

Modul M4 | P Experimentalpsychologisches Arbeiten

# Experimentalpsychologisches Arbeiten – Teil 2

## Methoden

### Wintersemester 2023/24

Christoph Scheffel, M.Sc.

[christoph\\_scheffel@tu-dresden.de](mailto:christoph_scheffel@tu-dresden.de) | BZW A437 | 0351 463-40336

Professur für Differentielle und Persönlichkeitspsychologie

# Gliederung

- ✓ Datenaufbereitung und Auswertung
- ✓ Wie schreibe ich einen Methodenteil?

# Datenaufbereitung und Datenauswertung

# Datenaufbereitung

## Arten von Daten

- Direkt nachdem eine Versuchsperson das Experiment beendet hat, erhalten wir **Rohdaten** (= Ursprungsaufzeichnungen)
- Diese Daten müssen in ein digitales Format übertragen werden. Sind diese Daten vollkommen unbearbeitet, handelt es sich um **Primärdaten** (= CSV Datei, die aus Labvanced exportiert werden kann → Auf OPAL zu finden)
- Primärdaten müssen verarbeitet werden. Häufig werden Daten, die den statistischen Analysen unterliegen zusammengefasst (= **Aggregierte Daten**), was aber immer mit Informationsverlust einher geht

# Datenaufbereitung

## Open Science

- um eine optimale Nachvollziehbarkeit im Sinne von Open Science zu erreichen und dabei möglichst wenig Informationsverlust erleiden möchten, sollten möglichst **Rohdaten** *oder* **Primärdaten** zur Verfügung gestellt werden.
- Aber: Um die Daten für statistische Analysen verwenden zu können, müssen sie vorverarbeitet werden, zum Beispiel aggregiert.
- Damit es auch hier nicht zu Informationsverlust kommt, sollten Bearbeitungsschritte der wissenschaftlichen Gemeinschaft zur Verfügung gestellt werden. Am einfachsten geht dies durch die Verwendung von R-Skripten

# Datenaufbereitung

## Rohdaten in Labvanced

# Datenaufbereitung

## Aggregierte Daten

- Primärdaten liefern einen sehr hohen Informationsgehalt, sind jedoch nicht geeignet um statistische Auswertungen durchzuführen. Dazu benötigt es aggregierte Daten, die folgender Struktur folgen (long format):

ID	$\Sigma$ Adv. Choices	Block	NFC Score	NFC Group	Geschlecht	Alter
101	3	1	3.5	1	2	20
101	7	2	3.5	1	2	20
101	10	3	3.5	1	2	20
101	12	4	3.5	1	2	20
101	13	5	3.5	1	2	20
102	5	1	2.8	2	2	21
102	8	2	2.8	2	2	21
102	10	3	2.8	2	2	21
102	14	4	2.8	2	2	21
102	17	5	2.8	2	2	21
103	4	1	2.7	2	1	19
...	...	...	...	...	...	...

# Datenaufbereitung

## Aggregierte Daten

- Wie müssen unsere Primärdaten zusammengefasst werden, bevor wir uns der statistischen Auswertung widmen können?
- Welche Zwischenschritte sind dabei nötig oder zumindest hilfreich?

# Datenauswertung

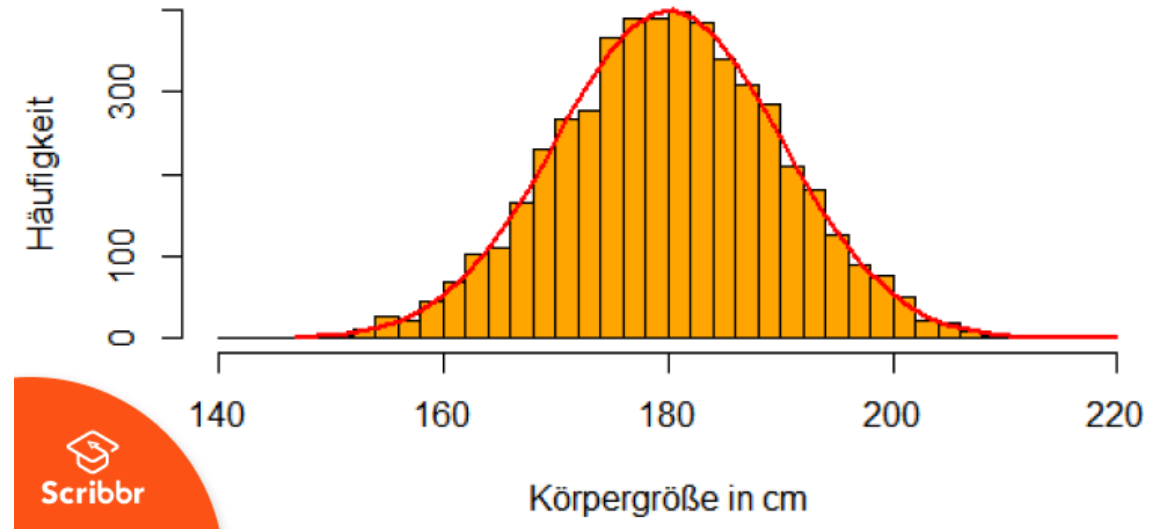
## ANOVA

ANOVA testet **Mittelwertsunterschiede** verschiedener Gruppen → Vergleich von **mehr als zwei Gruppen** gleichzeitig

Es werden nicht die Mittelwerte direkt, sondern **verschiedene Varianzanteile** analysiert  
→ ANOVA – ANalysis Of VAriance.

Beispiel: Varianz in Körpergröße

**Körpergrößen inklusive Normalverteilungskurve**

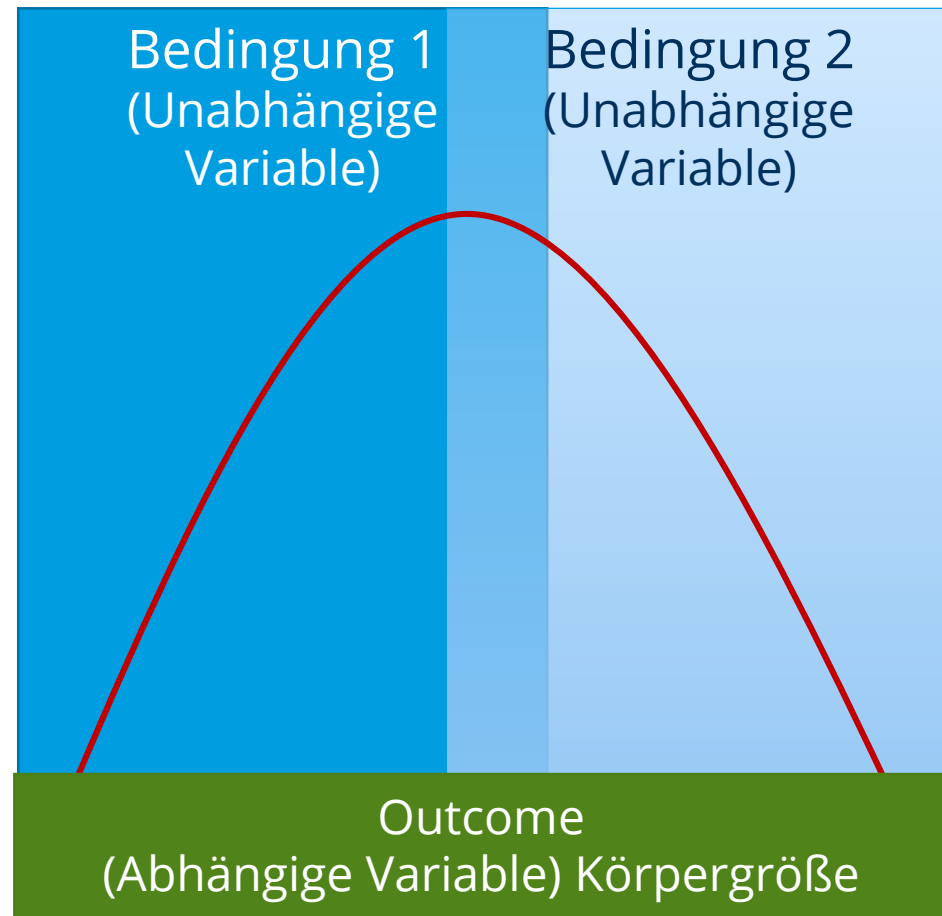


# Datenauswertung

## ANOVA

**Beispiel:** Woran könnte die Varianz in Körpergröße liegen?

Was könnte also die Unabhängige Variable sein?

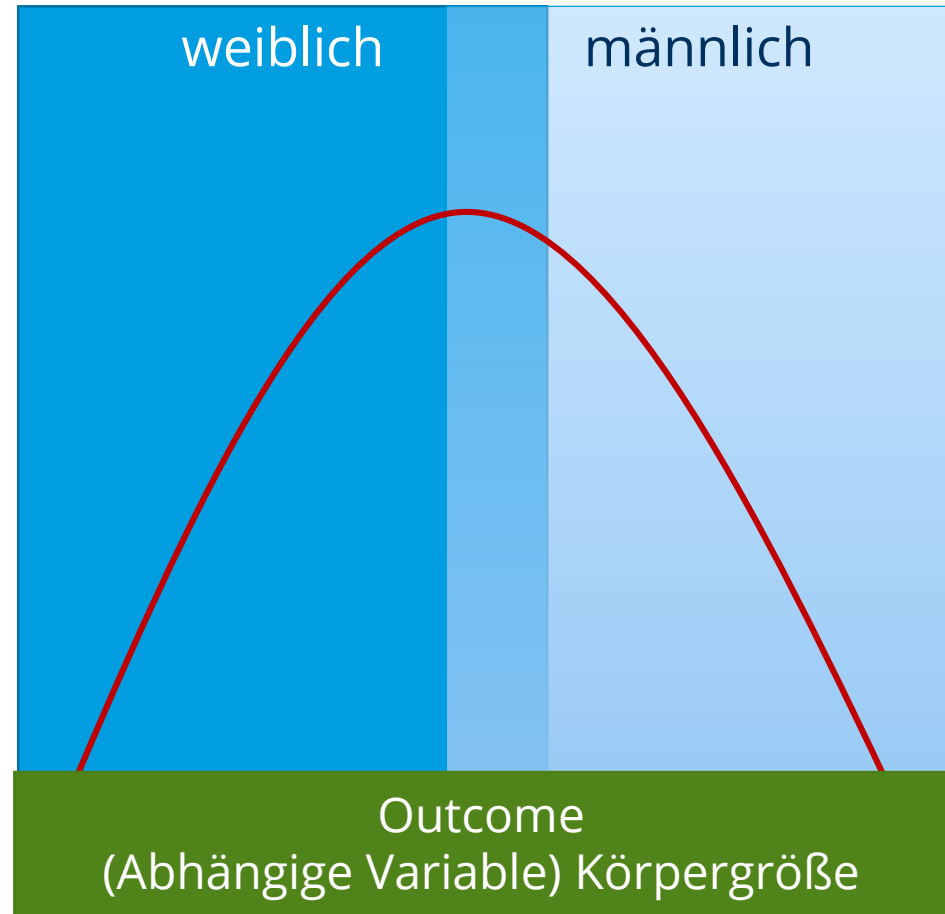


# Datenauswertung

## ANOVA

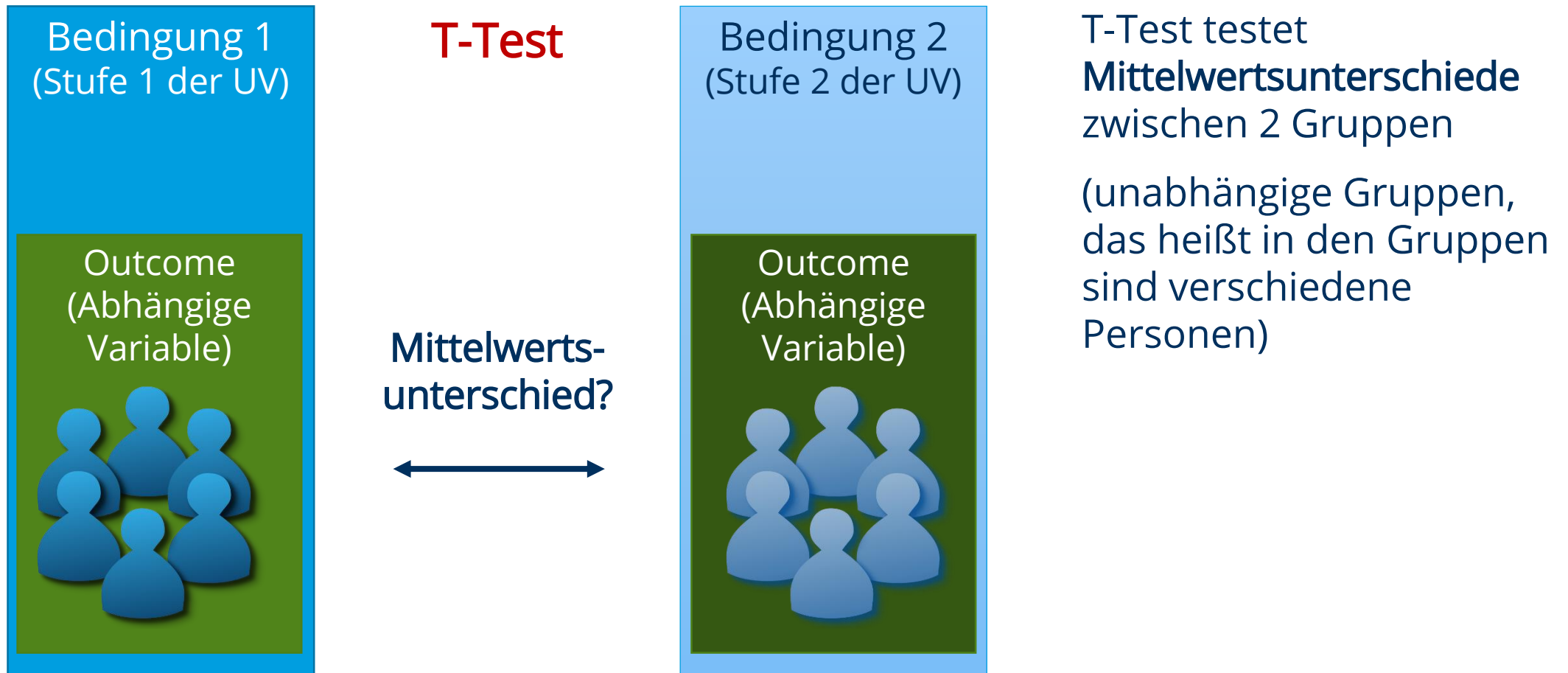
**Beispiel:** Woran könnte die Varianz in Körpergröße liegen?

Was könnte also die Unabhängige Variable sein?



# Datenauswertung

## ANOVA

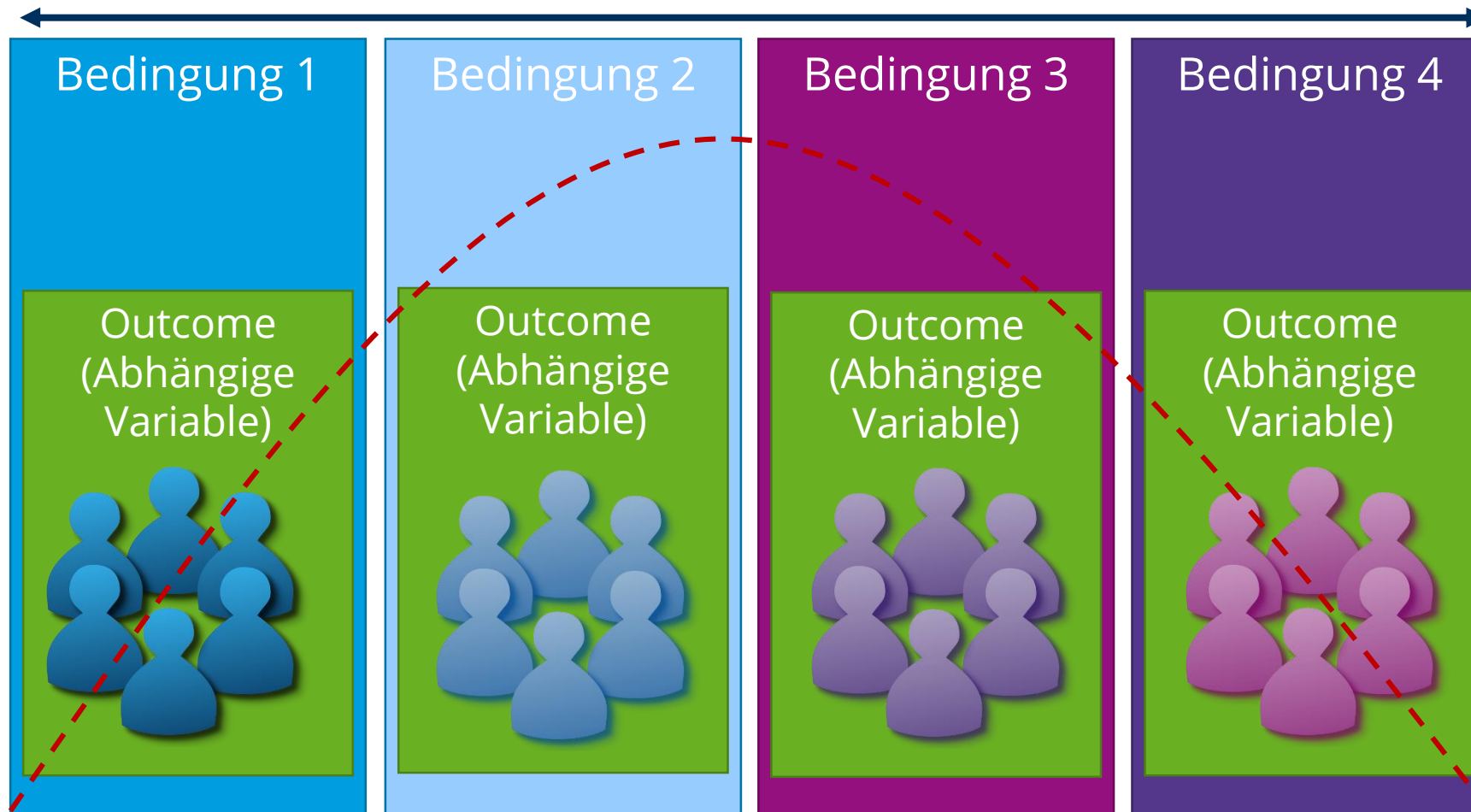


# Datenauswertung

## ANOVA

## ANOVA

Mittelwertsunterschiede?



# Datenauswertung

## ANOVA

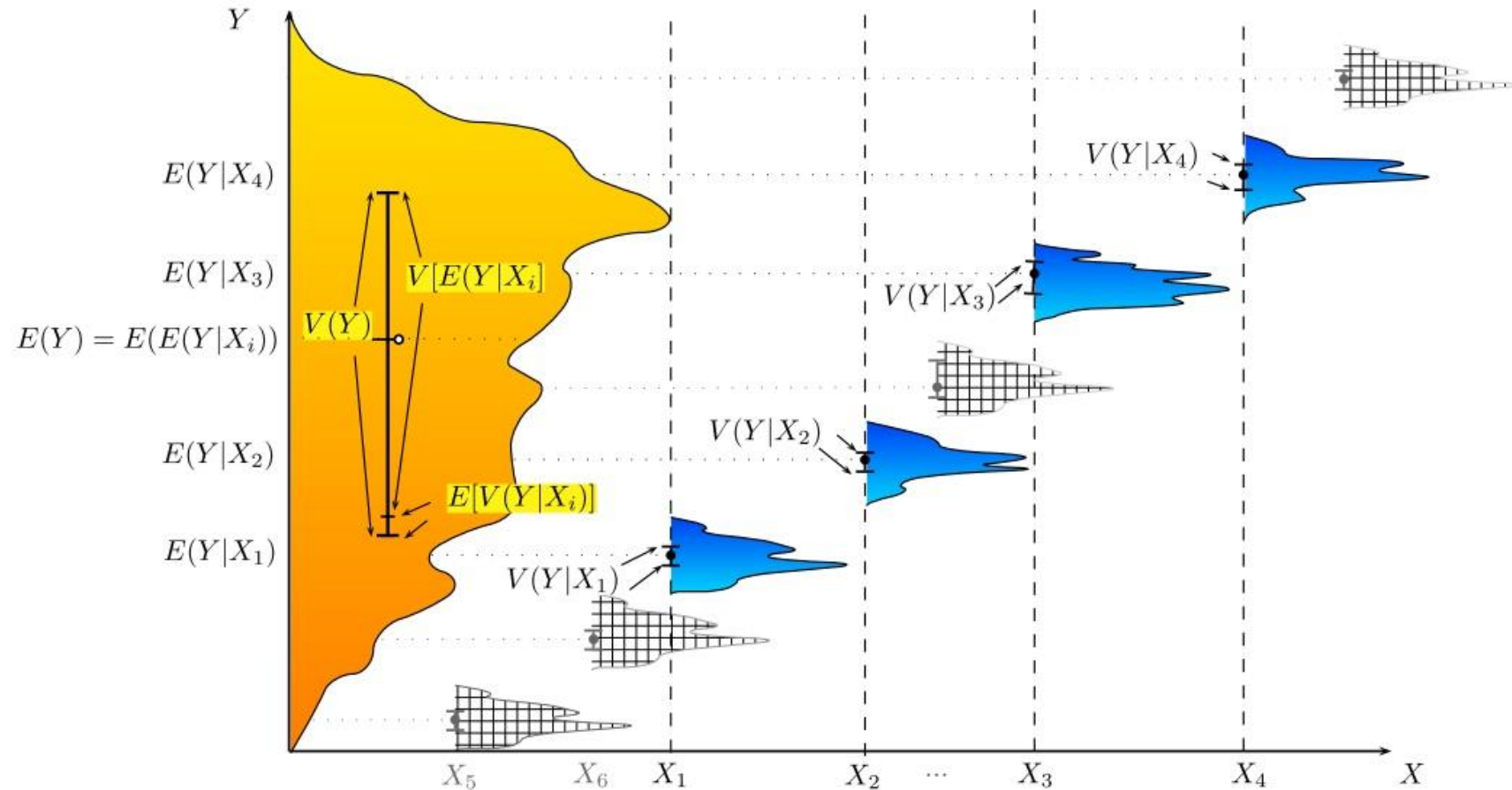


Figure 3: ANOVA : very good fit

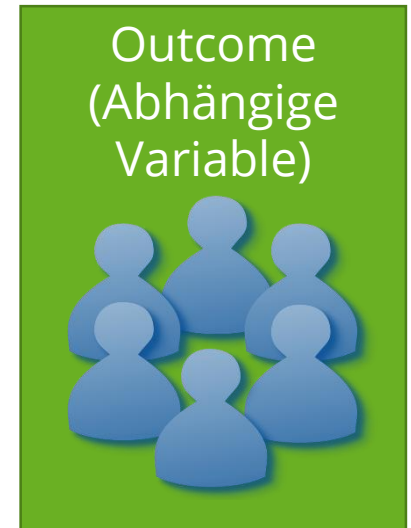
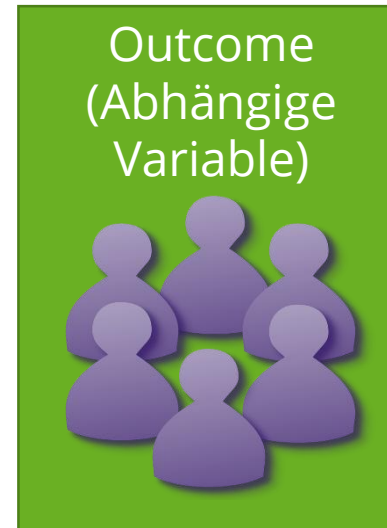
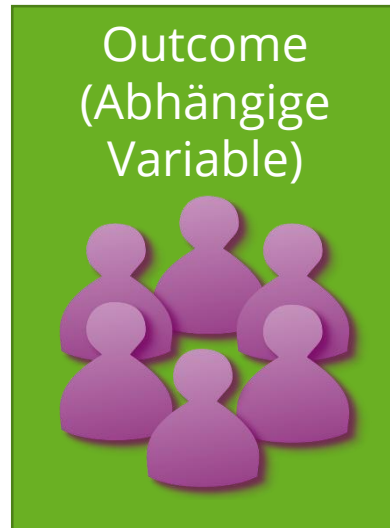
# Datenauswertung

## ANOVA

ANOVA testet **Mittelwertsunterschiede** in der Abhängigen Variable zwischen verschiedenen Gruppen → Vergleich von **mehr als zwei Gruppen** gleichzeitig

Es werden nicht die Mittelwerte direkt, sondern **verschiedene Varianzanteile** analysiert → ANOVA – ANalysis Of VAriance.

Ist eine Erklärung für die **Varianz** der Messwerte, dass Personen **bestimmten Gruppen/Bedingungen** zugeteilt wurden oder natürlich angehören?

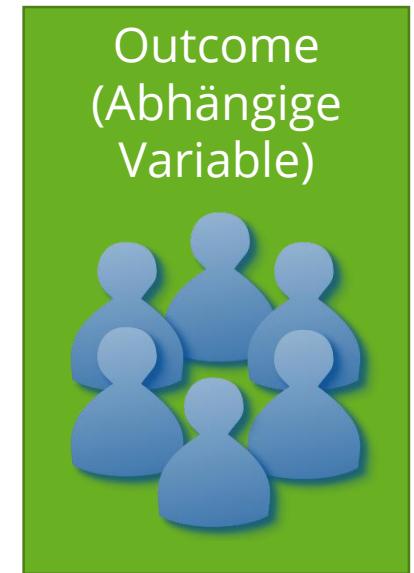
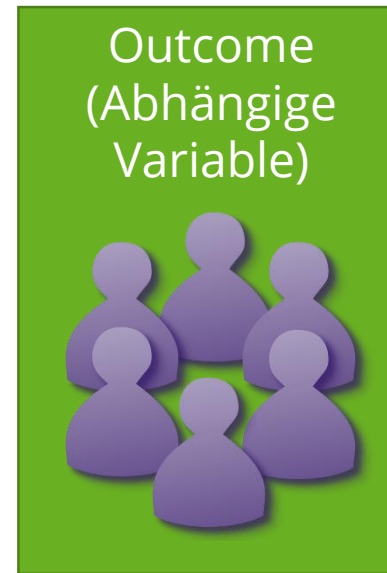
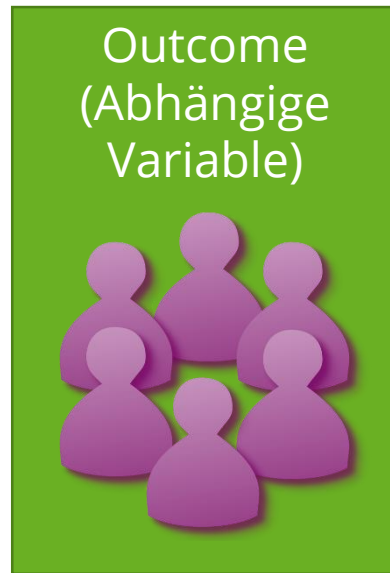
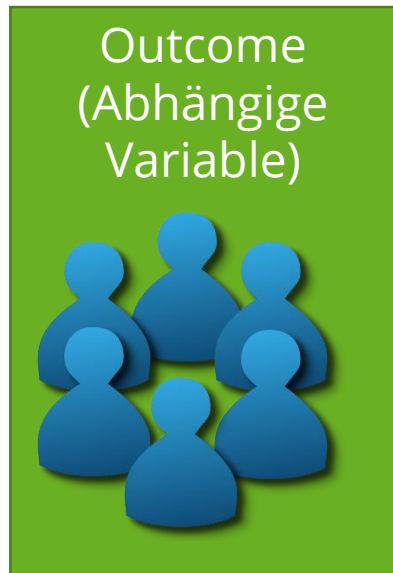


# Datenauswertung

## ANOVA

ANOVA testet **Mittelwertsunterschiede** in der Abhängigen Variable zwischen verschiedenen Gruppen → Vergleich von **mehr als zwei Gruppen** gleichzeitig

Diese Gruppen können **unabhängig** (verschiedene Personen in den Gruppen) oder **abhängig** (die selben Personen in den Gruppen) sein



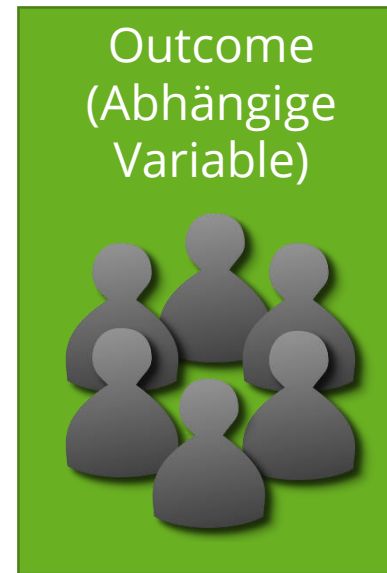
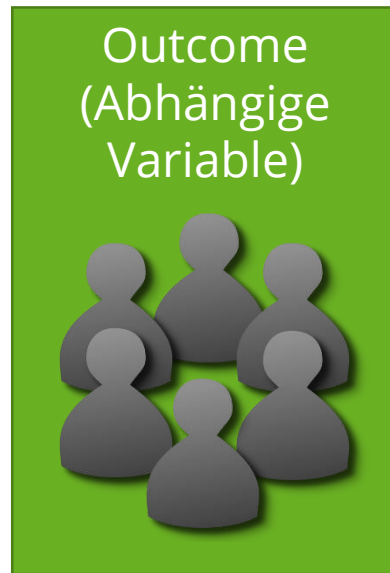
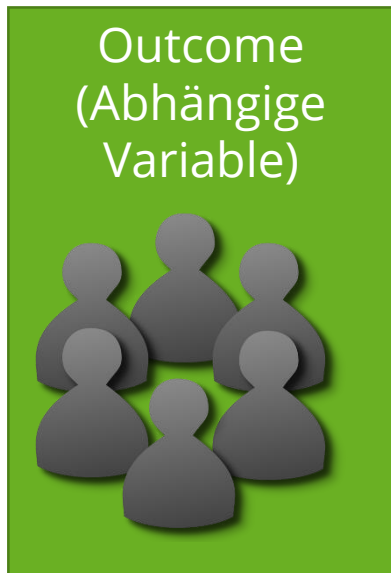
# Datenauswertung

## ANOVA

ANOVA testet **Mittelwertsunterschiede** in der Abhängigen Variable zwischen verschiedenen Gruppen → Vergleich von **mehr als zwei Gruppen** gleichzeitig

Diese Gruppen können **unabhängig** (verschiedene Personen in den Gruppen) oder **abhängig** (die selben Personen in den Gruppen) sein




**Abhängige Gruppen → ANOVA mit Messwiederholung (repeated measures ANOVA)**



# Datenauswertung

## ANOVA

### Zweifaktorielle ANOVA, 2x2 Design

	Stufe 1 der UV1	Stufe 2 der UV1
Stufe 1 der UV2		
Stufe 2 der UV2		

# Datenauswertung

## ANOVA



```
1 Datensatz %>%  
2   anova_test(  
3     AV ~ UV1 * UV2,  
4     type = 3,  
5     detailed = T,  
6     effect.size = "pes"  
7   )
```

```
1 Datensatz %>%  
2   tukey_hsd(AV ~ UV1 * UV2)
```

# Datenauswertung

## ANOVA



```
# compute ANOVA for arousal ratings
aov.arousal <- afex::aov_ez(id = "vp",      #ATTENTION: needs package "afex"
  dv = "Rating",
  data = Rating.Arousal.long,
  within = c("emot", "cond"),
  fun_aggregate = mean,
  include_aov = TRUE) |
afex::nice(aov.arousal, es = "pes")

# compute posthoc tests for both within measures
posthoc_test_strategy = emmeans::emmeans(aov.arousal$aov, specs = "cond")
pairs(posthoc_test_strategy, adjust = "bonferroni")
posthoc_test1 <- emmeans::emmeans(aov.arousal$aov, specs = "cond", by = "emot")
pairs(posthoc_test1, adjust = "bonferroni")
posthoc_test2 <- emmeans::emmeans(aov.arousal$aov, specs = "emot", by = "cond")
pairs(posthoc_test2, adjust = "bonferroni")
```

### Anova Table (Type 3 tests)

Response: Rating

	Effect	df	MSE	F	pes	p.value
1	emot	1, 109	7089.20	45.60 ***	.295	<.001
2	cond	1.85, 202.15	4377.72	26.93 ***	.198	<.001
3	emot:cond	1.98, 215.86	2396.19	3.83 *	.034	.024

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sphericity correction method: GG

# Datenauswertung

## ANOVA

### ABER:

- ANOVA liefert lediglich das Ergebnis, ob sich die Mittelwerte der betrachteten Gruppen signifikant voneinander unterscheiden
- Sie sagt nichts darüber aus, welche Gruppen sich voneinander unterscheiden
- Dafür benötigt es weitere Tests → **Post-Hoc Tests**
  
- (wird gelegentlich als Nachteil bezeichnet, für mich stellt das tatsächlich eher einen Vorteil dar)

# Datenauswertung

## R Script

### Empfehlungen:

- für eine einfachere Nutzung der Skripte durch Co-Autor:innen oder Dritte, empfiehlt es sich, am Anfange alle benötigten Packages zu laden
- Wenn Sie eine Funktion aus einem Package nutzen, schreiben Sie dieses dann direkt davor

```
# package names
packages <- c("tidyverse","afex","emmeans")

# install packages not yet installed

installed_packages <- packages %in% rownames(installed.packages())

if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])
}

# load packages

invisible(lapply(packages, library, character.only = TRUE))
```

```
library(tidyverse)
library(afex)
library(emmeans)
```

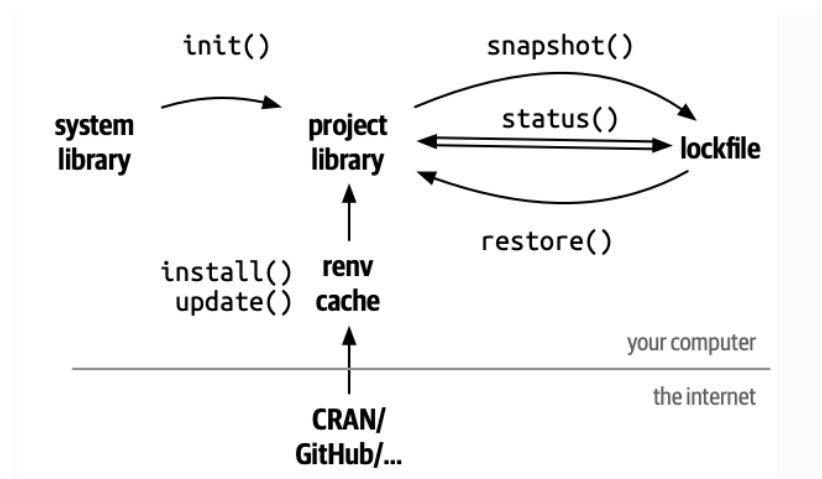
```
aov.arousal <- afex::aov_ez(id = "vp", #ATTENTION: needs package "afex"
```

# Datenauswertung

## R Script

### Exkurs:

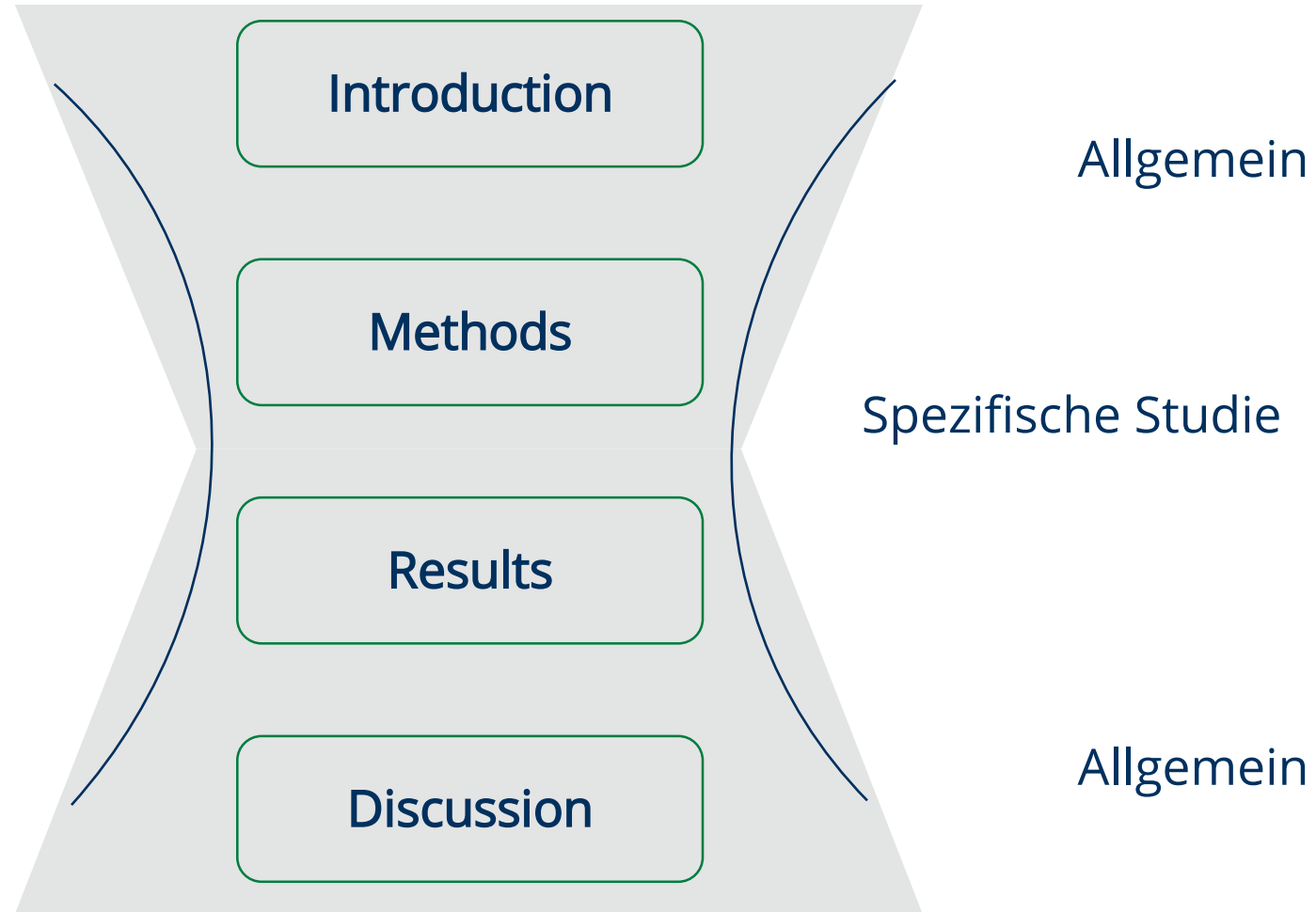
- Gelegentlich kann es zu leicht anderen Ergebnissen kommen, wenn man eine
- Um dabei computationale Reproduzierbarkeit zu gewährleisten, lohnt es sich reproduzierbare Umgebungen aufzubauen
- Bekannte Tools / Packages hierfür sind:
  - renv
  - groundhog
  - conda



# Wie schreibe ich einen Methodenteil?

# Methodenteil

Genereller Aufbau eines Artikels



# Methodenteil

## Warum schreibe ich einen Methodenteil?

- Kriterium für eine gute Studie ist deren Replizierbarkeit bzw. die Replizierbarkeit der Ergebnisse. Die Replizierbarkeit wird sehr gut im Methodenteil sichtbar
- Methodenteil ist eine möglichst genaue Beschreibung der Untersuchung – diese kann dazu dienen die verwendeten Methoden hinsichtlich der Gütekriterien zu bewerten, andererseits kann der Methodenteil als Anleitung verstanden werden, wenn ich die Studie wiederholen möchte

**RESEARCH ARTICLE**

**PSYCHOLOGY**

## **Estimating the reproducibility of psychological science**

**Open Science Collaboration\*†**

# Methodenteil

Warum schreibe ich einen Methodenteil?

Table of Definitions for Reproducibility

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

# Methodenteil

## Die technische Beschreibung der Studie

- **Participants:** Anzahl, Alter, Geschlecht, Rekrutierung, Entlohnung, Eigenschaften, Sample Size Rationale (→ Präregistrierung), Informed Consent, Declaration of Helsinki, Ethikvotum
- **Materials:** Fragebögen, Statements, etc...: Alle Materialien, die in der Studie eine Rolle spielten
- **Procedure:** Experimentalablauf, Software, Randomisierung → der komplette Studienablauf
  
- **Hinweis:** Gerade für Experimental- oder Trial Ablauf könnte sich eine Abbildung zur Veranschaulichung eignen, besonders bei komplizierteren Abläufen

# Methodenteil

## Die technische Beschreibung der Studie

- **Participants:** Anzahl, Alter, Geschlecht, Rekrutierung, Entlohnung, Eigenschaften, Sample Size Rationale (→ Präregistrierung), Informed Consent, Declaration of Helsinki, Ethikvotum
- **Materials:** Fragebögen, Statements, etc...: Alle Materialien, die in der Studie eine Rolle spielten
- **Procedure:** Experimentalablauf, Software, Randomisierung → der komplette Studienablauf
- **Measures (optional):** Abhängige Variablen und Transformation
- **Design (optional):** Art des Designs, das aus dem Procedure hervorgeht (→ evtl. schon unter der Überschrift Procedure abgehandelt)
- **Preprocessing (optional):** Transformationen, Ausschlusskriterien, Indexrechnungen, Software, etc.
- **Statistical Analyses (immer häufiger):**
  - Eingesetzte Analysen, alpha-Wert, Analyse der Vorannahmen, Korrekturverfahren (z.B. Greenhouse-Geisser für Spherizität bei ANOVA)
  - Am besten Unterteilung in: „Konfirmatorische Analysen“ und „Explorative Analysen“

# Ausblick

- Tauschen Sie sich in Ihren Kleingruppen aus:
  - Haben alle die Voraussetzungen geschaffen, um die Auswertung mit *R* anzugehen?
  - Wie werden Sie die Datenauswertung angehen? Wer fühlt sich dafür verantwortlich?
  - Wie dokumentieren wir am besten unsere Verarbeitungs- und Analyseschritte?
- Methodenteil:
  - Während der Auswertung bietet es sich an, die Schritte gut zu dokumentieren und bereits in Textform niederzuschreiben (spart Arbeit)
  - Wer fühlt sich für den Methodenteil verantwortlich?
  - Wie organisieren Sie sich die Zusammenarbeit am Dokument? (Cloud, Online-Dokument → wem Google nicht geheuer ist, dem kann ich CryptPad empfehlen <https://cryptpad.fr/>)

# Ausblick

nächster Termin: 24.10.23

„Schreibwerkstatt“ Auswertung

11:10 Uhr - BZW ASG14

nächster gemeinsamer Termin: 07.11.2023

Wie schreibe ich einen Ergebnisteil?

11:10 Uhr – BZW ASG14