- Übertragungseffekte (»carry over effects«)

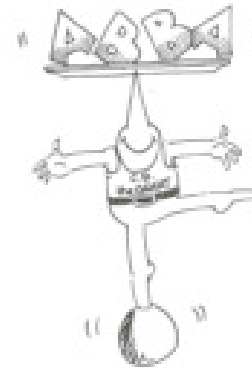
## Kontrolle der Störeffekte in within-subjects designs

### Positionseffekt:

„Ein Positionseffekt (Stellungseffekt) ist eine Störvariable, die von der Position einer experimentellen Bedingung in der Reihenfolge her bestimmt ist.“ (Huber, 2013, S. 171)

### Kontrolle von Positionseffekten:

- Ausbalancieren (vollständiges / unvollständiges)



Martin, 2008, S. 157

## Vollständiges Ausbalancieren (Huber, 2013, S. 175)

	Position			Reaktionszeit für Bedingung		
	1	2	3	B	L	T
Reihenfolge 1	B	L	T	RZ + PE(1)	RZ + PE(2)	RZ + PE(3)
Reihenfolge 2	B	T	L	RZ + PE(1)	RZ + PE(3)	RZ + PE(2)
Reihenfolge 3	L	B	T	RZ + PE(2)	RZ + PE(1)	RZ + PE(3)
Reihenfolge 4	L	T	B	RZ + PE(3)	RZ + PE(1)	RZ + PE(2)
Reihenfolge 5	T	B	L	RZ + PE(2)	RZ + PE(3)	RZ + PE(1)
Reihenfolge 6	T	L	B	RZ + PE(3)	RZ + PE(2)	RZ + PE(1)

Summe der Positionseffekte:

2 PE(1)	2 PE(1)	2 PE(1)
2 PE(2)	2 PE(2)	2 PE(2)
2 PE(3)	2 PE(3)	2 PE(3)

**Abbildung 10:** Summe der Positionseffekte für jede von drei experimentellen Bedingungen bei der Methode des vollständigen Ausbalancierens.

RZ = Reaktionszeit; PE(i) = Positionseffekt; L = Lichtsignal, T = Tonsignal, B = Berührungssignal  
Gemessene Reaktionszeit = RZ + PE(i)



## Problem des vollständigen Ausbalancierens

Zahl der möglichen  
Reihenfolgen bei  $n$   
experimentellen  
Bedingungen =  $n!$

■ TABLE 8-4

Completely Counterbalanced Design for Two-, Three-,  
and Four-Level Independent Variables

<i>Two levels of independent variable</i>		<i>Three levels of independent variable</i>	
<i>Number</i>	<i>Order of levels</i>	<i>Number</i>	<i>Order of levels</i>
1	AB*	1	ABC
2	BA	2	ACB
		3	BCA
		4	BAC
		5	CAB
		6	CBA
<i>Four levels of independent variable</i>			
<i>Number</i>	<i>Order of levels</i>	<i>Number</i>	<i>Order of levels</i>
1	ABCD	13	CABD
2	ABDC	14	CADB
3	ACBD	15	CBAD
4	ACDB	16	CBDA
5	ADCB	17	CDAB
6	ADBC	18	CDBA
7	BACD	19	DABC
8	BADC	20	DACB
9	BCAD	21	DBAC
10	BCDA	22	DBCA
11	BDAC	23	DCAB
12	BDCA	24	DCBA

\*The letters A, B, C, and D represent the levels.

Martin, 2008,  
S. 163





## Unvollständiges Ausbalancieren (Huber, 2013)

- Zufallsauswahl
- Spiegelbildmethode
- Lateinisches Quadrat

### Zufallsauswahl:

- randomisierte Auswahl aus allen möglichen Reihenfolgen
- jede VP erhält andere Reihenfolge
- bei grossen Stichproben
- Kontrolle über alle VPn hinweg



## Unvollständiges Ausbalancieren (Huber, 2013)

### Spiegelbildmethode (ABBA counterbalancing):

- Wahl einer Reihenfolge, die gespiegelt wird
- AB – BA , ABC – CBA , etc.
- alle VPn bekommen gleiche Sequenz
- dadurch Kontrolle des Positionseffekts pro Person

### Zwei Voraussetzungen:

- a) jede experimentelle Bedingung muss pro Person zweimal durchführbar sein
- b) der Positionseffekt muss linear sein  
(→ muss vorher bekannt sein, bzw. geprüft werden)  
(→ sehr gute Herleitung dieser Problemstellung bei Martin, 2008, S. 156 ff)



Huber, 2013,  
S. 176



## Unvollständiges Ausbalancieren (Huber, 2013)

### Lateinisches Quadrat

- Genauso viele Reihenfolgen wie experimentelle Bedingungen
- Gleichmässiges Verteilen der VPn auf Reihenfolgen
- Kontrolle über alle VPn hinweg



## Beispiel Lateinisches Quadrat

(Martin, 2008, S. 163)

	Order of Presentation			
	1st	2nd	3rd	4th
Participant 1	<b>Chicago</b>	Courier	Geneva	Times
Participant 2	Courier	Times	<b>Chicago</b>	Geneva
Participant 3	Times	Geneva	Courier	<b>Chicago</b>
Participant 4	Geneva	<b>Chicago</b>	Times	Courier

**FIGURE 8-6** A balanced Latin Square for ordering the presentation of four print fonts to at least four readers



## Unvollständiges Ausbalancieren (Huber, 2013)

### Lateinisches Quadrat - Vorteil:

kann in faktoriellen Versuchsplan integriert werden:

- Position der einzelnen experimentellen Bedingungen wird als UV eingeführt
  - VPn werden in Gruppen zugeteilt
- Überprüfung (nicht nur Kontrolle) möglich

## Beispiel lateinisches Quadrat als faktorieller Versuchsplan

	Position 1				Position 2				Position 3				Position 4			
	exp. Bed.				exp. Bed.				exp. Bed.				exp. Bed.			
	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
<b>Gruppe 1</b>	x					x					x					x
<b>Gruppe 2</b>		x						x	x						x	
<b>Gruppe 3</b>				x			x			x			x			
<b>Gruppe 4</b>			x		x							x		x		

**Abbildung 11:** Beispiel für ein Lateinisches Quadrat als faktorieller Versuchsplan

In diesem Versuchsplan werden nur die mit einem x gekennzeichneten Kombinationen realisiert. Dreiviertel der Zellen des Versuchsplanes sind also leer.

Huber, 2013



## Kontrolle von Carry-Over-Effekten (Huber, 2013)

### Carry-over Effekt:

«Ein carry-over Effekt (Übertragungseffekt) ist eine Störvariable, die davon herrührt, dass eine frühere experimentelle Bedingung eine spätere **inhaltlich** beeinflusst.» (Huber, 2013, S. 171)

- Kontrolle von Positionseffekten  $\neq$  Kontrolle von Carry-over-Effekten

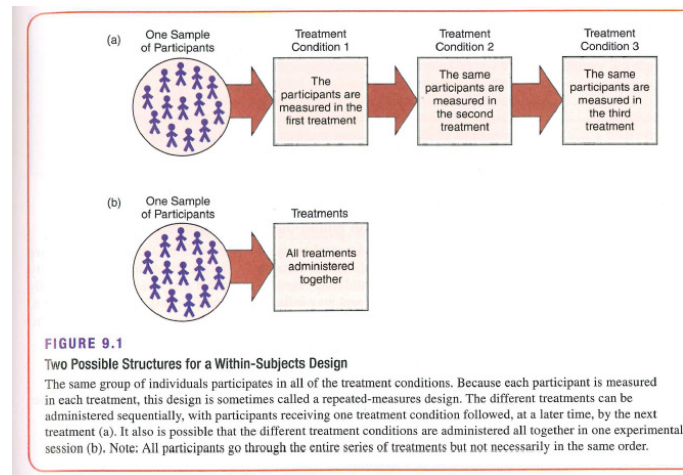
### Alternativen:

- Ursachen des carry-over-Effekts bekannt  $\rightarrow$  Versuch der Beseitigung der Ursache
- Between-subjects Design wählen
- Notfalls Zeit zwischen den Bedingungen einsetzen («wash-out period»)



## Fazit experimentelle Within-Subjects Designs

- Etliche Vorteile gegenüber between-subjects designs
- Aber: Problem der spezifischen Störvariablen (v.a. Positions- / Carry-over-Effekte)  
→ Kontrolle dieser Störvariablen oder, wenn nicht möglich, doch between-subjects design



Gravetter & Forzano,  
2018, S. 213



## Forschungsdesign wählen (Gravetter & Forzano, 2018)

### Forschungsdesigns - Arten:

- ✓ Deskriptiv → reine Beschreibung einzelner Merkmale
- ✓ Korrelativ → Zusammenhang / Zusammenhänge zwischen Variablen, keine Erklärung
- (✓) Nicht-experimentell → 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



## Non-experimentelle, quasiexperimentelle und experimentelle Designs (Gravetter & Forzano, 2018)

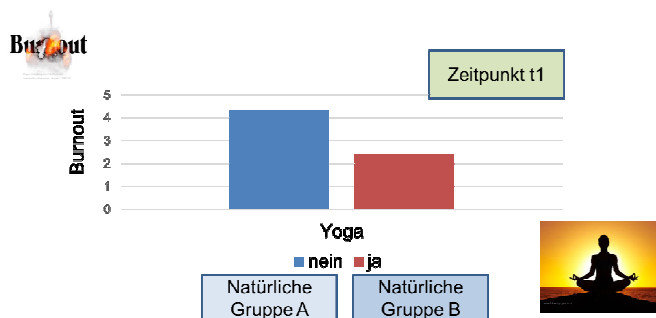
Nonexperimentell	Quasiexperimentell	Experimentell
Natürliche Gruppen / non-equivalent groups		Randomisierte Zuteilung zu Bedingungen
Keine Kontrolle von Störvariablen	Versuch der Kontrolle von Störvariablen	Kontrolle von Störvariablen

### Natürliche Gruppen?

- a) definiert über Charakteristika der Teilnehmenden  
→ between-subjects designs
- b) definiert über Zeit (prä-post)  
→ within-subjects designs

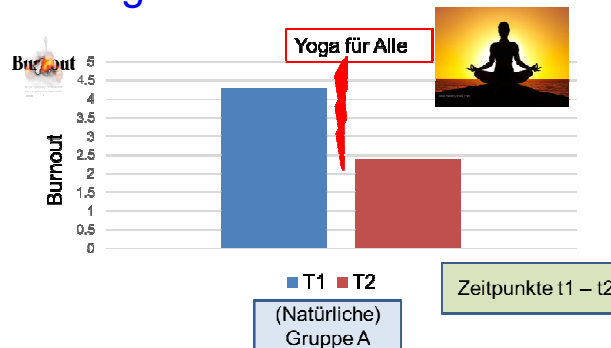
## Non-experimentelle Designs

Querschnitts-Design: Differential research design / Ex-post-facto design



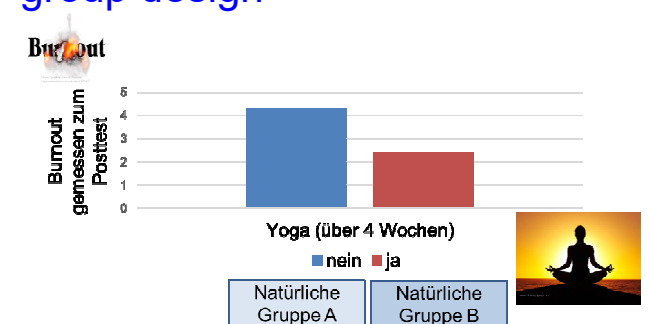
- Vergleich zweier Gruppen zu **einem** Zeitpunkt
  - Keinerlei Kontrolle möglicher Störvariablen
- nonexperimentell

Längsschnitt-Design: One group pretest-posttest design



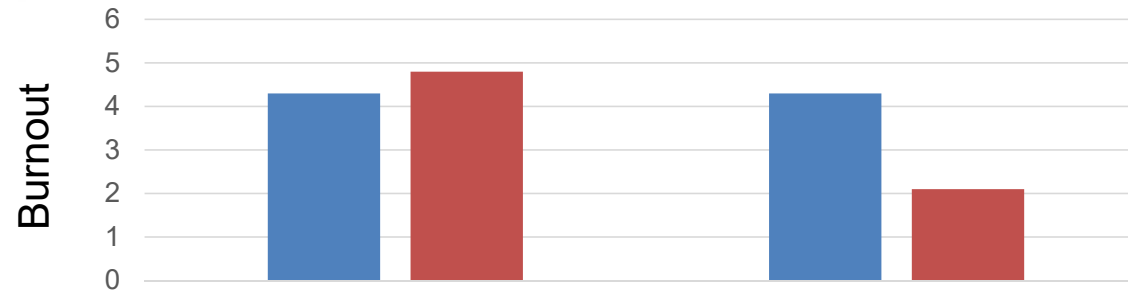
- Vorher-Nachher Vergleich nach Massnahme innerhalb einer Gruppe (within-subjects design)
  - Keinerlei Kontrolle möglicher Störvariablen
- nonexperimentell

Längsschnitt-Design: Posttest-only nonequivalent control group design



- Vergleich der Gruppen **nach** Massnahme
  - Keinerlei Kontrolle möglicher Störvariablen
- nonexperimentell

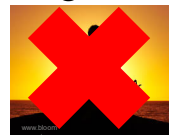
## Quasi-experimentelles Design: Pretest-posttest nonequivalent control group design (Gravetter & Forzano, 2018)



Achtung: beim  
quasiexperimentellen  
Vorgehen KEINE  
randomisierte Zuteilung  
zu Gruppen



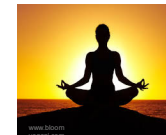
Yoga nein



Natürliche  
Gruppe A

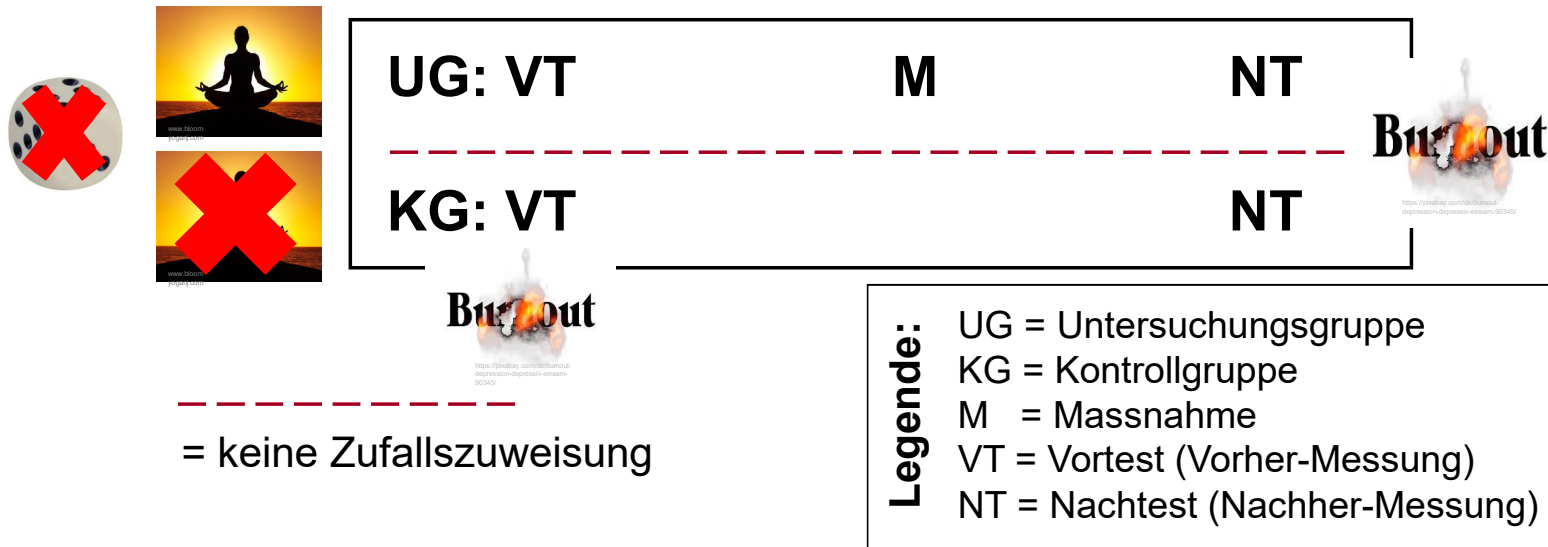
■ T1 ■ T2

Yoga ja



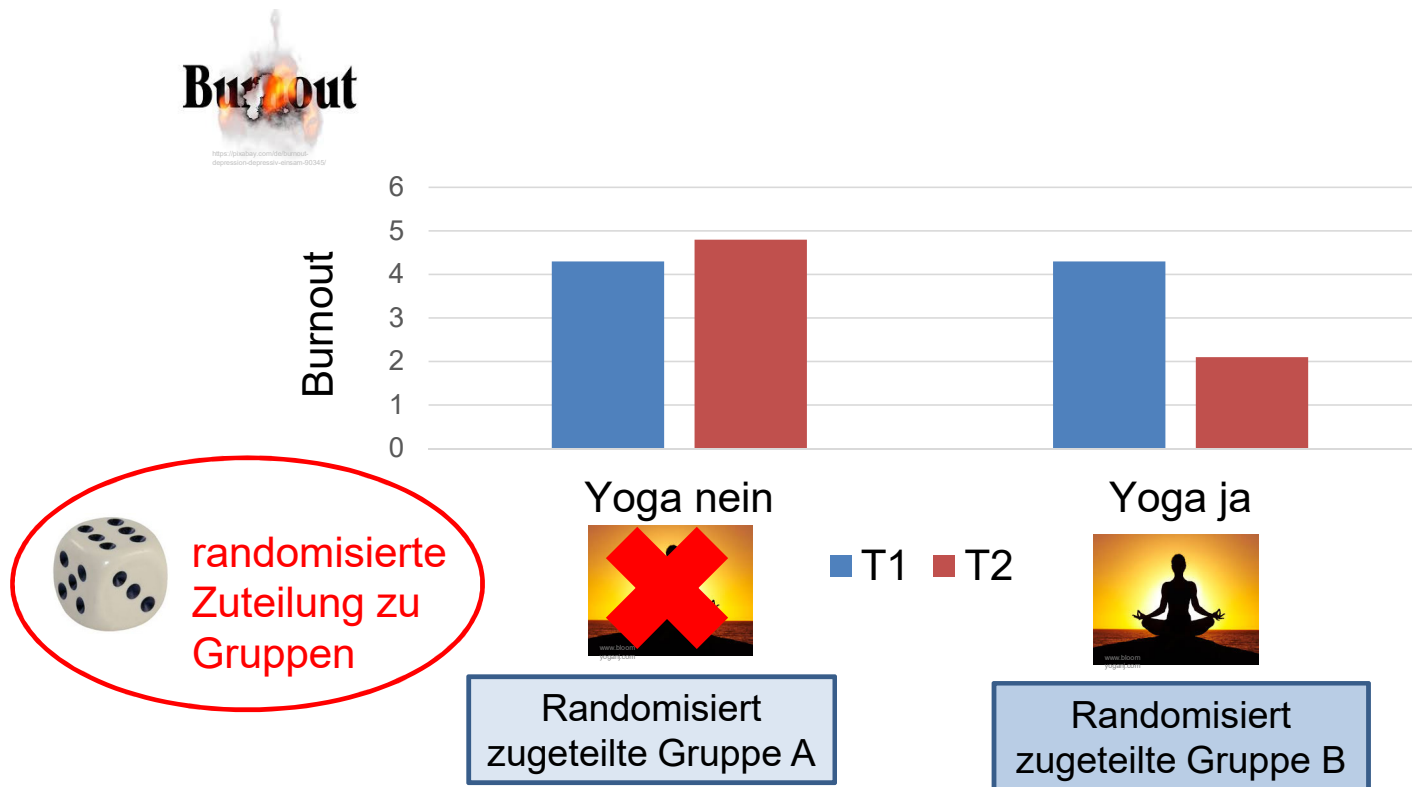
Natürliche  
Gruppe B

## Quasi-experimentelles Design: Pretest-posttest nonequivalent control group design



- Durch Vortest → Vergleich der Gruppen vor der Massnahme möglich
  - aktive Manipulation mind. 1 UV
  - Einfluss zeitlicher Veränderungen als Bedrohung der internen Validität minimiert, aber nicht ganz ausgeschaltet
- quasiexperimentell

## Zum Vergleich: Beispiel für experimentelles between-subjects Design





## Zum Vergleich: Beispiel für experimentelles between-subjects Design



randomisierte  
Zuteilung zu  
Gruppen



UG: VT

M

NT

KG: VT

NT



**Legende:**

UG = Untersuchungsgruppe

KG = Kontrollgruppe

M = Massnahme

VT = Vortest (Vorher-Messung)

NT = Nachtest (Nachher-Messung)

- randomisierte Zuteilung zu den Bedingungen
  - aktive Manipulation mind. 1 UV
  - Kontrolle von weiteren Störvariablen
- experimentell



## Non-experimentelle, quasiexperimentelle und experimentelle Designs

(Gravetter & Forzano, 2018)

Nonexperimentell	Quasiexperimentell	Experimentell
Natürliche Gruppen / non-equivalent groups		Randomisierte Zuteilung zu Bedingungen
Keine Kontrolle von Störvariablen	Versuch der Kontrolle von Störvariablen	Kontrolle von Störvariablen

### Natürliche Gruppen?

- a) definiert über Charakteristika der Teilnehmenden  
→ between-subjects designs
- b) definiert über Zeit (prä-post)  
→ within-subjects designs



Universi  
Zürich<sup>UZH</sup>

Psych

## Zusammenfassende Abgrenzung zwischen den verschiedenen Forschungsdesigns (Gravetter & Forzano, 2018)

Was	ja	nein
Deskriptiv	reine Beschreibung einzelner Merkmale (z.B. Häufigkeiten)	keine Zusammenhänge, keine Gruppenunterschiede, keine experimentelle Manipulation, keine Kausalitätsüberprüfung
Korrelativ	Zusammenhang / Zusammenhänge zwischen Variablen	keine Gruppenunterschiede, keine experimentelle Manipulation, keine Kausalitätsüberprüfung
Nicht-experimentell	Unterschiede natürlicher Gruppen auf AV; im within-subjects-Kontext: Unterschiede einer Gruppe über die Zeit	keine Randomisierung (sondern natürliche Gruppen); keinerlei Kontrolle von Störvariablen;
Quasi-experimentell	Versuch einer Annäherung an Kausalitätsüberprüfung durch experimentelle Manipulation mind. 1 UV und Versuch der Kontrolle von Störvariablen → mehr als reine Beschreibung von Mittelwertsunterschieden	keine Randomisierung (sondern natürliche Gruppen); keine <i>vollständige</i> Kontrolle von Störvariablen; deshalb Kausalaussagen nie abschliessend möglich
Experimentell	Kausalitätsüberprüfung (= Ursache-Wirkungs-Zusammenhänge) zwischen UVs und AVs durch randomisierte Zuteilung zu den Bedingungen (between oder within) und aktive Manipulation mind. 1 UV	Achtung, auch hier gewisse Gefährdungen der internen Validität durch entsprechende Störvariablen möglich



## Zusammenfassende Abgrenzung zwischen den verschiedenen Forschungsdesigns (Gravetter & Forzano, 2018)

Was	ja	nein	
Deskriptiv	reine Beschreibung einzelner Merkmale (z.B. Häufigkeiten)	keine Zusammenhänge, keine Gruppenunterschiede, keine experimentelle Manipulation, keine	
Korrelativ	<p>Im Lehrbuch von Hussy et al. (2013) werden unter dem Begriff nicht-experimentelle Forschung deskriptive und korrelative Ansätze zusammengefasst. Der non-experimentelle Ansatz (z.B. ex post facto design) findet sich bei den experimentellen Ansätzen.</p> <p>→ Die Differenzierung bei Gravetter &amp; Forzano (2018) ist für das Lernen der Unterschiede zwischen den Forschungsstrategien und –designs m. E. günstiger.</p>		telle
Nicht-experimentell			uppen);
Quasi-experimentell			uppen); ; möglich
	von Mittelwertsunterschieden		
Experimentell	Kausalitätsüberprüfung (= Ursache-Wirkungs-Zusammenhänge) zwischen UVs und AVs durch randomisierte Zuteilung zu den Bedingungen (between oder within) und aktive Manipulation mind. 1 UV	Achtung, auch hier gewisse Gefährdungen der internen Validität durch entsprechende Störvariablen möglich	



## Forschungsdesign wählen (Gravetter & Forzano, 2018)

### Forschungsdesigns - Arten:

- ✓ Deskriptiv → reine Beschreibung einzelner Merkmale
- ✓ Korrelativ → Zusammenhang / Zusammenhänge zwischen Variablen, keine Erklärung
- ✓ Nicht-experimentell → 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



**Kurze abschliessende Übung:**  
**Welches Forschungsdesign für welche Forschungsfrage? (bei mehreren Möglichkeiten, wählen Sie das Design mit der höchsten internen Validität)**

1. Gibt es Unterschiede in der Rechtschreibfähigkeit zwischen Jungen und Mädchen der 3. Klasse?
2. Wie hoch ist der Anteil an Schülerinnen und Schülern mit Migrationshintergrund im Kanton Zürich?
3. Fördert die Einnahme von Nikotinersatzpräparaten den Erfolg im Rauchstopp?
4. Ist die Teilnahme an einem Programm zur Förderung sozialer Kompetenzen effektiver für Nicht-Muttersprachlerinnen / Nicht-Muttersprachler als für Muttersprachlerinnen / Muttersprachler?
5. Wie stark ist der Zusammenhang zwischen sozialer Unterstützung und Wohlbefinden?



## Lernziele erreicht?

Am Ende der Veranstaltung ...

... können Sie die zentralen Unterschiede sowie die jeweiligen Vor- und Nachteile von within- und between-subjects designs einem Laien erklären.

... sind Sie in der Lage, mögliche Störeffekte bei within-subjects designs zu definieren und entsprechende Kontroll- bzw. Umgangsmöglichkeiten zu erklären.

... können Sie non-experimentelle, quasiexperimentelle und experimentelle Designs voneinander unterscheiden und Beispiele für verschiedene Arten dieser Designs herleiten.

... haben Sie einen umfassenden Überblick über die verschiedenen Forschungsdesigns und können entscheiden, welches Design Sie bei welcher Fragestellung anwenden.