



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (1) Einführung

Prof. Dr. Dirk Ostwald (dirk.ostwald@ovgu.de)



Seit 2021	W2 Professur Methodenlehre I
2014 - 2020	W1 Professur Freie Universität Berlin
2010 - 2014	Postdoc BCCN & MPIB Berlin
2007 - 2010	PhD Psychologie Birmingham
2004 - 2006	MSc Neurowissenschaften Tübingen
2005 - 2012	BSc Mathematik Hagen
2000 - 2003	BSc Medizin Hamburg

Forschung    Komputationale Kognitive Neurowissenschaften  
Lehre         Datenwissenschaft



OTTO VON GUERICKE  
UNIVERSITÄT  
MAGDEBURG

INSTITUT FÜR PSYCHOLOGIE

[Sitemap](#) [Impressum](#) [Kontakt](#)

Suchbegriff



INSTITUT | STUDIUM | FORSCHUNG | PERSONEN

DIREKTLINKS ▾

[Home](#) > [Institut](#) > [Abteilungen des I.](#) > [Methodenlehre I: Experimentelle und N.](#) > [Forschung](#) | [Lehre](#) | [CBBS Imaging Plattform](#) | [Team](#)

## Methodenlehre I: Experimentelle und Neurowissenschaftliche Psychologie



### Forschung



### Lehre



### CBBS Imaging Plattform



### Team



### Kontakt

#### Abteilungsleitung

> [Prof. Dr. Dirk Ostwald](#)

> [dirk.ostwald@ovgu.de](mailto:dirk.ostwald@ovgu.de)

Tel.: +49 391 67 57370

#### Abteilungsassistentz

> [Birgit Müller](#)

> [birgit.mueller@ovgu.de](mailto:birgit.mueller@ovgu.de)

Tel.: +49 391 67 58464

#### Anschrift

Otto-von-Guericke-Universität  
Magdeburg  
Institut für Psychologie  
Universitätsplatz 2  
Gebäude 24  
39106 Magdeburg

> [Anfahrt](#)

---

Formalia

Wissenschaft und Psychologie

Selbstkontrollfragen

---

## **Formalia**

Wissenschaft und Psychologie

Selbstkontrollfragen

## Psychologische Forschungsmethoden

- Klärung von Begrifflichkeiten psychologischer Forschung
- Vermittlung der psychologischen Fachsprache
- Vermittlung eines (unrealistischen und beschränkten) Wissenschaftsideals
- Keine Vermittlung konkret anwendbarer Forschungsmethodik (Design & Analyse)
- Keine umfassende Behandlung wissenschaftstheoretischer Hintergründe

⇒ Grundlegende Überarbeitung zum Sommersemester 2022

- **Modul “Einführung in die Forschungsmethoden der Psychologie”**
- Donnerstags 7 - 9 Uhr, G22A-H2
- Kursmaterialien (Folien, Videos) auf der **Kurswebseite**
- Ankündigungen über die **Moodleseite**
- Benotete Multiple Choice Klausur Ende Wintersemester 2021/22
- Klausurwiederholungstermin am Ende des Sommersemesters 2022
- Klausurtermin und Klausurort gemäß Prüfungsplan des **FNW Prüfungsamtes**

## Vorlesungsübersicht

Datum	Einheit	Thema
21.10.2021	Einführung	(1) Einführung
28.10.2021	Einführung	(2) Theorien, Hypothesen, Experimente
04.11.2021	Einführung	(3) Operationalisieren, Messen, Skalieren
11.11.2021	Studiendesign	(4) Prinzipien der Versuchsplanung
18.11.2021	Studiendesign	(5) Kriterien der Versuchsplanung
25.11.2021	Studiendesign	(6) Experimentelle Versuchspläne I
02.12.2021	Studiendesign	(7) Experimentelle Versuchspläne II
09.12.2021	Studiendesign	(8) Experimentelle Versuchspläne III
16.12.2021	Studiendesign	(9) Quasiexperimente und Korrelationsstudien
	Weihnachtspause	
06.01.2022	Auswahlverfahren	(10) Grundgesamtheit und Stichprobe
13.01.2022	Auswahlverfahren	(11) Stichprobentechniken
20.01.2022	Datenerhebung	(12) Qualitative Verfahren
27.01.2022	Datenerhebung	(13) Quantitative Verfahren
Feb 2022	Klausurtermin	
Jul 2022	Klausurwiederholungstermin	

- Vorlesungsfolien inklusive Selbstkontrollfragen sind klausurrelevant
- Vorlesungsfolien basierend auf früheren Vorlesungen durch
  - Dr. Melanie Baumgarten
  - Dr. Martin Krippel
- Zugrundeliegender Text
  - Reiß, S. & Sarris, V. (2012) Experimentelle Psychologie

## Weitere Kurse der Abteilung

- Vorkurs Grundlagen der Mathematik und Informatik
- Wahrscheinlichkeitstheorie und Frequentistische Inferenz
- Allgemeines Lineares Modell und Bayesianische Inferenz
- Multivariate Datenanalyse
- Programmierung und Deskriptive Statistik in R

---

Formalia

## **Wissenschaft und Psychologie**

Selbstkontrollfragen

## Wissenschaft

Wissenschaft bezeichnet den methodischen Prozess intersubjektiv nachvollziehbaren Forschens und Erkennens in einem bestimmten Bereich, der nach herkömmlichem Verständnis ein begründetes, geordnetes und gesichertes Wissen hervorbringt.

Wissenschaft kennzeichnet entsprechend das gesicherte und in einen rationalen Begründungszusammenhang gestellte Wissen, welches kommunizierbar und überprüfbar ist sowie bestimmten wissenschaftlichen Kriterien folgt.

Wissenschaft bezeichnet somit ein zusammenhängendes System von Aussagen, Theorien und Verfahrensweisen, das strengen Prüfungen der Geltung unterzogen wurde und mit dem Anspruch objektiver, überpersönlicher Gültigkeit verbunden ist.

*Carrier, M. (2011) Lexikon der Philosophie, Reclam, Stuttgart, S. 312*

## Naturwissenschaften | Science

- Empirische Erforschung der Natur mit dem Ziel, Regelmäßigkeiten zu erkennen
- Quantitatives Beobachten, messen, analysieren des Verhaltens der Natur
- Grundlage zur Nutzbarmachung der Natur in den Ingenieursdisziplinen
- Physik, Chemie, Biologie, Medizin, **Psychologie**, Geologie, etc.

## Geisteswissenschaften | Humanities

- Analytische Erforschung menschlicher Kultur
- Qualitative Sinnsuche, informelle Kritik, Spekulation
- Philosophie, Theologie, Geschichtswissenschaft, Literaturwissenschaft, etc.
- Naturwissenschaftliche Theoriebildung

## Formalwissenschaften | Formal Sciences

- Analyse formaler Systeme
- Sprachwerkzeuge
- Mathematik, Logik, theoretische Informatik, Rechtswissenschaft, etc.
- Naturwissenschaftliche Theoriebildung

## Prinzipien der Erkenntnisgewinnung

### Prinzip der Intuition

- Unmittelbare Eingebung
- Ökonomisch, aber risikobehaftet

### Prinzip der Autorität

- Übernahme von Erkenntnissen von Autoritäten (Expert:innen)
- Ökonomisch, aber risikobehaftet

### Prinzip der Vernunft

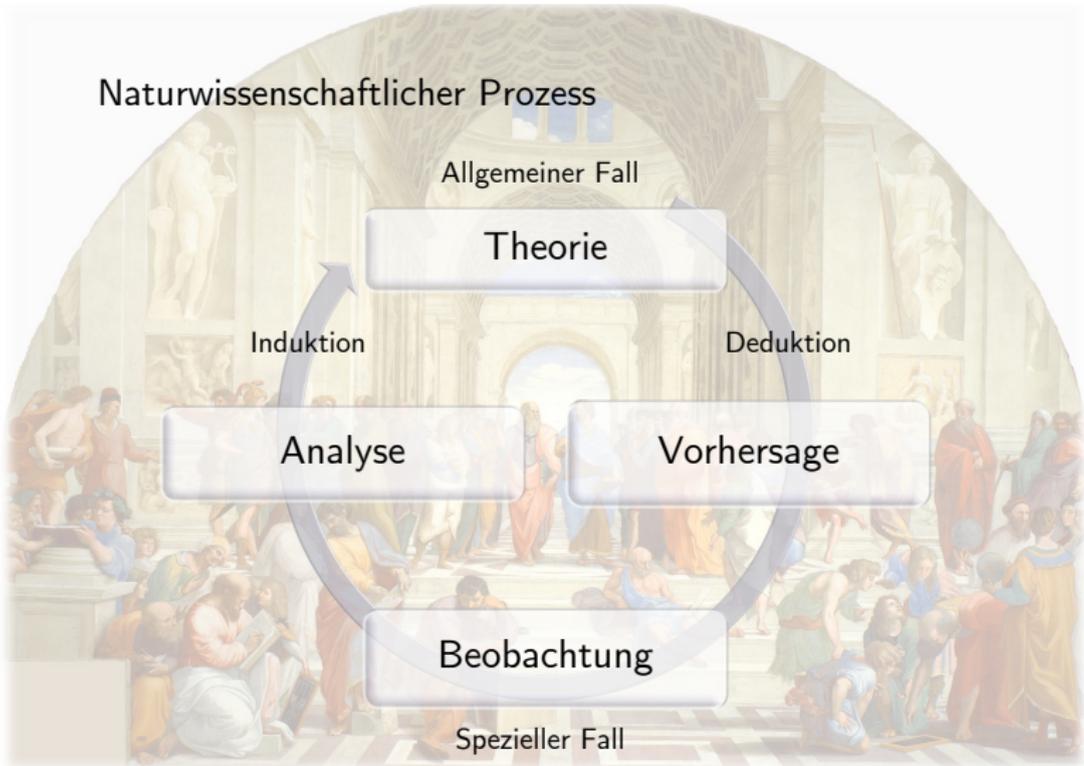
- Erkenntnisgewinn in der Theorie nach formalen Regeln (Logik und Mathematik)
- Intersubjektiv, aber modellbasiert

### Prinzip der Erfahrung

- Beobachtung und Experiment
- Intersubjektiv, aber *per se* unstrukturiert

*“Theorie ohne Erfahrung ist lediglich intellektuelles Spiel, Erfahrung ohne Theorie ist blind.”*

*nach Immanuel Kant (vielleicht)*

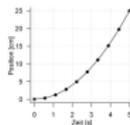




Isaac Newton  
1689

$$\ddot{x}(t) = \frac{F}{m}, F = -mg$$

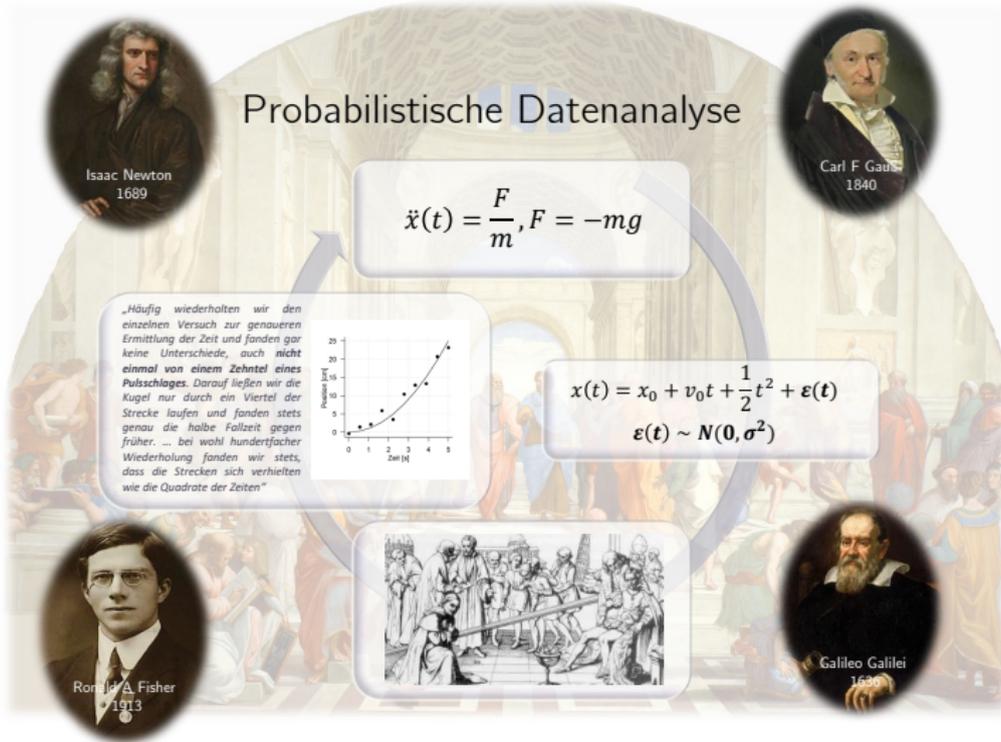
„Häufig wiederholten wir den einzelnen Versuch zur genaueren Ermittlung der Zeit und fanden gar keine Unterschiede, auch nicht einmal von einem Zehntel eines Pulsschlages. Darauf ließen wir die Kugel nur durch ein Viertel der Strecke laufen und fanden stets genau die halbe Fallzeit gegen früher. ... bei wohl hundertfacher Wiederholung fanden wir stets, dass die Strecken sich verhielten wie die Quadrate der Zeiten“



$$x(t) = x_0 + v_0 t + \frac{1}{2} t^2$$



Galileo Galilei  
1639



## Psychologie

Wissenschaft des menschlichen Erlebens, Verhaltens und Handelns

Beschreiben

- Benennen und Klassifizieren neuropsychologischer Phänomene

Erklären

- Entwicklung mechanistischer neuropsychologischer Modelle

Vorhersagen

- Prognose zukünftigen Erlebens, Verhaltens und Handelns

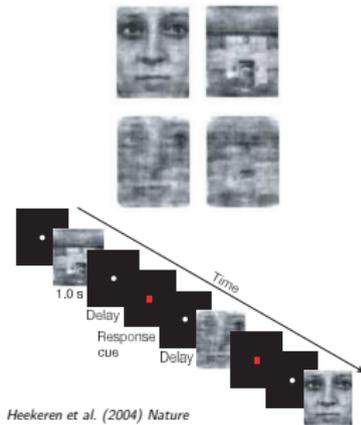
Verändern

- Prävention, Diagnose, Behandlung psychiatrischer Erkrankungen

## Neurobiologische Verarbeitung von Sinnesreizen

Wie werden visuelle Stimuli im Gehirn verarbeitet?

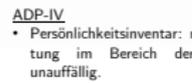
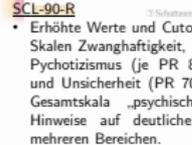
Wie entscheiden Menschen, ob sie ein Haus oder ein Gesicht wahrnehmen?



→ Allgemeine Psychologie, Biologische Psychologie, Kognitive Neurowissenschaften

## Psychologische Diagnostik im psychotherapeutischen Erstgespräch

Ist eine Psychotherapie indiziert?



- Summe 47, Hinweis auf schwere depressive Episode.

- Erhöhte Werte und Cutoff-Überschreitungen auf den Skalen Zwanghaftigkeit, Depressivität, Ängstlichkeit, Psychotizismus (je PR 80), Aggressivität (PR 73) und Unsicherheit (PR 70). Deutliche Belastung auf Gesamtskala „psychische Belastung“ (PR 80). Hinweise auf deutliche psychische Belastung in mehreren Bereichen.

- Persönlichkeitsinventar: minimale Cutoff-Überschreitung im Bereich der Paranoiden PS, sonst unauffällig.

→ Diagnostik, Persönlichkeitspsychologie, Verfahrenlehre, MSc Psychotherapie

## Evidenzbasierte Evaluation von Psychotherapieformen bei Depression

Welche Therapieform ist bei Depression wirksamer?

### Online Psychotherapie



### Klassische Psychotherapie



→ Klinische Psychologie, Klinische Diagnostik, MSc Psychotherapie

# Wissenschaft und Psychologie

## Evidenzbasierte Evaluation von Psychotherapieformen bei Depression

### Becks Depressions-Inventar (BDI) zur Depressionsdiagnostik

BDI-II Fragebogen			
Name	Alter	Geschlecht	Datum
<p><b>Anleitung:</b> Dieser Fragebogen enthält 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Gruppen von Aussagen sorgfältig durch und wählen Sie sich dann in jeder Gruppe eine Aussage heraus, die am besten beschreibt, wie Sie sich in der letzten zwei Wochen, einschließlich heute, gefühlt haben. Konzentrieren Sie die Aufmerksamkeit auf die Aussage, die Sie sich am häufigsten haben (0, 1, 2 oder 3). Falls in einer Gruppe mehrere Aussagen gleichwertig sind für zutreffend, können Sie für Aussage mit der höchsten Zahl ein „Achten Sie bitte darauf, dass Sie in jeder Gruppe nicht mehr als eine Aussage ankreuzen, die gut zu Ihnen passt.“ (Veränderungen der Schlafgewohnheiten) oder Gruppe 16 (Veränderungen des Appetits).</p>			
<p><b>1.) Traurigkeit</b></p> <p>0 Ich bin nicht traurig. 1 Ich bin oft traurig. 2 Ich bin ständig traurig. 3 Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.</p> <p><b>2.) pessimismus</b></p> <p>0 Ich sehe nicht mal in die Zukunft. 1 Ich sehe mühsamer in die Zukunft als sonst. 2 Ich bin müde und erwarte nicht, dass meine Situation besser wird. 3 Ich glaube, dass meine Zukunft hoffnungslos ist und nur noch schlechter wird.</p> <p><b>3.) Versagensgefühle</b></p> <p>0 Ich fühle mich nicht als Versager. 1 Ich habe häufiger Versagensgefühle. 2 Wenn ich zurückblicke, sehe ich eine Menge Fehlertage. 3 Ich habe das Gefühl, ich mache ein völliger Versager zu sein.</p> <p><b>4.) Verlust von Freude</b></p> <p>0 Ich kann die Dinge genauso gut genießen wie früher. 1 Ich kann die Dinge nicht mehr so genießen wie früher. 2 Dinge, die mir früher Freude gemacht haben, kann ich kaum mehr genießen. 3 Dinge, die mir früher Freude gemacht haben, kann ich überhaupt nicht mehr genießen.</p> <p><b>5.) Schuldgefühle</b></p> <p>0 Ich habe keine besonderen Schuldgefühle. 1 Ich habe oft Schuldgefühle wegen Dingen, die ich getan habe oder hätte tun sollen. 2 Ich habe die meiste Zeit Schuldgefühle. 3 Ich habe ständig Schuldgefühle.</p>	<p><b>6.) Bestrafungsgefühle</b></p> <p>0 Ich habe keine das Gefühl, für etwas bestraft zu sein. 1 Ich habe das Gefühl, vielleicht bestraft zu werden. 2 Ich erwarte, bestraft zu werden. 3 Ich habe das Gefühl, bestraft zu sein.</p> <p><b>7.) Selbsthöhnung</b></p> <p>0 Ich habe von mir genauso viel wie immer. 1 Ich habe Vertrauen in mich verloren. 2 Ich bin von mir enttäuscht. 3 Ich lehne mich völlig ab.</p> <p><b>8.) Selbstvorwürfe</b></p> <p>0 Ich kritisiere oder tadle mich nicht mehr als sonst. 1 Ich bin mir gegenüber kritischer als sonst. 2 Ich kritisiere mich für all meine Mängel. 3 Ich gebe mir die Schuld für alles Schlechte, was passiert.</p> <p><b>9.) Selbstmordgedanken</b></p> <p>0 Ich denke nicht daran, mir etwas anzutun. 1 Ich denke manchmal an Selbstmord, aber ich würde es nicht tun. 2 Ich möchte mich ein bisschen verletzen. 3 Ich würde mich umbringen, wenn ich die Gelegenheit dazu hätte.</p> <p><b>10.) Weinen</b></p> <p>0 Ich weine nicht öfter als früher. 1 Ich weine jetzt mehr als früher. 2 Ich weine beim geringsten Anlass. 3 Ich möchte gern weinen, aber ich kann nicht.</p>		

<p><b>11.) Unruhe</b></p> <p>0 Ich bin nicht unruhiger als sonst. 1 Ich bin unruhiger als sonst. 2 Ich bin so unruhig, dass es mir schwerfällt, still zu sitzen. 3 Ich bin so unruhig, dass ich mich ständig bewegen oder etwas tun muss.</p> <p><b>12.) Interessensverlust</b></p> <p>0 Ich habe das Interesse an anderen Menschen oder an Tätigkeiten nicht verloren. 1 Ich habe weniger Interesse an anderen Menschen oder an Dingen als sonst. 2 Ich habe das Interesse an anderen Menschen oder Dingen zum größten Teil verloren. 3 Es fällt mir schwer, mich überhaupt für irgend etwas zu interessieren.</p> <p><b>13.) Entschlussunfähigkeit</b></p> <p>0 Ich bin so entscheidungslos wie immer. 1 Es fällt mir schwerer als sonst, Entscheidungen zu treffen. 2 Es fällt mir sehr viel schwerer als sonst, Entscheidungen zu treffen. 3 Ich habe Mühe, überhaupt Entscheidungen zu treffen.</p> <p><b>14.) Wertlosigkeit</b></p> <p>0 Ich fühle mich nicht wertlos. 1 Ich fühle mich für weniger wertvoll und nützlich als sonst. 2 Vergleichen mit anderen Menschen fühle ich mich viel weniger wert. 3 Ich fühle mich völlig wertlos.</p> <p><b>15.) Energieverlust</b></p> <p>0 Ich habe so viel Energie wie immer. 1 Ich habe weniger Energie als sonst. 2 Ich habe so wenig Energie, dass ich kaum noch etwas schaffen. 3 Ich habe keine Energie mehr, um überhaupt noch etwas zu tun.</p> <p><b>16.) Veränderungen der Schlafgewohnheiten</b></p> <p>0 Meine Schlafgewohnheiten haben sich nicht verändert. 1a Ich schlafe etwas mehr als sonst. 1b Ich schlafe etwas weniger als sonst. 2a Ich schlafe viel mehr als sonst. 2b Ich schlafe viel weniger als sonst. 3a Ich schlafe fast den ganzen Tag. 3b Ich wache 1-2 Stunden früher auf als gewöhnlich und kann dann nicht mehr einschlafen.</p>	<p><b>17.) Reizbarkeit</b></p> <p>0 Ich bin nicht reizbarer als sonst. 1 Ich bin reizbarer als sonst. 2 Ich bin viel reizbarer als sonst. 3 Ich fühle mich dauernd gereizt.</p> <p><b>18.) Veränderungen des Appetits</b></p> <p>0 Mein Appetit hat sich nicht verändert. 1a Mein Appetit ist etwas schlechter als sonst. 1b Mein Appetit ist etwas größer als sonst. 2a Mein Appetit ist viel schlechter als sonst. 2b Mein Appetit ist viel größer als sonst. 3a Ich habe überhaupt keinen Appetit. 3b Ich habe ständig Heißhunger.</p> <p><b>19.) Konzentrationschwierigkeiten</b></p> <p>0 Ich kann mich so gut konzentrieren wie immer. 1 Ich kann mich nicht mehr so gut konzentrieren wie sonst. 2 Es fällt mir schwer, mich längere Zeit auf irgend etwas zu konzentrieren. 3 Ich kann mich überhaupt nicht mehr konzentrieren.</p> <p><b>20.) Ermüdung oder Erschöpfung</b></p> <p>0 Ich fühle mich nicht müde oder erschöpft als sonst. 1 Ich werde schneller müde oder erschöpft als sonst. 2 Für viele Dinge, die ich üblicherweise tue, bin ich zu müde oder erschöpft. 3 Ich bin so müde oder erschöpft, dass ich fast nichts mehr tun kann.</p> <p><b>21.) Verlust an sexuellem Interesse</b></p> <p>0 Mein Interesse an Sexualität hat sich in letzter Zeit verändert. 1 Ich interessiere mich weniger für Sexualität als früher. 2 Ich interessiere mich jetzt viel weniger für Sexualität. 3 Ich habe das Interesse an Sexualität völlig verloren.</p>
---	--

0 - 8 keine Depression

9 - 13 minimale Depression

14 - 19 leichte Depression

20 - 28 mittelschwere Depression

29 - 63 schwere Depression

## Qualitativer und quantitativer Ansatz

<b>Methodenklasse</b>	<b>Qualitativer Ansatz</b>	<b>Quantitativer Ansatz</b>
Forschungsansatz	Deskriptive Feldforschung	Laborexperimente
	Biografische Methoden	Quasiexperimente
	Handlungsforschung	Korrelationsstudien
Erhebungsmethoden	Interviews	Verhaltensmessungen
	Teilnehmende Beobachtung	Physiologische Messungen
Analysemethoden	Inhaltsanalysen	Deskriptive Statistiken
	Hermeneutik	Probabilistische Modellierung

## Psychologischer Methodenbegriff

Mittel und Wege, die dem Erkenntnisgewinn dienen

- Theoretischer Ausgangspunkt
- Fragestellung
- Hypothesenbildung
- Studiendesign
- Datenerhebung
- Datenauswertung
- Präsentation

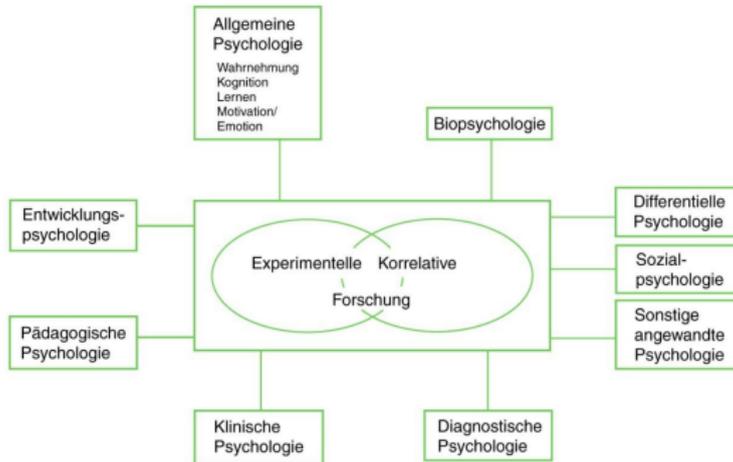
Die traditionelle deutschen Psychologie arbeitet pseudo-quantitativ

- Box-Arrows-Modelle
- Methoden nicht verstehen, SPSS durckklicken, SPSS Output “interpretieren”
- Oft: Methoden = Statistik

Die moderne psychologische Datenwissenschaft arbeitet quantitativ

- Mathematische Formulierung probabilistischer kognitiv-neurophysiologischer Modelle
- Intuitives und formal-mathematisches Verständnis
- Selbständige Software Implementation  $\Leftrightarrow$  Objektorientiere Programmierung

## Experimentell-korrelatives Grundmodell der wissenschaftlichen Psychologie



**Abbildung 1.1:** Das experimentell-korrelative Grundmodell. In der Psychologie muss vielfach die experimentelle Untersuchungsmethodologie durch den korrelativen Forschungsansatz ergänzt werden. (Modifiziert nach Sarris, 1995, 1999)

aus Reiß, S. & Sarris, V. (2012)

---

Formalia

Wissenschaft und Psychologie

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Diskutieren Sie die Begriffe Naturwissenschaft, Geisteswissenschaft und Formalwissenschaft.
2. Beschreiben Sie den naturwissenschaftlichen Prozess.
3. Nennen Sie vier Ziele der wissenschaftlichen Psychologie.
4. Beschreiben Sie Beispiele für die Anwendung wissenschaftlicher Methoden in der Psychologie.
5. Diskutieren Sie qualitative und quantitative Forschungsansätze in der Psychologie.



# Psychologische Forschungsmethoden

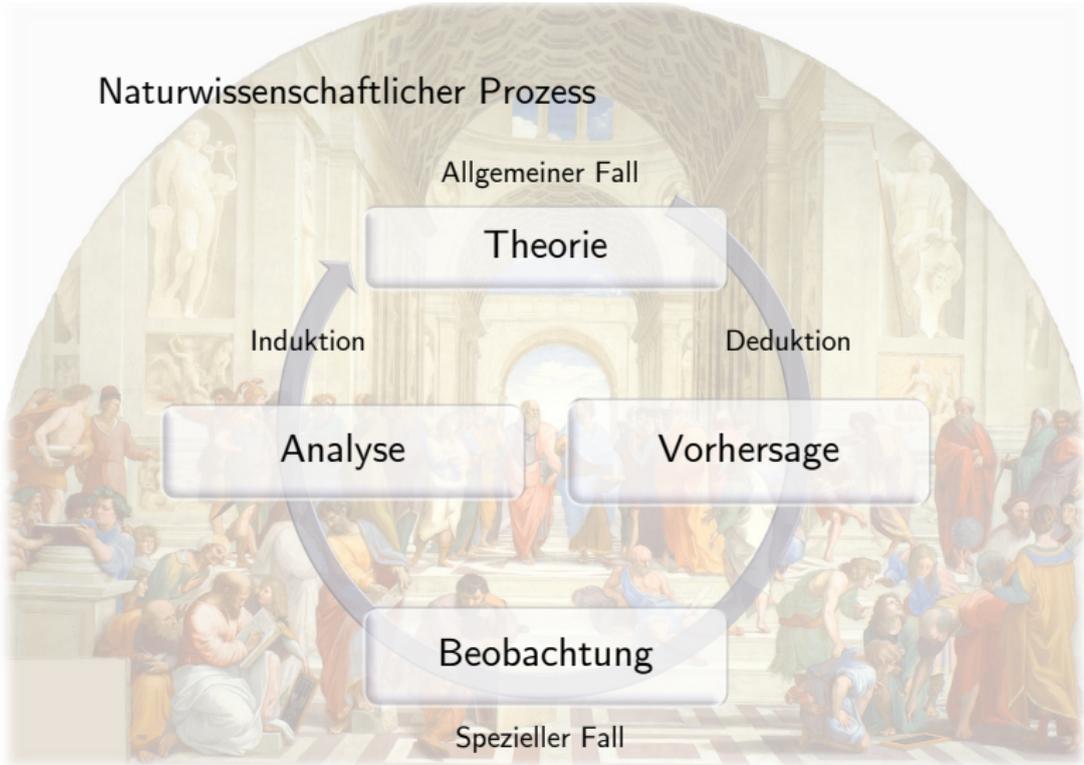
BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

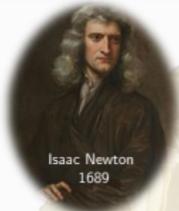
Prof. Dr. Dirk Ostwald

(2) Theorien, Hypothesen, Experimente

## Naturwissenschaftlicher Prozess



# Probabilistische Datenanalyse



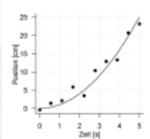
Isaac Newton  
1689



Carl F. Gauß  
1840

$$\ddot{x}(t) = \frac{F}{m}, F = -mg$$

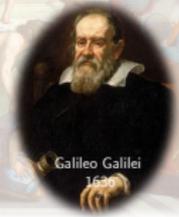
„Häufig wiederholen wir den einzelnen Versuch zur genaueren Ermittlung der Zeit und fanden gar keine Unterschiede, auch nicht einmal von einem Zehntel eines Pulschlags. Darauf ließen wir die Kugel nur durch ein Viertel der Strecke laufen und fanden stets genau die halbe Fallzeit gegen früher. ... bei wohl hundertfacher Wiederholung fanden wir stets, dass die Strecken sich verhielten wie die Quadrate der Zeiten“



$$x(t) = x_0 + v_0 t + \frac{1}{2} t^2 + \varepsilon(t)$$
$$\varepsilon(t) \sim N(0, \sigma^2)$$



Ronald A. Fisher  
1913



Galileo Galilei  
1632

---

Theorien

Hypothesen

Experimente

Wissenschaftstheorie

Selbstkontrollfragen

---

## **Theorien**

Hypothesen

Experimente

Beispiel

Wissenschaftstheorie

Selbstkontrollfragen

## Definitionsversuche

“Eine Theorie ist ein Erklärungsmodell, das auf bestimmten Prinzipien basiert. Eine Theorie stellt Einzelbeobachtungen in einen Zusammenhang. Mithilfe einer Theorie können Verhaltensweisen oder Ereignisse in ein System gebracht und Vorhersagen abgeleitet werden.”

Myers (2010)

“Eine Theorie ist eine geordnete Menge von Konzepten oder Aussagen, die ein Phänomen oder eine Gruppe von Phänomenen erklärt.”

Gerrig et al. (2020)

“Eine Theorie ist ein System von Definitionen, Annahmen und Schlussfolgerungen, welches einen Ordnungs- und Erklärungsversuch für ein oder mehrere Phänomene darstellt.”

Reiß and Sarris (2012)

## Theorien und Modelle

Psycholog:innen unterscheiden manchmal zwischen den Begriffen “Theorie” und “Modell”

- “Nicht selten sind Modelle Bestandteile von Theorien”
- “Die allgemeinen theoretischen Prinzipien sind dann durch ein Modell repräsentiert”
- “Man unterscheidet Mathematische Modelle, Computermodelle, . . .”

(vgl. Reiß and Sarris (2012), S. 34 - 37)

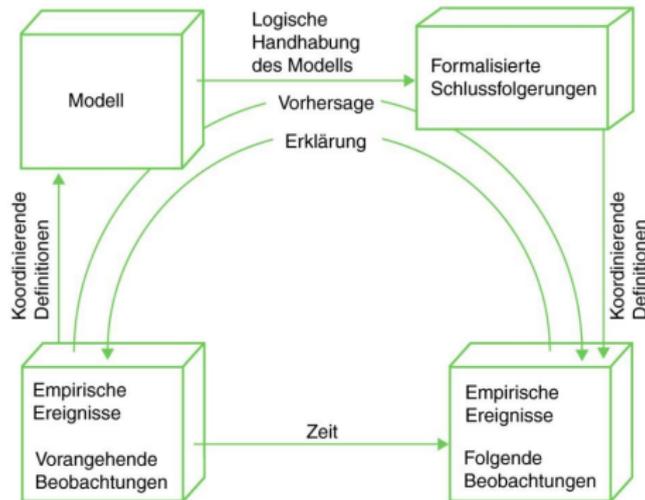
Psycholog:innen denken wohl, dass es Theorien auch ohne Modelle gibt (oder so).

Psycholog:innen haben meist keine Erfahrung mit mathematischer oder informatischer Modellbildung.

⇒ Eine Unterscheidung zwischen Theorie und Modell ist nicht zielführend.

⇒ Mathematische und informatische Modellbildung ist für naturwissenschaftliche Arbeit notwendig.

**Modell = Theorie und Theorie = Modell**



**Abbildung 2.2: Allgemeines Strukturschema einer Theorie.** Gemäß einer in der empirischen Einzelwissenschaft häufig vertretenen Auffassung wird ein „Modell“ – im Sinne einer Analogie zum „abzubildenden“ Phänomen – als ein Bestandteil einer „Theorie“ begriffen, mit dessen (meist anschaulicher) Hilfe zumindest einige ihrer Aspekte bzw. Teile in eine möglichst enge Verbindung zu der empirischen Datenwelt gebracht werden. – (Als eine Art „Flussdiagramm“ veranschaulicht das Strukturschema ferner den stark idealisierten zeitlichen Ablauf einer allgemeinen Modell- und Theorienbildung im Zusammenhang mit den komplementären wissenschaftstheoretischen Konzepten der „Vorhersage“ [prospektiv] und der „Erklärung“ [retrospektiv] von Beobachtungsdaten). – (Modifiziert nach Matheson, Bruce & Beauchamp, 1978; aus Sarris, 1999)

## Theoriegütekriterien nach Groeben and Westmeyer (1975)

Explizitheit	Intersubjektive Übereinstimmung
Widerspruchsfreiheit	Keine Widersprüche in abgeleitete Aussagen
Sparsamkeit	Möglichst wenig komplex
Vollständigkeit	Erklärung aller Phänomene des Gegenstandsbereichs
Prüfbarkeit	Empirische Überprüfung möglich
Empirische Verankerung	Vorhersage von beobachtbaren Daten
Produktivität	Generation neuer wissenschaftlicher Fragestellungen
Anwendbarkeit	Nutzen in der Praxis

Groeben and Westmeyer (1975)

## Definitionsversuch

“Eine Theorie/ein Modell ist ein System von intuitiv-verbal und mathematisch formulierten Definitionen und Theoremen, die elementare Grundannahmen zu Erklärung eines Phänomens der Wirklichkeit darstellen, das die Herleitung quantitativer, informatisch-implementierbarer Vorhersagen über beobachtbare Daten des Phänomens ermöglicht.”

Ostwald, 2021

---

Theorien

**Hypothesen**

Experimente

Beispiel

Wissenschaftstheorie

Selbstkontrollfragen

## Definitionsversuche

“Eine Hypothese ist eine überprüfbare Vorhersage, die aus einer Theorie abgeleitet wird.”

Myers (2010)

“Eine Hypothese ist eine vorläufige und prüfbare Erklärung der Beziehung zwischen zwei oder mehreren Ereignissen oder Variablen; oft als Vorhersage formuliert, dass bestimmte Ereignisse aufgrund spezifischer Bedingungen eintreten werden.

Gerrig et al. (2020)

“Eine Hypothese ist eine experimentell zu prüfende Tatsachenbehauptung bzw. präzise Angabe über die Art der erwarteten Abhängigkeitsbeziehung. Sie enthält die exakte Festlegung der variierten Bedingungen und der erwarteten Veränderungen, d.h. eine möglichst präzise Aussage (Vorhersage) über die empirische Beziehung zwischen Ereignissen”

Reiß and Sarris (2012)

## Hypothese und Statistische Hypothese

Psycholog:innen verwechseln oft die Begriffe “Hypothese” und “Statistische Hypothese.”

⇒ Der Begriff “Hypothese” und der Begriff der “Statistischen Hypothese” sind nicht gleich.

- “Hypothese” meint hier zunächst etwas wie “aus der Theorie abgeleitete Datenvorhersage.”
- Eine “Statistische Hypothese” (Nullhypothese, Alternativhypothese) ist eine Aussage über die Lage des wahren, aber unbekanntem, Parameters im Parameterraum eines statistischen Modells.

Hypothesen können manchmal als Statistische Hypothesen formuliert werden.

⇒ Hypothesen müssen nicht als Statistische Hypothesen formuliert werden.

⇒ Es gibt mehr datenanalytische Verfahren als Frequentistisches Hypothesentesten.

Die Annahme oder Ablehnung von Statistischen Hypothesen sind Quantifizierungen von Unsicherheit, keine abschließenden binären Urteile. Keine Hypothese wird sich jemals als absolut “falsch” oder absolut “richtig” erweisen.



## Probabilistische Datenvorhersage statt Hypothese

Der von Psycholog:innen propagierte Hypothesenfetisch ist problematisch:

- Verwechslung von Hypothesen und Statistischen Hypothesen
- Abwertung wichtiger naturwissenschaftlicher Beiträge wie zum Beispiel
  - (Mathematische) Modellentwicklung und (analytische/simulierende) Modellvalidation
  - Datenerhebung, Datenaufbereitung, Datenbereitstellung
  - Exploratorisch-charakterisierende Forschungsprojekte

Der Hypothesenfetisch hat die Qualität (psychologischer) Forschung gemindert:

- p-Hacking: Datenselektion bis zum Ablehnen der (statistischen) Nullhypothese
- HARKING: Hypothesizing After Results are Known (Pseudohypothesen)

Entscheidend für das naturwissenschaftliche Paradigma ist es, dass aus Theorien quantitative Vorhersagen beobachtbarer Daten abgeleitet werden, deren explanatorische Unsicherheit probabilistisch quantifiziert werden kann, und die Theorievergleiche ermöglicht.

⇒ An die Stelle des “Hypothesentestens” sollte das “Theorievergleichen” treten.

---

Theorien

Hypothesen

**Experimente**

Beispiel

Wissenschaftstheorie

Selbstkontrollfragen

## Definitionsversuche

“Ein Experiment ist eine Untersuchung, bei welcher Untersuchungsleitende aktiv und gezielt eine Intervention durchführen, um die Effekte der Intervention zu beobachten”

Shadish, Cook, and Campbell (2001)

“Ein Experiment ist eine empirische Untersuchung, bei der gezielt bestimmte Bedingungen (Stufen der *unabhängigen Variable*) hergestellt werden und in ihren Auswirkungen auf ausgewählte *abhängige Variablen* beobachtet werden.”

Bortz and Döring (2006)

“Ein Experiment ist ein systematischer Beobachtungsvorgang, aufgrund dessen der Versuchsleiter (sic) das jeweils interessierende Phänomen planmäßig erzeugt und variiert sowie gleichzeitig systematische und/oder unsystematische Störfaktoren durch hierfür geeignete Techniken kontrolliert”

Reiß and Sarris (2012)

## Definitionsversuche



Wilhelm Wundt (1832 - 1920)

## Experimentkriterien nach Wundt

- Willkürlichkeit
- Variierbarkeit
- Wiederholbarkeit

## Experimente vs. Quasiexperimente vs. Korrelationsstudien

### Experiment

- Randomisierte kontrollierte Studie
- Die Untersuchungseinheiten werden den Versuchsbedingungen zufällig zugeordnet
- Beispiel: Online Psychotherapie vs. Klassische Psychotherapie bei Depression

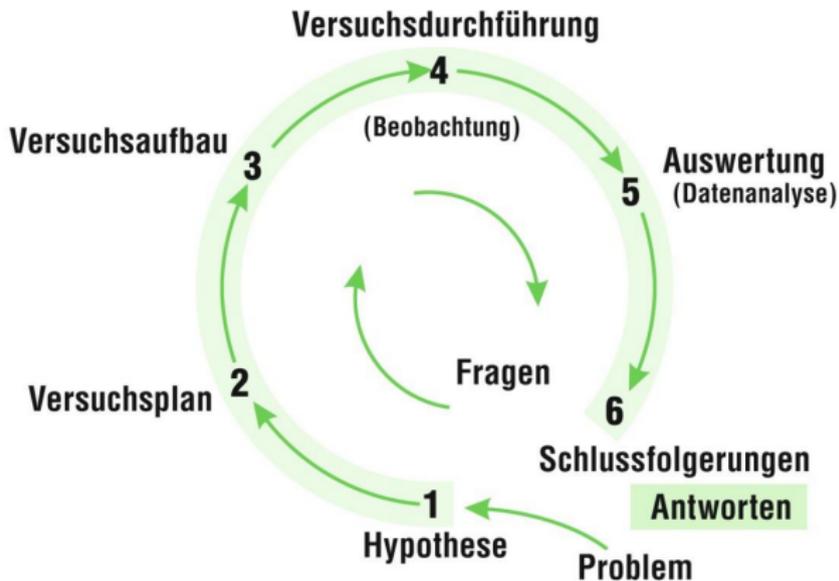
### Quasiexperiment

- Nicht-randomisierte kontrollierte Studie
- Untersuchung natürlich bzw. bereits bestehender Gruppen
- Beispiel: Online Psychotherapie bei Depression vs. Schizophrenie

### Korrelationsstudie

- Nicht-randomisierte, nicht kontrollierte Studie
- Beobachtungsstudie ohne Intervention
- Beispiel: Analyse von Paneldaten

## Stadien einer experimental-psychologischen Untersuchung



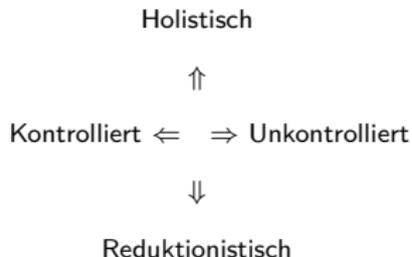
Reiß and Sarris (2012)

## Datenerhebung statt Experiment

Der von (Experimental)Psycholog:innen propagierte Experimentenfetisch ist problematisch:

- Abwertung anderer wichtiger naturwissenschaftlicher Beiträge
- Fokus auf experimentelles Design anstatt integrierter Betrachtung von Design und Analyse
- Es gibt keinen prinzipiellen Unterschied zwischen “Hypothesentests” und “Korrelation”
- Experimentaldatenbesitzgier verhindert wissenschaftlichen Fortschritt

Datenerhebungen finden (mindestens) in einem zweidimensionalen Erhebungsraum statt



Entscheidend ist die kritische Evaluation der Lage einer Datenerhebung in diesem Kontinuum

Niemand sagt mehr “Experiment,“ jeder sagt “Studie“ heutzutage.

---

Theorien

Hypothesen

Experimente

**Beispiel**

Wissenschaftstheorie

Selbstkontrollfragen



## Human Belief State-Based Exploration and Exploitation in an Information-Selective Symmetric Reversal Bandit Task

Lilla Horvath<sup>1</sup> · Stanley Colcombe<sup>2</sup> · Michael Milham<sup>2</sup> · Shruti Ray<sup>3</sup> · Philipp Schwartenbeck<sup>4</sup> · Dirk Ostwald<sup>5,6</sup>

Accepted: 24 May 2021  
© The Author(s) 2021

### Abstract

Humans often face sequential decision-making problems, in which information about the environmental reward structure is detached from rewards for a subset of actions. In the current exploratory study, we introduce an information-selective symmetric reversal bandit task to model such situations and obtained choice data on this task from 24 participants. To arbitrate between different decision-making strategies that participants may use on this task, we developed a set of probabilistic agent-based behavioral models, including exploitative and explorative Bayesian agents, as well as heuristic control agents. Upon validating the model and parameter recovery properties of our model set and summarizing the participants' choice data in a descriptive way, we used a maximum likelihood approach to evaluate the participants' choice data from the perspective of our model set. In brief, we provide quantitative evidence that participants employ a belief state-based hybrid explorative-exploitative strategy on the information-selective symmetric reversal bandit task, lending further support to the finding that humans are guided by their subjective uncertainty when solving exploration-exploitation dilemmas.

**Keywords** Bandit problem · Agent-based behavioral modeling · Exploration · Exploitation

# Beispiel

## Gegenstandsbereich und Phänomen

Menschen müssen oft Entscheidungen unter Unsicherheit treffen

Menschen müssen manchmal informations- und gewinnbringende Handlungen abwägen

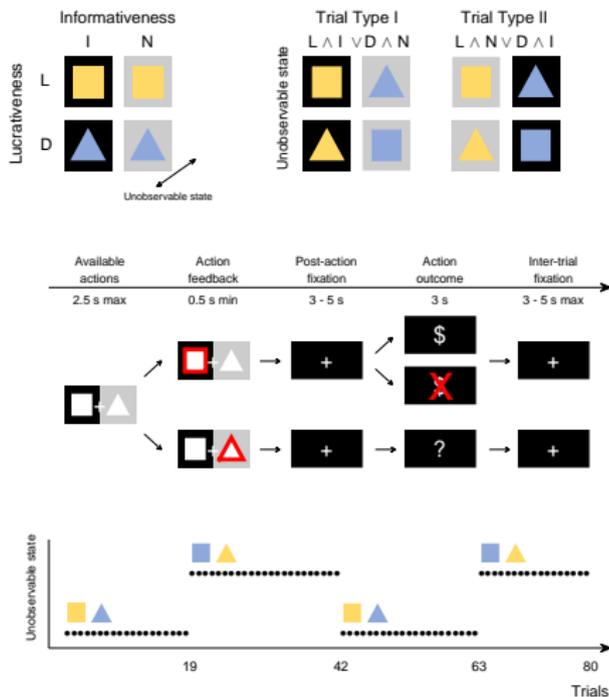


- Wie gehen Menschen dabei vor?
- Wie lernen Menschen in solchen Situationen Entscheidungen zu treffen?

Horvath et al. (2021)

# Beispiel

## Experimentelle Simulation



Horvath et al. (2021)

## Theorie

### Künstliche Intelligenz - Artificial Agent

$$M_{\text{Agent}} := (S, A, R, O, p(s_1^1), p(s_{t+1}^1 | s_t^1), p^{at}(o_t | s_t^1), p^{at}(r_t | s_t^1), v, d)$$

- Dynamisches handlungsabhängiges generatives Modell

$$p^{a_{1:T}}(s_{1:T}^1, o_{1:T}) = p(s_1^1) \prod_{t=1}^T p^{at}(o_t | s_t^1) p(s_{t+1}^1 | s_t^1)$$

- Handlungsabhängige Zustandsschätzung (Belief State)

$$p^{a_{1:t-1}}(s_t^1 | o_{1:t-1}) = \frac{\sum_{s_{t-1}^1} p(s_t^1 | s_{t-1}^1) p^{a_{t-1}}(o_{t-1} | s_{t-1}^1) p^{a_{1:t-2}}(s_{t-1}^1 | o_{1:t-2})}{\sum_{s_t^1} \sum_{s_{t-1}^1} p(s_t^1 | s_{t-1}^1) p^{a_{t-1}}(o_{t-1} | s_{t-1}^1) p^{a_{1:t-2}}(s_{t-1}^1 | o_{1:t-2})}$$

Horvath et al. (2021)

## Theorie

### Künstliche Intelligenz - Artificial Agent

$$M_{\text{Agent}} := (S, A, R, O, p(s_1^1), p(s_{t+1}^1 | s_t^1), p^{a_t}(o_t | s_t^1), p^{a_t}(r_t | s_t^1), v, d)$$

- Handlungswertungsfunktion

$$v : A \times [0, 1] \rightarrow \mathbb{R}, (a, b) \mapsto v(a, b)$$

- Entscheidungsfunktion

$$d : \mathbb{R} \rightarrow A, v(\cdot, b) \mapsto d(v(\cdot, b)) := \arg \max_{a \in A} v(a, b)$$

Horvath et al. (2021)

## Theorievarianten

### A1 | Gewinnmaximierender Agent

$$v_{A1}(a, b) := b\mathbb{E}_{p^a(r_t|s_t^1=1)}(r_t) + (1-b)\mathbb{E}_{p^a(r_t|s_t^1=2)}(r_t)$$

⇒ Erwartete Belohnung von  $a$  unter momentaner Zustandsschätzung  $b_t = b$

### A2 | Informationsmaximierender Agent

$$v_{A2}(a, b) := \sum_{o_t} p_{a_{1:t-1}, a_t=a}(o_t|o_{1:t-1}) \text{KL} \left( p_{a_{t-1}, a_t=a}(s_{t+1}^1|o_{1:t-1}, o_t) \parallel p_{a_{1:t-1}}(s_t^1|o_{1:t-1}) \right)$$

⇒ Erwartete Bayesianische Überraschung von  $a$  unter momentaner Zustandsschätzung  $b_t = b$

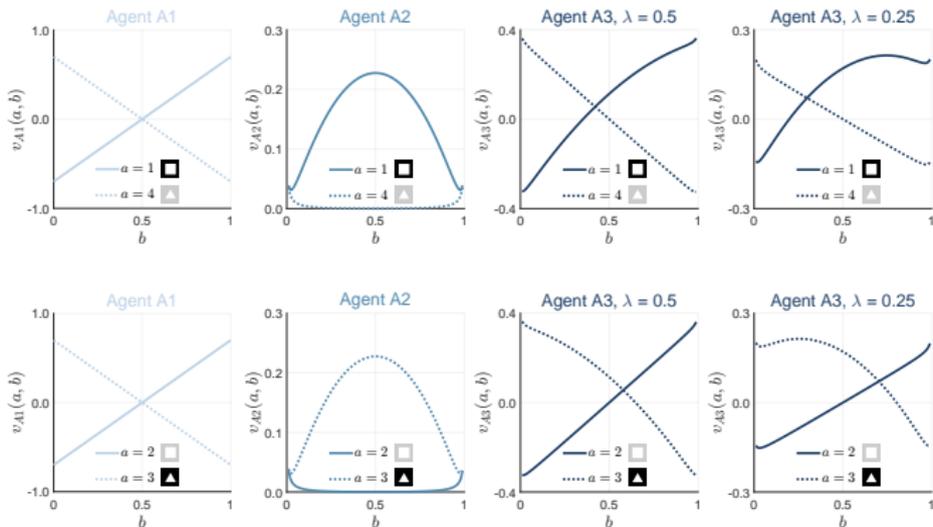
### A3 | Gewinn- und informationsmaximierender Agent

$$v_{A3}(a, b) := \lambda v_{A1}(a, b) + (1-\lambda)v_{A2}(a, b)$$

⇒ Gewichtete Kombination der beiden Theoriealternativen

Horvath et al. (2021)

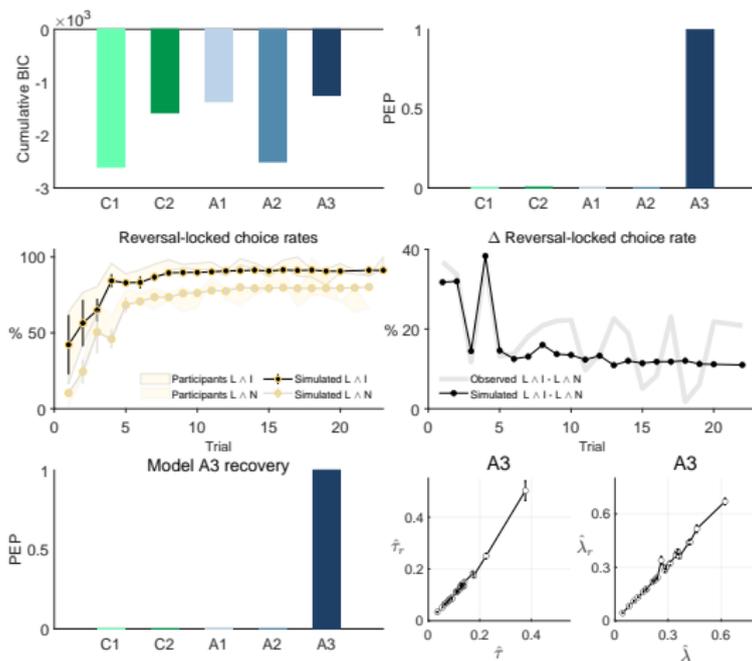
## Datenvorhersage - Hypothesenraum



Horvath et al. (2021)

# Beispiel

## Datenanalyse



Horvath et al. (2021)

---

Theorien

Hypothesen

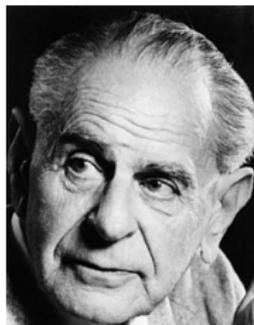
Experimente

Beispiel

**Wissenschaftstheorie**

Selbstkontrollfragen

## Kritischer Rationalismus



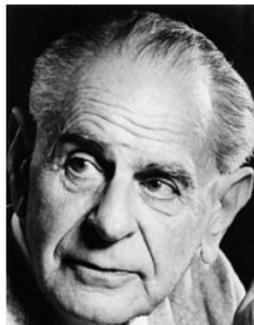
Karl Popper (1902 - 1994)

### Logik der Forschung (1935)

- Es gibt eine vom Menschen unabhängige externe Welt.
- Die externe Welt ist zumindest teilweise empirisch erfahrbar.
- Beobachtungen sind immer schon "theoriegeladen".
- Fortlaufende Prüfung eigener Grundannahmen und Methoden.

Popper (1935)

## Kritischer Rationalismus



Karl Popper (1902 - 1994)

### Falsifikationsprinzip

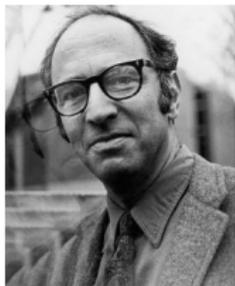
- Theorien können grundsätzlich nicht verifiziert werden.
- ⇒ Es mag immer noch eine bessere Theorie geben.
- Man sollte danach streben, Theorien zu widerlegen.

“Nicht nach Wahrheitsbeweisen ist in den Erfahrungswissenschaften zu suchen, denn diese sind dort grundsätzlich unmöglich; vielmehr müssen sogenannte Naturgesetze ausschließlich als Hypothesen betrachtet werden, die so lange beibehalten werden, als sie nicht falsifiziert sind.”

“Ein empirisch-wissenschaftliches System muss an der Erfahrung scheitern können.”

Popper (1935)

## Kritik des kritischen Rationalismus



Thomas S. Kuhn  
(1922 - 1996)

### The Structure of Scientific Revolutions (1962)

- Es gibt keine Logik der Forschung im Sinne Poppers.
  - Es gibt Normalzeiten, in diesen herrschen *Paradigmen* vor.
- ⇒ Man macht frequentistische Hypothesentests.
- Es gibt Krisen, in diesen finden *Paradigmenwechsel* statt.
- ⇒ Frequentistische Hypothesentests sind problematisch.
- ⇒ Man macht Modellvergleiche.

Kuhn (1962)

## Kritik des kritischen Rationalismus



Stuart Ritchie

### Science Fictions (2020)

- How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth.
- Das akademische Belohnungssystem fördert unwissenschaftliche Forschung.
- High Impact und Drittmittelinwerbungen wichtiger als Inhalt und Lehre.

⇒ Gehaltszulage €300 für vier Paper/Jahr (Erstautor? Koautor?)

Ritchie (2020)

⇒ Open Science ⇐

---

Theorien

Hypothesen

Experimente

Beispiel

Wissenschaftstheorie

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Geben Sie die Definition einer Theorie nach Reiß and Sarris (2012) wieder.
2. Geben Sie die Definition einer Theorie nach Ostwald (2021) wieder.
3. Nennen und erläutern Sie die Theoriegütekriterien nach Groeben and Westmeyer (1975).
4. Geben Sie die Definition einer Hypothese nach Reiß and Sarris (2012) wieder.
5. Geben Sie die Definition eines Experimentes nach Reiß and Sarris (2012) wieder.
6. Nennen und erläutern Sie die Experimentkriterien nach Wundt.
7. Erläutern Sie die Begriffe Experiment, Quasiexperiment und Korrelationsstudie.
8. Nennen Sie drei Grundideen des Kritischen Rationalismus nach Popper (1935).
9. Skizzieren Sie das Falsifikationsprinzip nach Popper (1935).
10. Nennen Sie zwei Grundideen der Arbeit von Kuhn (1962).

## References

---

- Bortz, Jürgen, and Nicola Döring. 2006. *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. 4., überarb. Aufl., [Nachdr.]. Springer-Lehrbuch Bachelor, Master. Heidelberg: Springer-Medizin-Verl.
- Gerrig, Richard J., Philip G. Zimbardo, Ralf Graf, and Richard J. Gerrig. 2020. *Psychologie*. 18., aktualis. Aufl., [Nachdr.]. ps psychologie. München: Pearson Studium.
- Groeben, Norbert, and Hans Westmeyer. 1975. *Kriterien psychologischer Forschung*. Grundfragen der Psychologie. München: Juventa Verlag.
- Horvath, Lilla, Stanley Colcombe, Michael Milham, Shruti Ray, Philipp Schwartenbeck, and Dirk Ostwald. 2021. "Human Belief State-Based Exploration and Exploitation in an Information-Selective Symmetric Reversal Bandit Task." *Computational Brain & Behavior*, August. <https://doi.org/10.1007/s42113-021-00112-3>.
- Kuhn, Thomas S. 1962. "The Structure of Scientific Revolutions," no. 2: 112.
- Myers, David G. 2010. *Psychology*. 9th ed. New York: Worth Publishers.
- Popper, Karl. 1935. *Logik der Forschung*. Vienna: Springer Vienna. <https://doi.org/10.1007/978-3-7091-4177-9>.
- Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.
- Ritchie, Stuart. 2020. *Science Fictions - Exposing Fraud, Bias, Negligence and Hype in Science*. Penguin.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2001. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.



# Psychologische Forschungsmethoden

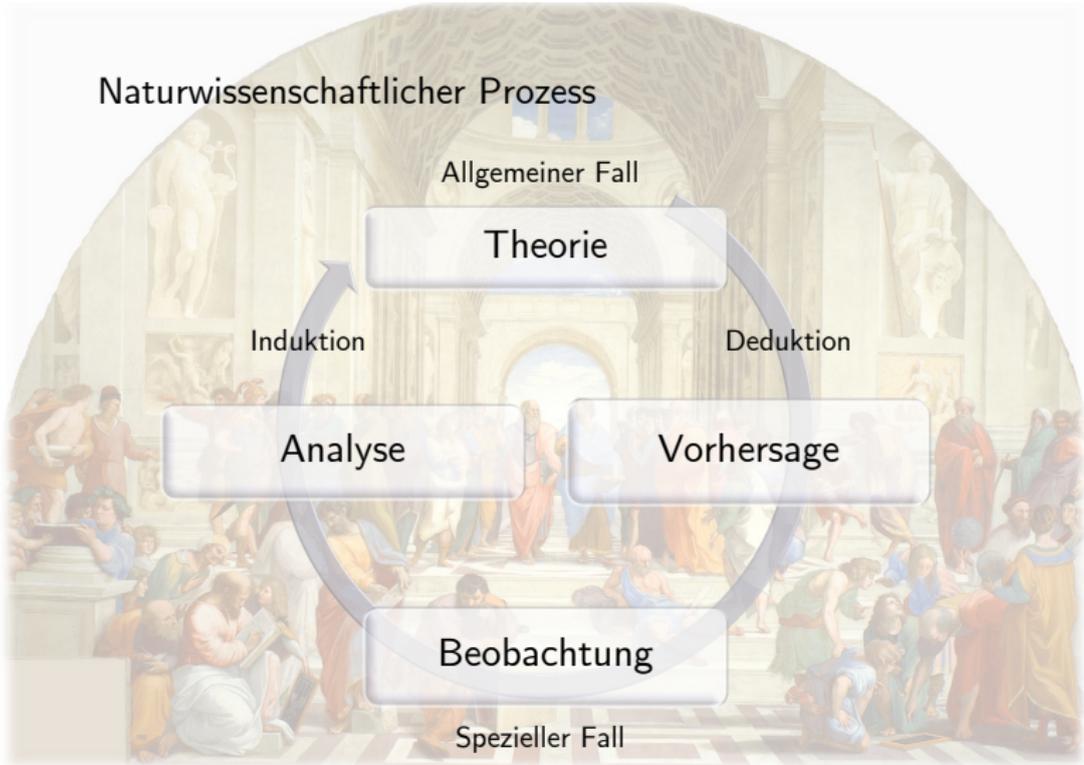
BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

(3) Operationalisieren, Messen, Skalieren

## Naturwissenschaftlicher Prozess



---

Operationalisierung

Variablentypen

Messen

Skalenniveaus

Selbstkontrollfragen

---

## **Operationalisierung**

Variablentypen

Messen

Skalenniveaus

Selbstkontrollfragen

## Variable

Eine Variable ist etwas, das durch Veränderlichkeit charakterisiert ist.

## Konstrukt

Ein Konstrukt ist ein in der Theorie generierter Erklärungs begriff, der sich nur indirekt und unter Zuhilfenahme operationaler Definitionen empirisch erfassen lässt.

## Operationalisieren

Operationalisieren bezeichnet den Prozess der Umsetzung eines Konstrukts in eine empirisch messbare Variable.

Reiß and Sarris (2012)

## Konstrukt | Entscheidungsfindung unter Unsicherheit

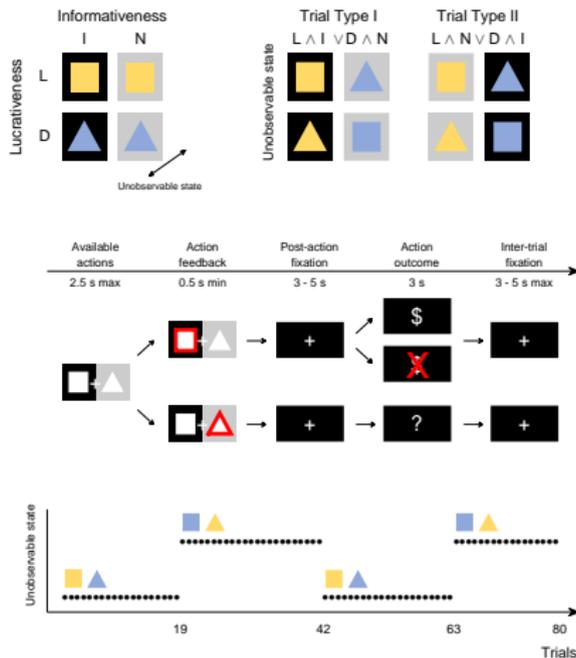
Menschen müssen oft Entscheidungen unter Unsicherheit treffen

Menschen müssen manchmal informations- und gewinnbringende Handlungen abwägen



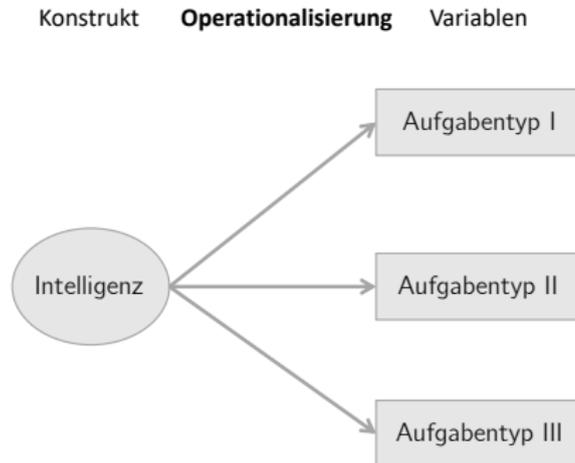
- Wie gehen Menschen dabei vor?
- Wie lernen Menschen in solchen Situationen Entscheidungen zu treffen?

Horvath et al. (2021)

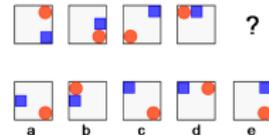


Horvath et al. (2021)

## Konstrukt und Operationalisierung | Intelligenz



Wähle die Figur, welche die Reihe fortsetzt



Zahlenreihen

Welche Zahl ist die logische Fortsetzung dieser Reihe:

4 6 9 6 6 14 6 ...

- 6  
 17  
 19  
 21

Finden Sie das Wort heraus,  
das nicht in diese Reihe passt.

- essen  
 trinken  
 fernsehen  
 schlafen  
 atmen

---

Operationalisierung

## **Variablentypen**

Messen

Skalenniveaus

Selbstkontrollfragen

## Unabhängige Variable (UV)

Etwas, das in einem Experiment systematisch variiert wird, um seine Auswirkung auf eine oder mehrere abhängige Variable(n) zu untersuchen.

## Abhängige Variable (AV)

Etwas, das in einem Experiment erfasst wird, um zu überprüfen, wie sich systematisch variierte unabhängige Variablen Auswirkungen.

## Beispiele

- Einfluss von Alkoholkonsum (UV) auf Reaktionszeiten (AV)
- Einfluss des Erziehungsstils (UV) auf die Kreativität von Kindern (AV)
- Einfluss der Belohnungsanzahl (UV) auf die Leistungsmotivation (AV)
- Einfluss des Entscheidungskontext (UV) auf das Entscheidungsverhalten (AV)

## Diskrete Variablen

Diskrete (kategoriale) Variablen sind Variablen, die nur eine endliche Anzahl an verschiedenen Werten annehmen und meist durch ganze Zahlen repräsentiert sind.

## Kontinuierliche Variablen

Kontinuierliche Variablen sind Variablen, die unendlich viele Werte annehmen können und meist durch die reellen Zahlen repräsentiert sind.

Einordnung einer Variable als diskret oder kontinuierlich ist eine Modellierungsannahme

Geschlecht	m/w vs. m/w/d vs. Kontinuum
Alter	Zeit als reelle Zahl vs. 20, 21, 22, ..., 100
Reaktionszeiten	Zeit als reelle Zahl vs. floating point numbers

# Variablentypen

---

## Spezielle Variablentypen

### Organismusvariablen

Variablen, die an individuelle experimentelle Einheiten gebunden sind.

- Alter, Geschlecht, Augenfarbe, ...

### Reizvariablen

Variablen, die als Umweltreiz auf den Organismus einwirken.

- Sensorischer Stimulus, Aufgabenstellung, ...

### Störvariablen

Variablen, die im Rahmen einer Studie nicht kontrolliert werden.

- Psychophysiologische Dynamiken, Lebenssituation, ...

### Moderatorvariablen

Variablen, die den Einfluss von UV auf AV beeinflussen.

## Beispiel

Ein Forscherteam möchte untersuchen, ob die wöchentlich aufgewendete Zeit für soziale Netzwerke die Qualität von Offline-Freundschaften bei Jugendlichen beeinflusst. Bekannt ist, dass die Qualität von Offline-Freundschaften von der Dauer der Bekanntschaft, Ausmaß der Selbstoffenbarung und dem Verbundenheitsgefühl abhängt, nicht jedoch vom Geschlecht. Außerdem vermutet das Forscherteam, dass die Enge des Zusammenhangs von aufgewendeter Zeit in sozialen Netzwerken und Qualität von Offline-Freundschaften von dem Bedürfnis nach Nähe beeinflusst wird.

Zeit in sozialen Netzwerken	→ Unabhängige Variable
Qualität von Offline Freundschaften	→ Abhängige Variable
Geschlecht	→ Organismusvariable
Bekanntschafsdauer	→ Störvariable
Verbundenheitsgefühl	→ Störvariable
Selbstoffenbarungsausmaß	→ Störvariable
Nähebedürfnis	→ Moderatorvariable

## Anmerkungen

Die Begriffsdefinitionen UV und AV sind kontraintuitiv.

Variablentypen sollten nicht mit mathematischen Variablen verwechselt werden.

Alle Variablentypen können als Zufallsvariablen modelliert werden oder auch nicht.

Die Zuteilung messbarer Entitäten zu Variablentypen ist ein subjektiv-kreativer Prozess:

“One researcher's signal is another researcher's noise.”

---

Operationalisierung

Variablentypen

**Messen**

Skalenniveaus

Selbstkontrollfragen

## Messen

Unter *Messen* versteht man die Zuordnung von Zahlen zu Objekten nach bestimmten Regeln, die gewährleisten, dass bestimmte interessierende Relationen in der Menge der Objekte in der Menge der Zahlen erhalten bleiben. Die theoretische Untersuchung von Messvorgängen heißt *Messtheorie*.

## Empirisches Relativ

Die Menge von zu messenden Objekten mit einer zugehörigen Relation

- Äquivalenzrelationen (A und B sind dem Wesen nach gleich, A und C nicht)
- Ordnungsrelationen (A ist größer als B, C ist größer als A)

## Numerisches Relativ

Eine Menge von Zahlen mit einer zugehörigen Relation

- Äquivalenzrelationen ( $\{A, B, C\}, A \sim B, A \not\sim C$ )
- Ordnungsrelationen ( $(\mathbb{R}, =, <)$ )

## Homomorphismus

Ein Homomorphismus ist eine Abbildung, die die Relation des empirischen Relativs im numerischen Relativ erhält. Ein Homomorphismus ist also *strukturerhaltend*.

Beispiel (1)

Empirische Relation: A studiert Psychologie, B studiert Psychologie, C studiert PNK

$$\begin{array}{ll} \text{Menge des empirisches Relativs} & E := \{A, B, C\} \\ \text{Menge des numerischen Relativs} & R := \{0, 1\} \end{array}$$

Homomorphismus

$$h : E \rightarrow R, \begin{cases} A \mapsto 0 \\ B \mapsto 0 \\ C \mapsto 1 \end{cases} \quad (1)$$

Beispiel (2)

Empirische Relation: A ist größer als B, B ist größer als C

$$\begin{array}{ll} \text{Menge des empirisches Relativs} & E := \{A, B, C\} \\ \text{Menge des numerischen Relativs} & R := \mathbb{R} \end{array}$$

Homomorphismus

$$h : E \rightarrow R, \begin{cases} A \mapsto 174 \\ B \mapsto 170 \\ C \mapsto 165 \end{cases} \quad (2)$$

---

Operationalisierung

Variablentypen

Messen

**Skalenniveaus**

Selbstkontrollfragen

## Skala

Eine Skala ist die Einheit eines empirischen Relativs, eines numerischen Relativs und einer homomorphen Abbildung. Die spezielle Form des empirischen und numerischen Relativs bestimmen dabei das sogenannte **Skalenniveau**.

### Skalenniveaus nach Stevens (1946)

Nominalskala	Nur Äquivalenzrelationen, keine Ordnungsrelationen
Ordinalskala	Ordnungsrelationen
Intervallskala	Ordnungsrelationen mit gleichen Abständen zwischen Skalenpunkten
Verhältnisskala	Ordnungsrelationen mit gleichen Abständen und empirischem Nullpunkt

## Nominalskala

### Definition

Die Nominalskala ordnet den Objekten des empirischen Relativs Zahlen zu, die so geartet sind, dass Objekte mit gleicher Merkmalsausprägung gleiche Zahlen und Objekte mit verschiedenen Merkmalsausprägungen verschiedene Zahlen erhalten.

Bortz and Döring (2006)

### Eigenschaften

- Die Zahlen einer Nominalskala sind Namen für Äquivalenzklassen ohne quantitative Bedeutung.
- Je zwei Objekte des empirischen und numerischen Relativs sind äquivalent oder nicht äquivalent.

### Beispiel

#### Studienfach

- A studiert Psychologie, B studiert Psychologie, C studiert PNK.
- A und B sind äquivalent, A und C sind nicht äquivalent, B und C sind nicht äquivalent
- Nominalskala  $\{0, 1\}$  mit  $0 = \text{Psychologie}$ ,  $1 = \text{PNK}$
- $A \rightarrow 0$ ,  $B \rightarrow 0$ ,  $C \rightarrow 1$ .

## Ordinalskala

### Definition

Die Ordinalskala ordnet den Objekten eines empirischen Relativs Zahlen zu, die so geartet sind, dass von jeweils zwei Objekten das Objekt mit der größeren Merkmalsausprägung die größere (manchmal auch kleinere) Zahl erhält.

Bortz and Döring (2006)

### Eigenschaften

- Die Zahlen einer Ordinalskals bilden eine Rangordnung im empirischen Relativ ab.
- Die Abstände zwischen zwei Rängen müssen nicht numerisch gleich sein.

### Beispiel

#### ESC 1974 Plätze

- 1. Platz: ABBA, 2. Platz: Cinquetti, 3. Platz: MacNeal.
- Ordnungsrelation im empirischen Relativ:  $ABBA > Gigliola Cinquetti > Mouth and MacNeal$ .
- Ordinalskala  $\{1, 2, 3\}$  mit  $1 = 1. \text{ Platz}$ ,  $2 = 2. \text{ Platz}$ ,  $3 = 3. \text{ Platz}$ .
- $ABBA \rightarrow 1$ ,  $Cinquetti \rightarrow 2$ ,  $MacNeal \rightarrow 3$ .
- ABER:  $ABBA \rightarrow 24 \text{ Punkte}$ ,  $Cinquetti \rightarrow 18 \text{ Punkte}$ ,  $MacNeal \rightarrow 15 \text{ Punkte}$ .

## Intervallskala

### Definition

Die Intervallskala ordnet den Objekten des empirischen Relativs Zahlen zu, die so geartet sind, dass die Verhältnisse der Differenzen zwischen je zwei Merkmalsausprägungen den Verhältnissen der Differenzen zwischen je zwei Zahlen der Skala entspricht. Eine Intervallskala zeichnet sich also durch Äquidistanz der Messwerte aus.

Bortz and Döring (2006)

### Beispiel und Eigenschaften

- Die Celsius Temperaturskala und die Fahrenheit Temperaturskala sind Intervallskalen
- Es gilt  $T_F = 1.8 \cdot T_C + 32$ , also z.B.  $10^\circ\text{C} = 50^\circ\text{F}$ ,  $20^\circ\text{C} = 68^\circ\text{F}$ ,  $30^\circ\text{C} = 86^\circ\text{F}$ .
- Bei Intervallskalen sind die Verhältnisse von Wertdifferenzen invariant, z.B.

$$\frac{30^\circ\text{C} - 10^\circ\text{C}}{30^\circ\text{C} - 20^\circ\text{C}} = \frac{20^\circ\text{C}}{10^\circ\text{C}} = 2 \text{ und } \frac{86^\circ\text{F} - 50^\circ\text{F}}{86^\circ\text{F} - 68^\circ\text{F}} = \frac{36^\circ\text{F}}{18^\circ\text{F}} = 2. \quad (3)$$

- Bei Intervallskalen sind die Verhältnisse von Werten allerdings variant, z.B.

$$\frac{20^\circ\text{C}}{10^\circ\text{C}} = 2.00 \text{ und } \frac{68^\circ\text{F}}{50^\circ\text{F}} = 1.36. \quad (4)$$

## Verhältnisskala

### Definition

Die Verhältnisskala ordnet den Objekten des empirischen Relativs Zahlen zu, die so geartet sind, dass die Verhältnisse zwischen je zwei Merkmalsausprägungen den Verhältnissen zwischen je zwei Zahlen der Skala entspricht. Eine Verhältnisskala benötigt einen natürlichen Nullpunkt im empirischen und numerischen Relativ.

Bortz and Döring (2006)

### Beispiele

- Die Kelvin Temperaturskala bildet physikalische Energiezustände auf Zahlen ab.
- Die Längenskala in Meter bildet die physikalische Länge auf Zahlen ab.
- Psychologiemethodenbücher führen auch die Reaktionszeit als Verhältnisskala auf.

Eine durchdachte Einführung in die Messtheorie gibt es im SoSe 2022!

---

Operationalisierung

Variablentypen

Messen

Skalenniveaus

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Definieren Sie die Begriffe Variable, Konstrukt, und Operationalisierung nach Reiß and Sarris (2012).
2. Definieren Sie die Begriffe Unabhängige Variable und Abhängige Variable.
3. Definieren Sie die Begriffe Diskrete Variable und Kontinuierliche Variable.
4. Definieren Sie die Begriffe Organismusvariable, Reizvariable, Störvariable und Moderatorvariable.
5. Definieren Sie den Begriff Messen.
6. Erläutern Sie die Begriffe der Äquivalenzrelation und der Ordnungsrelation.
7. Definieren Sie die Begriffe Empirischen Relativ und Numerisches Relativ.
8. Erläutern Sie den Begriff des Homomorphismus.
9. Definieren Sie den Begriff der Skala.
10. Geben Sie die Definition einer Nominalskala nach Bortz and Döring (2006) wieder.
11. Geben Sie die Definition einer Ordinalskala nach Bortz and Döring (2006) wieder.
12. Geben Sie die Definition einer Intervallskala nach Bortz and Döring (2006) wieder.
13. Geben Sie die Definition einer Verhältnisskala nach Bortz and Döring (2006) wieder.

## References

---

- Bortz, Jürgen, and Nicola Döring. 2006. *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. 4., überarb. Aufl., [Nachdr.]. Springer-Lehrbuch Bachelor, Master. Heidelberg: Springer-Medizin-Verl.
- Horvath, Lilla, Stanley Colcombe, Michael Milham, Shruti Ray, Philipp Schwartenbeck, and Dirk Ostwald. 2021. "Human Belief State-Based Exploration and Exploitation in an Information-Selective Symmetric Reversal Bandit Task." *Computational Brain & Behavior*, August. <https://doi.org/10.1007/s42113-021-00112-3>.
- Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.
- Stevens, S. S. 1946. "On the Theory of Scales of Measurement." *Science, New Series* 103 (2684): 677–80.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (4) Prinzipien der Versuchsplanung

---

Versuchsplan

Varianzquellen

Max-Kon-Min Prinzip

Selbstkontrollfragen

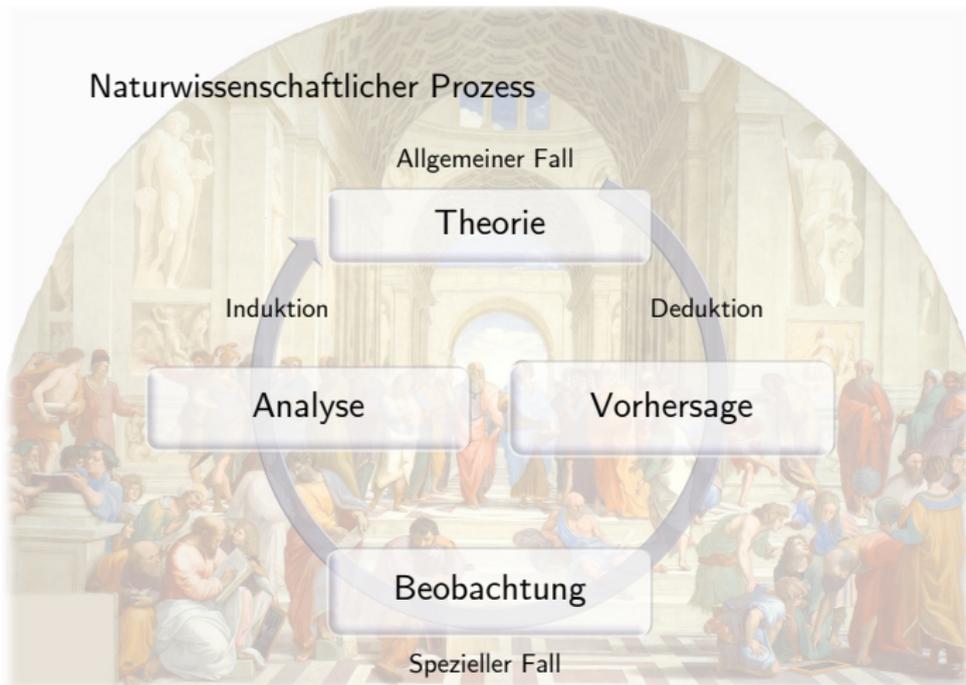
---

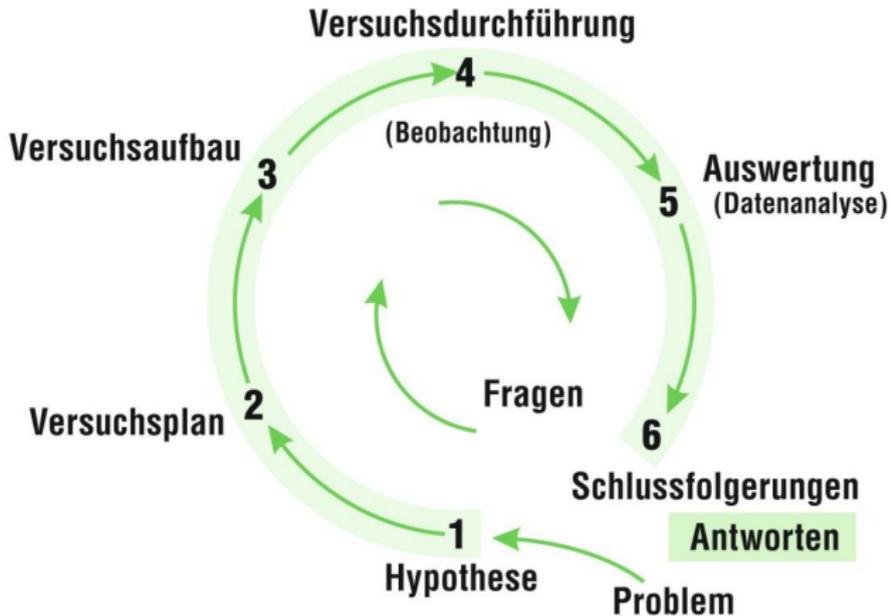
## **Versuchsplan**

Varianzquellen

Max-Kon-Min Prinzip

Selbstkontrollfragen





## Definition

Unter einem Versuchsplan versteht man ein standardisiertes, routinemäßig anwendbares Strukturschema, das dem Aufbau, der Kontrolle und der methodologischen Bewertung einer empirischen Untersuchung von unabhängigen und abhängigen Variablen sachlogisch zugrunde liegt.

Reiß and Sarris (2012)

## Beispiel: Evaluation von Psychotherapieformen bei Depression

### Online Psychotherapie



### Klassische Psychotherapie



# Versuchsplan

## Beispiel: Evaluation von Psychotherapieformen bei Depression

### Becks Depressions-Inventar (BDI) zur Depressionsdiagnostik

BDI-II Fragebogen	
Name	Wiev. erreichte Punkte
	von 0 bis 63
<p><b>Anleitung:</b> Dieser Fragebogen enthält 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Gruppen von Aussagen sorgfältig durch und wählen Sie sich dann in jeder Gruppe eine Aussage heraus, die am besten beschreibt, wie Sie sich in der letzten zwei Wochen, einschließlich heute, gefühlt haben. Kennen Sie die Zahl neben der Aussage an, die Sie als beste Aussage wählen (0, 1, 2 oder 3). Falls in einer Gruppe mehrere Aussagen gleichwertig auf Sie zutreffen, können Sie für Aussage mit der höchsten Zahl an. Achten Sie bitte darauf, dass Sie in jeder Gruppe nicht mehr als eine Aussage ankreuzen, die die gleiche BDI Gruppe (Veränderungen der Schlafgewohnheiten) oder Gruppe 16 (Veränderungen des Appetits).</p>	
<p><b>1.) Traurigkeit</b></p> <p>0 Ich bin nicht traurig. 1 Ich bin oft traurig. 2 Ich bin ständig traurig. 3 Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.</p> <p><b>2.) pessimismus</b></p> <p>0 Ich sehe nicht mal in die Zukunft. Ich sehe nur das Schlimme in der Zukunft als Folge. 1 Ich bin müde und erwarte nicht, dass meine Situation besser wird. 2 Ich glaube, dass meine Zukunft hoffnungslos ist und nur noch schlechter wird.</p> <p><b>3.) Versagensgefühle</b></p> <p>0 Ich fühle mich nicht als Versager. 1 Ich habe häufiger Versagensgefühle. 2 Wenn ich zurückblicke, sehe ich eine Menge Fehlertage. 3 Ich habe das Gefühl, ich mache ein völliger Versager zu sein.</p> <p><b>4.) Verlust von Freude</b></p> <p>0 Ich kann die Dinge genauso gut genießen wie früher. 1 Ich kann die Dinge nicht mehr so genießen wie früher. 2 Dinge, die mir früher Freude gemacht haben, kann ich kaum mehr genießen. 3 Dinge, die mir früher Freude gemacht haben, kann ich überhaupt nicht mehr genießen.</p> <p><b>5.) Schuldgefühle</b></p> <p>0 Ich habe keine besonderen Schuldgefühle. 1 Ich habe oft Schuldgefühle wegen Dingen, die ich getan habe oder hätte tun sollen. 2 Ich habe die meisten Zeit Schuldgefühle. 3 Ich habe ständig Schuldgefühle.</p>	<p><b>6.) Bestürzungsgefühle</b></p> <p>0 Ich habe viele das Gefühl, für etwas bestraft zu sein. 1 Ich habe das Gefühl, vielleicht bestraft zu werden. 2 Ich erwarte, bestraft zu werden. 3 Ich habe das Gefühl, bestraft zu sein.</p> <p><b>7.) Selbsthöhnung</b></p> <p>0 Ich habe von mir genauso viel wie immer. 1 Ich habe Vertrauen in mich verloren. 2 Ich bin von mir enttäuscht. 3 Ich lehne mich völlig ab.</p> <p><b>8.) Selbstvorwürfe</b></p> <p>0 Ich kritisiere oder tadle mich nicht mal als Versager. 1 Ich bin mir gegenüber kritischer als sonst. 2 Ich kritisiere mich für all meine Mängel. 3 Ich gebe mir die Schuld für alles Schlechte, was passiert.</p> <p><b>9.) Selbstmordgedanken</b></p> <p>0 Ich denke nicht daran, mir etwas anzutun. 1 Ich denke manchmal an Selbstmord. 2 Ich möchte mich umbringen, umfrage ich würde mich umbringen, wenn ich die Gelegenheit dazu hätte.</p> <p><b>10.) Weinen</b></p> <p>0 Ich weine nicht öfter als früher. 1 Ich weine jetzt mehr als früher. 2 Ich weine beim geringsten Anlass. 3 Ich möchte gar weinen, aber ich kann nicht.</p>

<p><b>11.) Unruhe</b></p> <p>0 Ich bin nicht unruhiger als sonst. 1 Ich bin unruhiger als sonst. 2 Ich bin so unruhig, dass es mir schwerfällt, still zu sitzen. 3 Ich bin so unruhig, dass ich mich ständig bewegen oder etwas tun muss.</p> <p><b>12.) Interessensverlust</b></p> <p>0 Ich habe das Interesse an anderen Menschen oder an Tätigkeiten nicht verloren. 1 Ich habe weniger Interesse an anderen Menschen oder an Dingen als sonst. 2 Ich habe das Interesse an anderen Menschen oder Dingen zum größten Teil verloren. 3 Es fällt mir schwer, mich überhaupt für irgend etwas zu interessieren.</p> <p><b>13.) Entschlussunfähigkeit</b></p> <p>0 Ich bin so entscheidungsfreudig wie immer. 1 Es fällt mir schwerer als sonst, Entscheidungen zu treffen. 2 Es fällt mir sehr viel schwerer als sonst, Entscheidungen zu treffen. 3 Ich habe Mühe, überhaupt Entscheidungen zu treffen.</p> <p><b>14.) Wertlosigkeit</b></p> <p>0 Ich fühle mich nicht wertlos. 1 Ich fühle mich für weniger wertvoll und nützlich als sonst. 2 Vergleichen mit anderen Menschen fühle ich mich viel weniger wert. 3 Ich fühle mich völlig wertlos.</p> <p><b>15.) Energieverlust</b></p> <p>0 Ich habe so viel Energie wie immer. 1 Ich habe weniger Energie als sonst. 2 Ich habe so wenig Energie, dass ich kaum noch etwas schaffe. 3 Ich habe keine Energie mehr, um überhaupt noch etwas zu tun.</p> <p><b>16.) Veränderungen der Schlafgewohnheiten</b></p> <p>0 Meine Schlafgewohnheiten haben sich nicht verändert. 1a Ich schlafe etwas mehr als sonst. 1b Ich schlafe etwas weniger als sonst. 2a Ich schlafe viel mehr als sonst. 2b Ich schlafe viel weniger als sonst. 3a Ich schlafe fast den ganzen Tag. 3b Ich wache 1-2 Stunden früher auf als gewöhnlich und kann dann nicht mehr einschlafen.</p>	<p><b>17.) Reizbarkeit</b></p> <p>0 Ich bin nicht reizbarer als sonst. 1 Ich bin reizbarer als sonst. 2 Ich bin viel reizbarer als sonst. 3 Ich fühle mich dauernd gereizt.</p> <p><b>18.) Veränderungen des Appetits</b></p> <p>0 Mein Appetit hat sich nicht verändert. 1a Mein Appetit ist etwas schlechter als sonst. 1b Mein Appetit ist etwas größer als sonst. 2a Mein Appetit ist viel schlechter als sonst. 2b Mein Appetit ist viel größer als sonst. 3a Ich habe überhaupt keinen Appetit. 3b Ich habe ständig Heißhunger.</p> <p><b>19.) Konzentrationschwierigkeiten</b></p> <p>0 Ich kann mich so gut konzentrieren wie immer. 1 Ich kann mich nicht mehr so gut konzentrieren wie sonst. 2 Es fällt mir schwer, mich längere Zeit auf irgend etwas zu konzentrieren. 3 Ich kann mich überhaupt nicht mehr konzentrieren.</p> <p><b>20.) Ermüdung oder Erschöpfung</b></p> <p>0 Ich fühle mich nicht müde oder erschöpft als sonst. 1 Ich werde schneller müde oder erschöpft als sonst. 2 Für viele Dinge, die ich üblicherweise tue, bin ich zu müde oder erschöpft. 3 Ich bin so müde oder erschöpft, dass ich fast nichts mehr tun kann.</p> <p><b>21.) Verlust an sexuellem Interesse</b></p> <p>0 Mein Interesse an Sexualität hat sich in letzter Zeit nicht verändert. 1 Ich interessiere mich weniger für Sexualität als früher. 2 Ich interessiere mich jetzt viel weniger für Sexualität. 3 Ich habe das Interesse an Sexualität völlig verloren.</p>
--	--

0 - 8 keine Depression

9 - 13 minimale Depression

14 - 19 leichte Depression

20 - 28 mittelschwere Depression

29 - 63 schwere Depression

## Beispiel: Evaluation von Psychotherapieformen bei Depression

Experimentelle Bedingung  
(Gruppen von  $n = 12$ )

Psychotherapie

Klassisch

Prä-BDI



Post-BDI

Online

Prä-BDI



Post-BDI

## Beispiel: Evaluation von Psychotherapieformen bei Depression

Proband:in	Bedingung	Prä-BDI	Post-BDI	BDI Differenz
1	Klassisch	11	9	-2
2	Klassisch	14	7	-7
3	Klassisch	10	10	0
4	Klassisch	12	11	-1
5	Klassisch	15	7	-8
6	Klassisch	11	7	-4
7	Klassisch	11	9	-2
8	Klassisch	11	10	-1
9	Klassisch	12	12	0
10	Klassisch	12	10	-2
11	Klassisch	14	11	-3
12	Klassisch	11	10	-1
13	Online	10	8	-2
14	Online	12	7	-5
15	Online	10	7	-3
16	Online	12	10	-2
17	Online	11	13	2
18	Online	8	9	1
19	Online	12	10	-2
20	Online	12	13	1
21	Online	14	9	-5
22	Online	14	11	-3
23	Online	15	8	-7
24	Online	13	9	-4

## Beispiel: Evaluation von Psychotherapieformen bei Depression

Klassisch	Online
-2	-2
-7	-5
0	-3
-1	-2
-8	2
-4	1
-2	-2
-1	1
0	-5
-2	-3
-3	-7
-1	-4

### Unabhängige Variable

Therapieform mit Werten "Klassisch" und "Online"

### Abhängige Variable

BDI Differenzwert

### Datenwertkodierung

$X_{ij}$  mit  $i = 1, 2$  und  $j = 1, \dots, 12$

$i$ : Proband:in Nummer,  $j$ : Therapieform

---

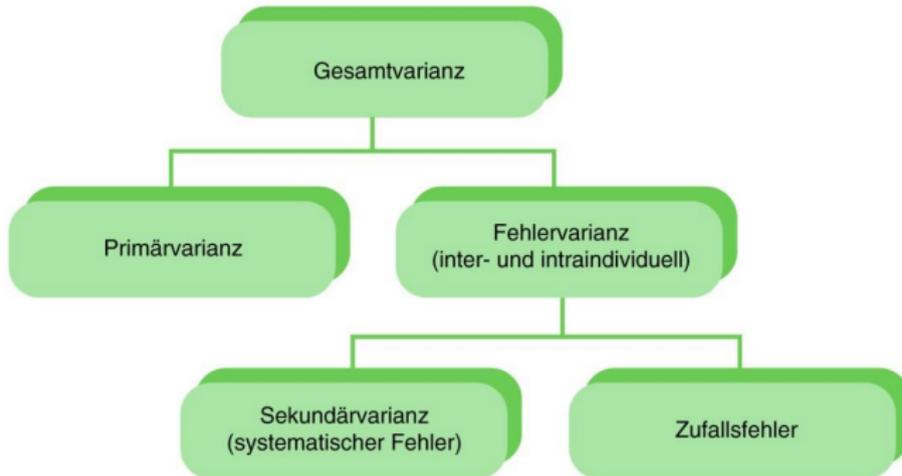
Versuchsplan

## **Varianzquellen**

Max-Kon-Min Prinzip

Selbstkontrollfragen

## Varianzquellen



Reiß and Sarris (2012)

# Varianzquellen

---

Gesamtvarianz = Primärvarianz + Fehlervarianz

## Primärvarianz

Systematische Veränderung der AV, die allein auf Variation der UV zurückzuführen ist.

Fehlervarianz = Sekundärvarianz + Zufallsfehler

## Sekundärvarianz

Systematische Veränderung der AV, die auf die Wirkung von unkontrollierten Störvariablen, nicht aber auf die Variation der UV, zurückzuführen ist.

## Zufallsfehler

Unsystematische Veränderung der AV, die weder auf die Variation der UV, noch auf den Einfluss von Störvariablen zurückzuführen ist.

Reiß and Sarris (2012)

## Das eigentliche Thema ist die einfaktorielle (Ko)Varianzanalyse

### Theorem (Quadratsummenzerlegung)

Für  $i = 1, \dots, p$  und  $j = 1, \dots, m$  bezeichne  $X_{ij}$  die  $j$ te Stichprobenvariable der  $i$ ten Stichprobengruppe. Weiterhin seien

$$\bar{X} := \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^m X_{ij} \quad \text{und} \quad \bar{X}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \quad (1)$$

den *Gesamtmittelwert* (*grand mean*) und das  $i$ te Stichprobenmittel, respektive. Schließlich seien

$$\text{SQT} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X})^2 \quad \text{die Total Sum of Squares}$$

$$\text{SQB} := \sum_{i=1}^p m(\bar{X}_i - \bar{X})^2 \quad \text{die Between-Group Sum of Squares}$$

$$\text{SQW} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 \quad \text{die Within-Group Sum of Squares}$$

Dann gilt

$$\text{SQT} = \text{SQW} + \text{SQB}. \quad (2)$$

SQT ist die Gesamtvarianz, SQB ist die Primärvarianz, SQW ist die Fehlervarianz

---

Versuchsplan

Varianzquellen

**Max-Kon-Min Prinzip**

Selbstkontrollfragen

MAXimiere die Primärvarianz

KONTrolliere die Sekundärvarianz

MINimiere die Fehlervarianz

Kerlinger (1986)

## MAXimiere die Primärvarianz

- Wahl von Extremgruppen
- Wahl von optimalen Stufen mit maximaler experimenteller Stimulation
- Umwandlung eines Störfaktors in eine experimentelle UV

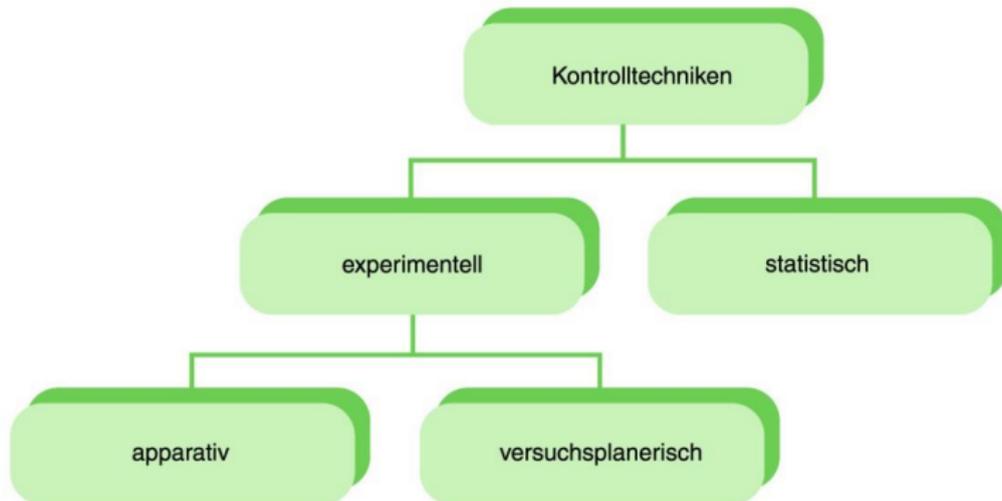
## KONtrolliere die Sekundärvarianz

- Eliminierung von Störfaktoren
- Konstanthaltung von Störfaktoren über experimentelle Bedingungen
- Randomisierung von Proband:innen und experimentellen Bedingungen
- Statistische Kontrolle durch Kovarianzanalyse
- Umwandlung eines Störfaktors in eine experimentelle UV

## MINimiere die Fehlervarianz

- Erhöhung der Standardisierung der Versuchsdurchführung
- Erhöhung der Zuverlässigkeit des Messinstrumentes
- Wahl eines Versuchsplans mit wiederholter Messung

## Max-Kon-Min Kontrolltechniken



Reiß and Sarris (2012)

---

Versuchsplan

Varianzquellen

Max-Kon-Min Prinzip

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Definieren Sie den Begriff des Versuchsplans nach Reiß and Sarris (2012).
2. Definieren Sie die Begriffe Primärvarianz und Sekundärvarianz.
3. Definieren Sie die Begriffe Fehlervarianz und Zufallsfehler.
4. Erläutern Sie den Zusammenhang von Gesamtvarianz, Primärvarianz und Sekundärvarianz.
5. Geben Sie Max-Kon-Min Prinzip wieder.
6. Nennen Sie drei Möglichkeiten zum Erreichen des MAX-Ziels im Max-Kon-Min Prinzip.
7. Nennen Sie drei Möglichkeiten zum Erreichen des KON-Ziels im Max-Kon-Min Prinzip.
8. Nennen Sie drei Möglichkeiten zum Erreichen des MIN-Ziels im Max-Kon-Min Prinzip.

## References

---

Kerlinger, Fred N. 1986. *Foundations of Behavioral Research*. 3rd ed. New York: Holt, Rinehart and Winston.

Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (5) Kriterien der Versuchsplanung

---

Überblick

Interne und externe Validität

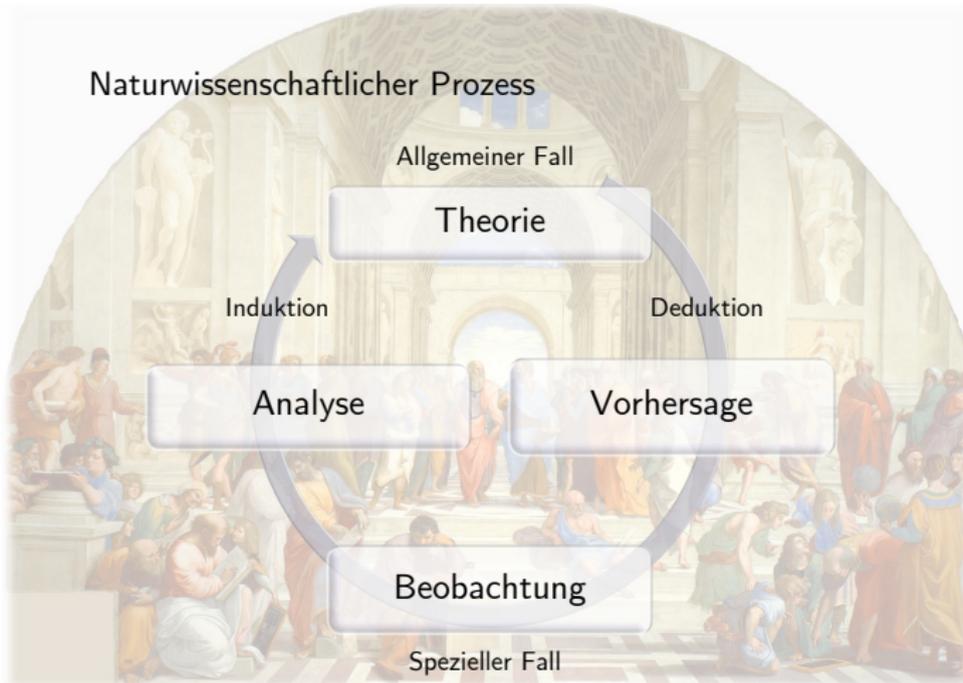
Selbstkontrollfragen

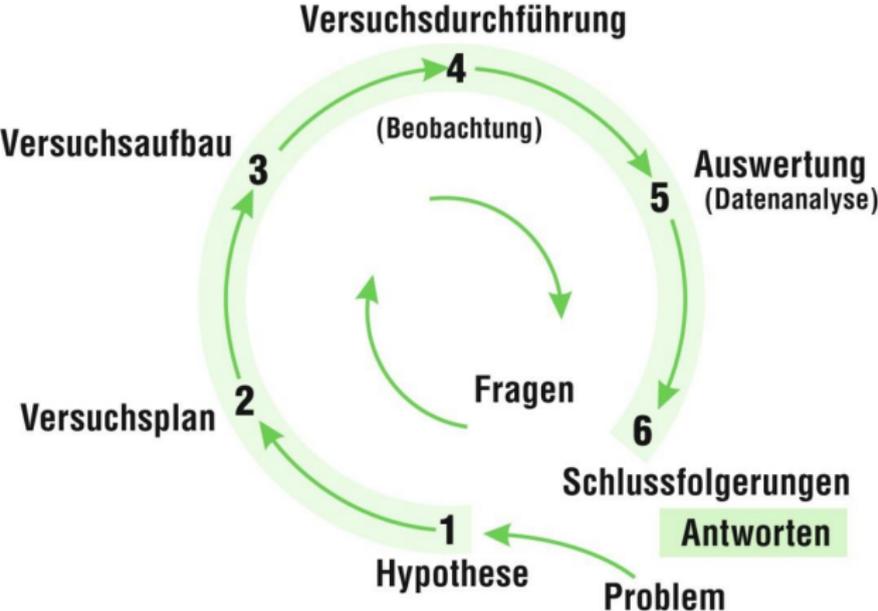
---

## Überblick

Interne und externe Validität

Selbstkontrollfragen





## Definition

Unter einem Versuchsplan versteht man ein standardisiertes, routinemäßig anwendbares Strukturschema, das dem Aufbau, der Kontrolle und der methodologischen Bewertung einer empirischen Untersuchung von unabhängigen und abhängigen Variablen sachlogisch zugrunde liegt.

Reiß and Sarris (2012)

⇒ Welche Gütekriterien erfüllen unterschiedliche Versuchspläne im Allgemeinen?

## Allgemeine Systematik von Versuchsplänen

### Experiment

- Randomisierte kontrollierte Studie
- Die Untersuchungseinheiten werden den Versuchsbedingungen zufällig zugeordnet
- Beispiel: Online Psychotherapie vs. Klassische Psychotherapie bei Depression

### Quasiexperiment

- Nicht-randomisierte kontrollierte Studie
- Untersuchung natürlich bzw. bereits bestehender Gruppen
- Beispiel: Online Psychotherapie bei Depression vs. Schizophrenie

### Korrelationsstudie

- Nicht-randomisierte, nicht kontrollierte Studie
- Beobachtungsstudie ohne Intervention
- Beispiel: Analyse von Paneldaten

## Allgemeine Gütemerkmale von Versuchsplänen

### Vorhandensein einer kausaltheoretischen Hypothese vor Versuchsbeginn

- Klarer Fokus auf einen Mechanismus
- Ausrichtung der Versuchsplanung auf Hypothesentest
- Maximierung der Primärvarianz
- Typisch für bereits gut erschlossene Gegenstandsbereiche

### Manipulierbarkeit der unabhängigen Variable

- Explizite Formulierung des untersuchten Konstrukts
- Explizite Operationalisierung des untersuchten Konstrukts
- Max-Kon-Min Ansatz
- Typisch für bereits gut erschlossene Gegenstandsbereiche

### Kontrollierbarkeit aller übrigen Versuchsbedingungen

- Ausschalten von erwarteten Störgrößen
- Kontrolle von Sekundärvarianz, Minimierung des Zufallsfehlers
- Typisch für bereits gut erschlossene Gegenstandsbereiche

## Allgemeine Systematik von Versuchsplänen und Güte Merkmalen

	1	2	3		4
	Strenges Experiment	Quasi-Experiment	Ex post facto Untersuchung	Korrelative Untersuchung	Vorexperimentelle Untersuchung
I Kausaltheoretische Hypothese vor Versuchsbeginn vorhanden und hinreichend begründet	+	+	(+)	-	-
II Experimentelle Variablen manipulierbar bzw. manipuliert	+	+	-	-	-
III Alle übrigen Versuchsbedingungen kontrollierbar bzw. kontrolliert	+	-	-	-	-

**Abbildung 5.2: Allgemeines Gliederungsschema für eine Systematik der Versuchspläne nach den vier allgemeinen Designtypen und deren Bewertung nach verschiedenen Güte Merkmalen (I, II, III).** (Modifiziert nach Boesch & Eckensberger, 1969)

---

Überblick

**Interne und externe Validität**

Selbstkontrollfragen

## Allgemeine Gütekriterien empirischer Untersuchungen\*

### Objektivität

- Ausmaß der Unabhängigkeit des Untersuchungsergebnisses vom Untersuchenden

### Reliabilität

- Reproduzierbarkeit des Untersuchungsergebnisses bei Wiederholung

### Validität

- Korrespondenz zwischen dem beabsichtigten und dem tatsächlichen Untersuchten
- Genauigkeit einer Untersuchung, das zu Messen, was sie messen soll
- Ausmaß der Angemessenheit von Schlussfolgerungen aus der Untersuchung

\* Psychodiagnostische Tests, Experimente, Studien, ...

## Validitätsarten

### Interne Validität

- Eine Untersuchung ist **intern valide**, wenn Veränderungen in den abhängigen Variablen eindeutig auf den Einfluss der unabhängigen Variablen zurückzuführen sind (Campbell and Stanley 1963). Interne Validität liegt also dann vor, wenn das Ergebnis einer Untersuchung eindeutig interpretierbar ist. Die interne Validität sinkt mit der Anzahl plausibler Alternativerklärungen für das erhaltene Ergebnis (Reiß and Sarris 2012). Die interne Validität ist eng verwandt mit der **Konstruktvalidität**, die die Güte der Operationalisierung von UV und AV hinsichtlich ihrer zugrundeliegenden theoretischen Konzeption beschreibt.

### Externe Validität

- Eine Untersuchung ist **extern valide**, wenn das in einer Stichprobenuntersuchung gefundene Ergebnis auf andere Personen, Situationen, oder Zeitpunkte generalisiert werden kann. Man spricht auch von **inferenzstatistischer Validität**.

## Mögliche Konfundierungen der internen Validität

Tabelle 5.3

**Neun typische Faktoren, welche die interne Validität von experimentellen Befunden im Sinne von Variablenkonfundierungen beeinträchtigen („Artefakte“).** (Nach Campbell & Stanley, 1966; vgl. Sarris, 1999)

### Validitätsbeeinträchtigender Störfaktor

### Kurze Charakterisierung des Artefakts

1. Zeitgeschehen  
(history)

Die beobachteten Effekte gehen nicht allein auf die experimentelle Bedingungskonstellation, sondern (zusätzlich) auf unkontrollierte *zwischenzeitliche* Ereignisse zurück.

2. Reifung  
(maturation)

Wenn sich der zu untersuchende Sachverhalt bezüglich biologischer (oder/und psychosozialer) Reifungsmerkmale verändert, ist mit reifungsbedingten Effekten zu rechnen, die den eigentlichen experimentellen Befund überlagern.

3. Mehrfache Testung  
(test sophistication)

Besonders bei mehrfacher Erhebung derselben Messdaten an ein und demselben Individuum können die während des zweiten (dritten, ...) Messzeitpunkts erhobenen Daten aufgrund vorangegangener Testung beeinflusst sein (z.B. sensibilisierende Erfahrung im Umgang mit einem Test).

4. 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).

## Mögliche Konfundierungen der internen Validität

5. Statistische Regression (regression)	Werden mehr oder weniger extrem verschiedene Leistungsgruppen z.B. mithilfe eines Vortests gebildet, dann kann die mangelnde Reliabilität (Testzuverlässigkeit) des Messinstruments zu einer statistischen „Regression“ zur Mitte bei der zweiten (experimentellen) Testung führen.
6. Auswahlverzerrung (selection)	Bei nicht-zufälliger Bildung von Versuchsgruppen können die damit von Anfang an bestehenden systematischen Ausgangsdifferenzen zwischen den Gruppen den eigentlichen experimentellen Effekt überlagern.
7. Ausfalleffekte (experimental mortality)	Fallen im Untersuchungsverlauf Probanden von verschiedenen Versuchsgruppen aus, so kann das die eigentlichen experimentellen Effekte beeinflussen, wenn die Ausfallquote für die Gruppen systematisch verschieden ist.
8. Versuchsleitereffekte (experimenter-bias effects, sog. Rosenthal-Effekte)	Bleiben die Eigenschaften, Verhaltensweisen oder/und Versuchserwartungen des Untersuchers unkontrolliert, kann das eine systematische Beeinträchtigung der eigentlichen experimentellen Befunde nach sich ziehen.
9. Interaktive Effekte (interactive effects; carry-over effects)	Wird ein Individuum unter verschiedenen experimentellen Bedingungen untersucht und bleiben dabei Übertragungseffekte unkontrolliert, können dadurch die experimentellen Befunde verfälscht werden.

## Mögliche Konfundierungen der externen Validität

### Stichprobeneffekte

- Effekte gelten nur für die spezifische vorliegende Stichprobe

### Treatmenteffekte

- Effekte gelten nur für die spezifische Variation der unabhängigen Variable

### Outcomeeffekte

- Effekte gelten nur für die spezifische Variation der abhängigen Variablen

### Kontexteffekte

- Effekte gelten nur für den spezifischen Studienkontext

## Laborexperiment

- Experiment in einer hochkontrollierbaren Umgebung
- Kontrolle von Störvariablen ideal möglich
- Generalisierung auf alltägliche Erfahrungswelt schwierig

⇒ Interne Validität eher hoch, externe Validität eher niedrig

## Feldexperiment

- Untersuchung in natürlicher Umgebung
- Kontrolle von Störvariablen nur eingeschränkt möglich
- Generalisierung auf alltägliche Erfahrungswelt eher möglich

⇒ Interne Validität eher niedrig, externe Validität eher hoch

---

Überblick

Interne und externe Validität

**Selbstkontrollfragen**

1. Definieren Sie die Begriffe der Objektivität, Reliabilität und Validität einer Untersuchung.
2. Definieren Sie die Begriffe interne Validität und externe Validität einer Untersuchung..
3. Nennen Sie vier mögliche Konfundierungen der internen Validität einer Untersuchung.
4. Nennen Sie vier mögliche Konfundierungen der externen Validität einer Untersuchung.
5. Diskutieren Sie Labor- und Feldexperimente im Kontext ihrer internen und externen Validität.

## References

---

- Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Belmont, CA: Wadsworth.
- Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (6) Experimentelle Versuchspläne I

Datum	Einheit	Thema
21.10.2021	Einführung	(1) Einführung
28.10.2021	Einführung	(2) Theorien, Hypothesen, Experimente
04.11.2021	Einführung	(3) Operationalisieren, Messen, Skalieren
11.11.2021	Studiendesign	(4) Prinzipien der Versuchsplanung
18.11.2021	Studiendesign	(5) Kriterien der Versuchsplanung
<b>25.11.2021</b>	<b>Studiendesign</b>	<b>(6) Experimentelle Versuchspläne I</b>
02.12.2021	Studiendesign	(7) Experimentelle Versuchspläne II
09.12.2021	Studiendesign	(8) Experimentelle Versuchspläne III
16.12.2021	Studiendesign	(9) Experimentelle Versuchspläne IV
	Weihnachtspause	
06.01.2022	Auswahlverfahren	(10) Grundgesamtheit und Stichprobe
13.01.2022	Auswahlverfahren	(11) Stichprobentechniken
20.01.2022	Datenerhebung	(12) Qualitative Verfahren
27.01.2022	Datenerhebung	(13) Quantitative Verfahren
30.03.2022	Klausurtermin	12:00 - 13:00 Uhr G16-H5
Jul 2022	Klausurwiederholungstermin	

---

Überblick

Randomisierte einfaktorische Versuchspläne

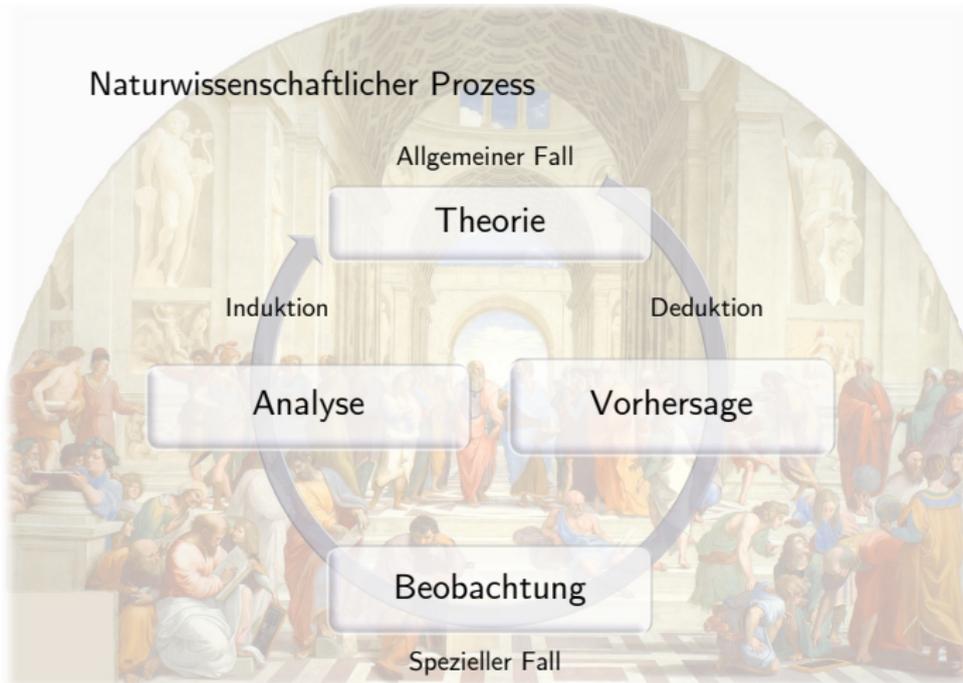
Selbstkontrollfragen

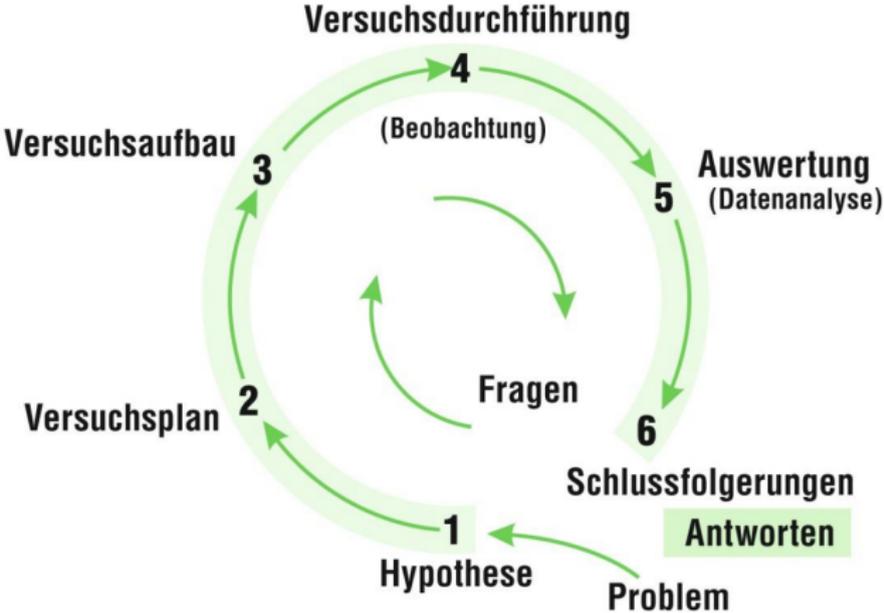
---

## Überblick

Randomisierte einfaktorielle Versuchspläne

Selbstkontrollfragen





## Definition

Unter einem Versuchsplan versteht man ein standardisiertes, routinemäßig anwendbares Strukturschema, das dem Aufbau, der Kontrolle und der methodologischen Bewertung einer empirischen Untersuchung von unabhängigen und abhängigen Variablen sachlogisch zugrunde liegt.

Reiß and Sarris (2012)

⇒ Überblick über häufig verwendete Versuchspläne

## Faktorielle Versuchspläne

- Kategoriale unabhängige Variable
- Die Werte der unabhängigen Variablen werden Level genannt
- Einfaktorielle oder mehrfaktoriell

## Parametrische Versuchspläne

- Kontinuierliche unabhängige Variable
- Die Werte der unabhängigen Variablen werden oft Level genannt
- Meist einfaktoriell

## Between-Group Designs | Versuchspläne mit Randomisierung

- Gesamtgruppe wird zufällig auf experimentelle Bedingungen aufgeteilt
- Einfaktoriell, mehrfaktoriell
- Häufig in der klinischen Forschung verwendet

## Within-Group Designs | Versuchspläne mit Wiederholungsmessung

- Eine Gesamtgruppe wird sämtlichen experimentellen Bedingungen unterzogen
- Einfaktoriell, mehrfaktoriell, parametrisch
- Häufig in der psychologischen Grundlagenforschung verwendet

## Block Designs | Versuchspläne mit parallelisierten Gruppen

- Gesamtgruppe wird gesteuert auf experimentelle Bedingungen aufgeteilt
- Experimentelle Gruppen werden anhand bestimmter Merkmale parallelisiert
- Häufiger in der klinischen Forschung verwendet

## Mixed Designs | Mischdesigns

- Mischungen aller obiger Versuchsplantypen

## Randomisierung bei Between-Group Designs

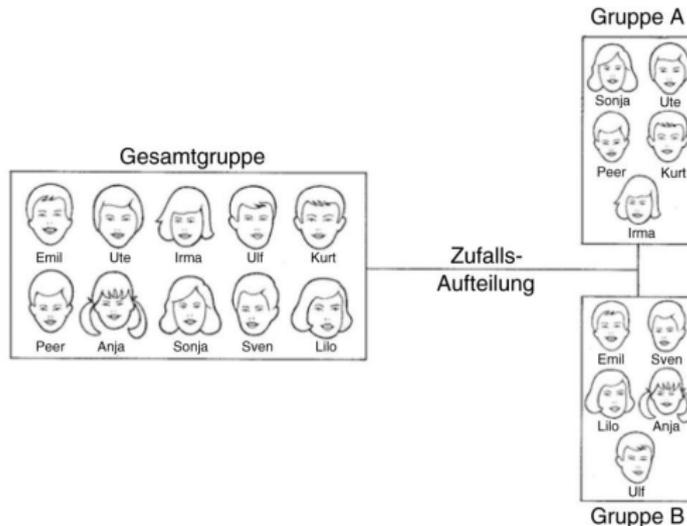


Abbildung 5.1: Zufallsaufteilung einer Gesamtgruppe von  $N=10$  Probanden (links) in zwei gleiche große Untergruppen (rechts). (Aus Sarris, 1999)

## Designschemata

- R: Randomisierung
- O: Observation (Test, Messung)
- X: Exposition experimenteller Bedingung
- Experimentelle Bedingungen von oben nach unten
- Zeitliche Abfolge von links nach rechts

## Beispiel

R	X	O
R		O

- Bedingungszuweisung erfolgt durch Randomisierung
- Nur eine Gruppe erhält das Treatment
- Beide Gruppen absolvieren die Messung

---

Überblick

## **Randomisierte einfaktorielle Versuchspläne**

Selbstkontrollfragen

## Randomisierte einfaktorielle Versuchspläne

- Gesamtgruppe wird zufällig auf experimentelle Bedingungen aufgeteilt
- Eine unabhängige Variable mit zwei oder mehr Leveln
- Populäres Designs in der klinischen Forschung
- Varianten
  - o Treatment- und Kontrollgruppe
  - o Treatment- und Placebogruppe
  - o Zwei Treatmentgruppen (und Kontrollgruppe)
  - o Pretest-Posttest Designs

## No-Treatment Kontrollgruppe

R	X	O
R		O

- Vergleich eines Treatments zu keinem Treatment

## Placebo Kontrollgruppe

R	X	O
R	X <sub>P</sub>	O

- Placebo = Scheintreatment
- Vergleich eines Treatments zu keinem Treatment
- Kontrolle studieninduzierter Effekte (Placeboeffekte)

## Vergleich zweier Treatments

R	$X_A$	O
R	$X_B$	O

- Vergleich Standardtreatment A und neues Treatment B
- Keine Aussage über Effektivität des Standardtreatments

## Zwei-Treatment Vergleich mit Placebo-Kontrollgruppe

R	$X_A$	O
R	$X_B$	O
R	$X_P$	O

- Vergleich Standardtreatment A und neues Treatment B
- Aussage über Effektivität des Standardtreatments möglich
- Placebotreatment kann ethisch nicht vertretbar sein

Beispiel: Einfluss von Psychotherapie auf Depressionssymptomatik

- Klassische Psychotherapie (A)
- Online Psychotherapie (B)
- Seelsorge (P)

→ Keine Aussagen über Pre-Treatment Gruppenunterschiede möglich

→ Keine Aussage über Dropout Charakteristika möglich

## Pre-Posttest Designs

R	O	$X_A$	O
R	O	$X_B$	O
R	O		O

- Fokus auf Treatment-induzierte Verbesserungen/Verschlechterungen
- Subtraktion von Pre-Test-Gruppenunterschieden möglich
- Untersuchung von Dropout Charakteristika möglich
- Mögliches Auftreten von Testeffekten (Lernen, Gewöhnung, Ermüdung)
- Höherer Zeit- und Kostenaufwand

# Randomisierte einfaktorielle Versuchspläne

## Beispiel: Evaluation von Psychotherapieformen bei Depression

Experimentelle Bedingung  
(Gruppen von  $n = 12$ )

Psychotherapie

Klassisch

Prä-BDI



Post-BDI

Online

Prä-BDI



Post-BDI

⇒ Randomisiertes einfaktorielles Preposttest Design ohne Kontrollgruppe

---

Überblick

Randomisierte einfaktorische Versuchspläne

**Selbstkontrollfragen**

1. Erläutern Sie die Begriffe “Kategorialer Versuchsplan” und “Parametrischer Versuchsplan”.
2. Erläutern Sie die Begriffe “Between-Group Design” und “Within-Group Design”.
3. Erläutern Sie den Begriff “Block Design”.
4. Nennen Sie drei Varianten von randomisierten einfaktoriellen Versuchsplänen.
5. Erläutern Sie die Bedeutung einer Placebo-Kontrollgruppe.
6. Nennen Sie Vor- und Nachteile von Zwei-Treatment Vergleichsdesigns ohne Pre-Tests.
7. Nennen Sie Vor- und Nachteile von Zwei-Treatment Vergleichsdesigns mit Pre- und Post-Tests.

## References

---

Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (7) Experimentelle Versuchspläne II

---

Randomisierte mehrfaktorielle Versuchspläne

Selbstkontrollfragen

---

## **Randomisierte mehrfaktorielle Versuchspläne**

Selbstkontrollfragen

## Mehrfaktorielle Versuchspläne

- Kombination mehrerer experimenteller Faktoren in einem Versuchsplan

## Crossed Design

- Jedes Level jedes Faktors wird mit allen Leveln aller Faktoren kombiniert (gekreuzt).

## Nested Design

- Einige Level eines Faktors werden nicht mit allen anderen Faktorleveln kombiniert.

⇒ Prototypisch sind zweifaktorielle Versuchspläne mit crossed design

## Randomisierter zweifaktorieller Versuchspläne mit crossed design

- Eine univariate abhängige Variable bestimmt an individuellen experimentellen Einheiten.
- Zwei diskrete unabhängige Variablen, die mindestens zweistufig sind.
- Die unabhängigen Variablen werden *Faktoren* genannt.
- Die Stufen der Faktoren werden *Level* genannt.
- Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert

## Zweifaktorielle Versuchspläne werden üblicherweise anhand ihrer Faktorlevel bezeichnet

2 × 2 Design: Faktor A mit Level 1,2	Faktor B mit Level 1,2
2 × 3 Design: Faktor A mit Level 1,2	Faktor B mit Level 1,2,3
4 × 2 Design: Faktor A mit Level 1,2,3,4	Faktor B mit Level 1,2
3 × 1 Design: Faktor A mit Level 1,2,3	Faktor B mit Level 1

- 2 × 2 Versuchspläne sind sehr populär, wir fokussieren auf diesen Fall.

⇒ Das entsprechende datenanalytische Verfahren ist die Varianzanalyse (ANOVA)

## Konzeptuelles Design eines 2 x 2 Versuchsplans

		Faktor B	
		Level 1	Level 2
Faktor A	Level 1	A1B1	A1B2
	Level 2	A2B1	A2B2

# Randomisierte mehrfaktorielle Versuchspläne

## Anwendungsbeispiel eines randomisierten 2 x 2 Versuchsplans

- Ist Psychotherapie bei Depression im klassischen oder im online Setting wirksamer?
- Ist Psychoanalyse oder Kognitive-Verhaltenstherapie bei Depression wirksamer?

	Setting	
	Klassisch	Online
Psychoanalyse		
Verfahren		
Verhaltenstherapie		

- Ist die Wirksamkeit der Psychoanalyse vom Setting abhängig?
- Ist die Wirksamkeit der Kognitiven-Verhaltenstherapie vom Setting abhängig?

## Verlaufsschema eines randomisierten $2 \times 2$ Versuchsplans

R	$X_{A1B1}$	O
R	$X_{A1B2}$	O
R	$X_{A2B1}$	O
R	$X_{A2B2}$	O

- Pre-Posttest Designs möglich
- Placebo Kontrollgruppen möglich

# Randomisierte mehrfaktorielle Versuchspläne

## Datennotation in 2 x 2 Versuchsplänen

		Faktor B	
		Level 1	Level 2
Faktor A	Level 1	$x_{111} = 1.2$ $x_{112} = 0.4$ $x_{113} = 1.7$ $\vdots$ $x_{11m} = 2.1$	$x_{121} = 4.1$ $x_{122} = 1.8$ $x_{123} = 3.3$ $\vdots$ $x_{12m} = 5.9$
	Level 2	$x_{211} = 0.1$ $x_{212} = 0.8$ $x_{213} = 2.7$ $\vdots$ $x_{21m} = 1.4$	$x_{221} = 7.4$ $x_{222} = 6.2$ $x_{223} = 9.5$ $\vdots$ $x_{22m} = 6.1$

$x_{ijk}$  bezeichnet den Datenpunkt der  $k$ ten experimentellen Einheit ( $k = 1, \dots, m$ )  
im  $i$ ten Level von Faktor A und  $j$ ten Level von Faktor B ( $i = 1, 2, j = 1, 2$ )

## Tabellarische Form von 2 x 2 Versuchsplandaten

	$x_{11k}$	$x_{12k}$	$x_{21k}$	$x_{22k}$
1	2.86	5.04	3.29	6.88
2	2.62	4.96	3.69	7.33
3	3.01	4.58	2.94	6.76
4	3.20	5.05	3.61	6.57
5	2.75	5.05	3.68	6.84
6	2.88	5.01	3.89	7.45
7	2.85	4.35	3.43	7.06
8	3.23	5.07	3.72	6.99
9	2.99	5.84	3.28	6.93
10	2.57	4.61	3.35	7.46
11	2.69	5.26	2.94	7.07
12	2.83	5.17	3.56	7.03

$x_{ijk}$  bezeichnet den Datenpunkt der  $k$ ten experimentellen Einheit ( $k = 1, \dots, m$ )  
im  $i$ ten Level von Faktor A und  $j$ ten Level von Faktor B ( $i = 1, 2, j = 1, 2$ )

## Haupteffekte und Interaktionen

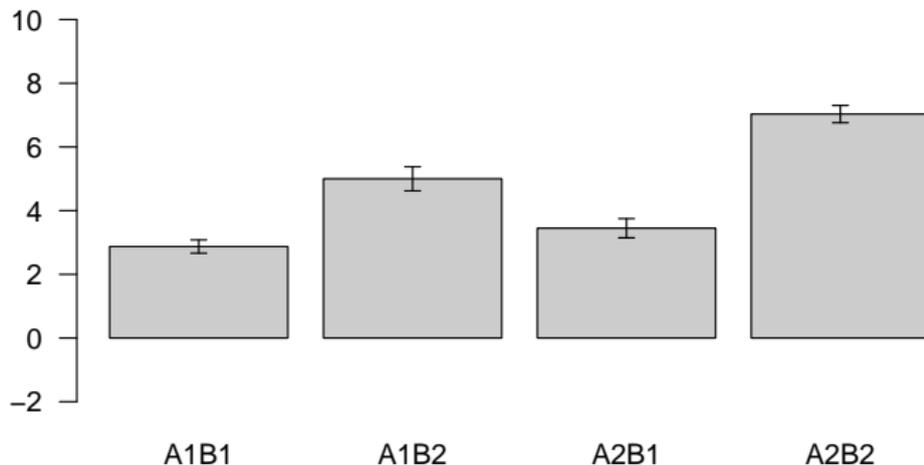
Hinsichtlich der Gruppenmittelwerte bei  $2 \times 2$  Versuchsplänen unterscheidet man **Haupteffekte** und **Interaktionen**

- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor A*, wenn sich die Gruppenmittelwerte zwischen Level 1 und Level 2 von Faktor A, jeweils gemittelt über die zwei Level von Faktor B, unterscheiden.
- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor B*, wenn sich die Gruppenmittelwerte zwischen Level 1 und Level 2 von Faktor B, jeweils gemittelt über die zwei Level von Faktor A, unterscheiden.
- Intuitiv spricht man vom Vorliegen einer *Interaktion der Faktoren A und B*, wenn der Unterschied der Gruppenmittelwerte von Faktor A zwischen Level 1 und 2 unterschiedlich für Level 1 und Level 2 von Faktor B ausgeprägt ist bzw. wenn der Unterschied der Gruppenmittelwerte von Faktor B zwischen Level 1 und 2 unterschiedlich für Level 1 und Level 2 von Faktor A ausgeprägt ist.

Intuitiv beziehen sich Haupteffekte also auf (marginale) Unterschiede (Differenzen), während sich Interaktionen auf Unterschiede von Unterschieden (Differenzen von Differenzen) beziehen.

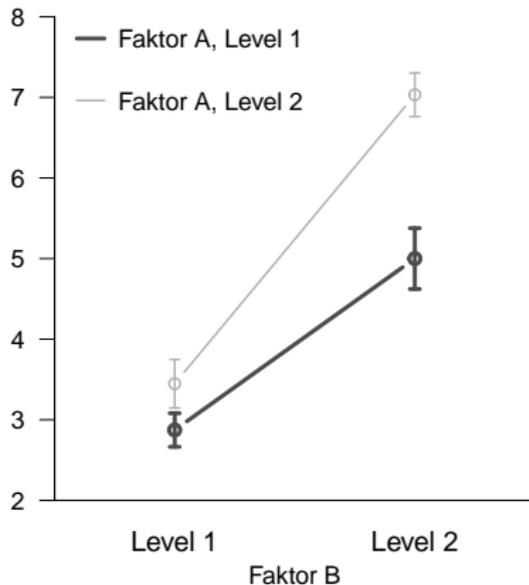
Das Vorhandensein einer Interaktion besagt lediglich, dass sich die Unterschiede der Gruppenmittelwerte zwischen den Leveln eines experimentellen Faktors in Abhängigkeit von den Leveln des anderen experimentellen Faktors ändern, es macht aber keine Aussage darüber, warum dies so ist. Haupteffekte und Interaktionen sind lediglich Datenmuster, keine mechanistischen wissenschaftlichen Theorien.

## Darstellung von Gruppenmittelwerten durch Balkendiagramm



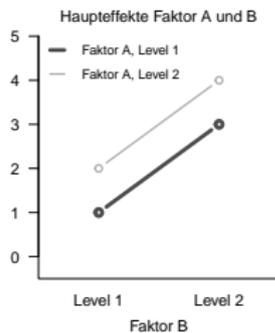
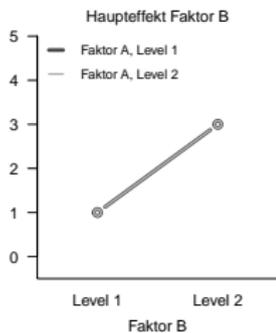
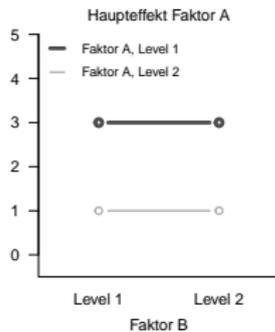
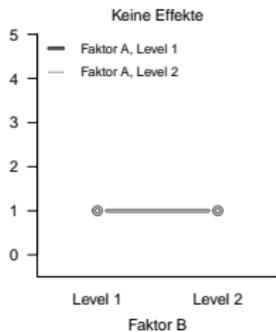
# Randomisierte mehrfaktorielle Versuchspläne

## Darstellung von Gruppenmittelwerten durch Liniendiagramm



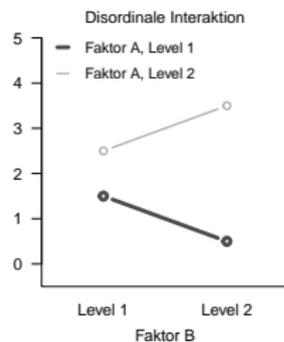
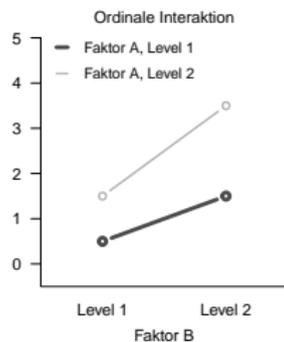
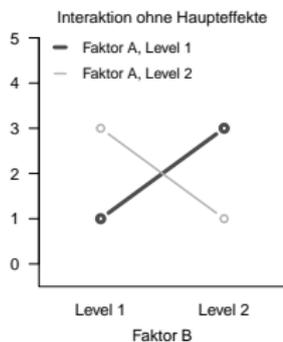
# Randomisierte mehrfaktorielle Versuchspläne

## Haupteffektmuster bei Abwesenheit von Interaktionen ( $\Rightarrow$ Parallele Linien)



# Randomisierte mehrfaktorielle Versuchspläne

## Haupteffekte und Interaktionen ( $\Rightarrow$ Nicht-Parallele Linien)



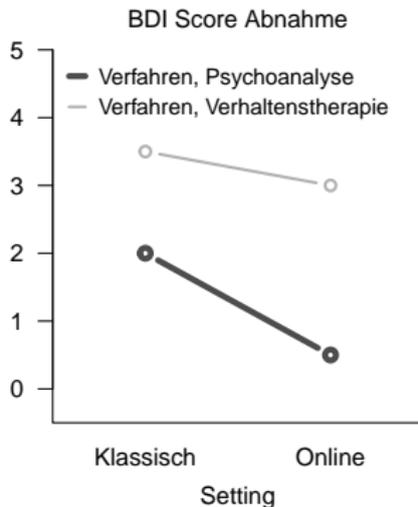
### Ordinale Interaktion

- Gleiches Steigungsvorzeichen für beide Faktoren

### Disordinale Interaktion

- Unterschiedliche Steigungsvorzeichen der Faktoren

## Hypothetisches Ergebnis des Anwendungsbeispiels



Haupteffekt Verfahren  $\Rightarrow$  Verhaltenstherapie  $>$  Psychoanalyse

Haupteffekt Setting  $\Rightarrow$  Klassisch  $>$  Online

Interaktion  $\Rightarrow$  Wirksamkeitsminderung durch Online Setting stärker bei Psychoanalyse

---

Randomisierte mehrfaktorielle Versuchspläne

**Selbstkontrollfragen**

1. Erläutern Sie die Begriffe Mehrfaktorieller Versuchsplan, Crossed design und Nested design.
2. Erläutern Sie den Begriff des randomisiertern zweifaktoriellen Versuchsplans mit Crossed design.
3. Erläutern Sie die Begriffe Haupteffekt und Interaktion am Beispiel eines  $2 \times 2$  Versuchsplans.
4. Skizzieren Sie Liniendiagramme von  $2 \times 2$  Versuchsplänen mit
  - o einem Haupteffekt und keiner Interaktion,
  - o Haupteffekten und keiner Interaktion,
  - o Haupteffekten und ordinaler Interaktion,
  - o Haupteffekten und disordinaler Interaktion.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (8) Experimentelle Versuchspläne III

---

Block Designs

Quasi-Experimentelle Designs

Ex-Post-Facto Designs

Selbstkontrollfragen

---

## **Block Designs**

Quasi-Experimentelle Designs

Ex-Post-Facto Designs

Selbstkontrollfragen

## Ziel und Grundidee

- Reduktion der Sekundärvarianz durch Differenzbildung bei unterschiedlichen Referenzwerten

## Block Designs | Versuchspläne mit parallelisierten Gruppen

- Gesamtgruppe wird gesteuert auf experimentelle Bedingungen aufgeteilt
- Experimentelle Gruppen werden anhand bestimmter Merkmale parallelisiert
- Häufiger in der klinischen Forschung verwendet

## Repeated Measures Designs | Versuchspläne mit Wiederholungsmessungen

- Ähnliche Grundidee wie bei Block Designs, aber anderes Vorgehen
- Within-group Designs mit Messwiederholungen
- Häufiger in der experimentalpsychologischen Grundlagenforschung verwendet

## Blockbildung durch Parallelisierung

- (1) Erfassung von relevanten Merkmalsausprägungen in einem Vortest.
- (2) Erstellen einer Rangreihe.
- (3) Bilden von Blöcken von experimentellen Einheiten mit benachbarten Rängen.
- (4) Randomisierte blockweise Zuordnung zu den experimentellen Bedingungen

## Einfluss der Digitalaffinität beim Vergleich von Klassischer vs. Online Psychotherapie

Keine Blockbildung I - Online Bedingung dominiert von digitalaffinen Proband:innen

Proband:in	Digitalaffinität	Bedingung	Therapieeffekt	Digitalaffinität x Bedingung	BDI.Verbesserung
1	-4	Klassisch	2	0	2
2	-4	Klassisch	3	0	3
3	-4	Klassisch	3	0	3
4	-4	Klassisch	2	0	2
5	-2	Klassisch	3	0	3
6	-2	Klassisch	3	0	3
7	-2	Klassisch	3	0	3
8	-2	Klassisch	4	0	4
9	2	Online	2	2	4
10	2	Online	4	2	6
11	2	Online	2	2	4
12	2	Online	2	2	4
13	4	Online	2	4	6
14	4	Online	3	4	7
15	4	Online	3	4	7
16	4	Online	3	4	7

> Mittelwert Therapieeffekt Klassisch: 2.88 , Mittelwert BDI Verbesserung Klassisch: 2.88

> Mittelwert Therapieeffekt Online : 2.62 , Mittelwert BDI Verbesserung Online : 5.62

## Einfluss der Digitalaffinität beim Vergleich von Klassischer vs. Online Psychotherapie

Keine Blockbildung II - Online Bedingung dominiert von nicht digitalaffinen Proband:innen

Proband:in	Digitalaffinität	Bedingung	Therapieeffekt	Digitalaffinität x Bedingung	BDI.Verbesserung
1	4	Klassisch	2	0	2
2	4	Klassisch	3	0	3
3	4	Klassisch	3	0	3
4	4	Klassisch	2	0	2
5	2	Klassisch	3	0	3
6	2	Klassisch	3	0	3
7	2	Klassisch	3	0	3
8	2	Klassisch	4	0	4
9	-2	Online	2	-2	0
10	-2	Online	4	-2	2
11	-2	Online	2	-2	0
12	-2	Online	2	-2	0
13	-4	Online	2	-4	-2
14	-4	Online	3	-4	-1
15	-4	Online	3	-4	-1
16	-4	Online	3	-4	-1

> Mittelwert Therapieeffekt Klassisch: 2.88 , Mittelwert BDI Verbesserung Klassisch: 2.88

> Mittelwert Therapieeffekt Online : 2.62 , Mittelwert BDI Verbesserung Online : -0.375

## Einfluss der Digitalaffinität beim Vergleich von Klassischer vs. Online Psychotherapie

Blockbildung - Gleiche Verteilung von digitalaffinen und nicht-digitalaffinen Proband:innen über Bedingungen

Proband:in	Digitalaffinität	Bedingung	Therapieeffekt	Digitalaffinität x Bedingung	BDI-Verbesserung
1	4	Klassisch	2	0	2
2	4	Klassisch	3	0	3
3	-4	Klassisch	3	0	3
4	-4	Klassisch	2	0	2
5	2	Klassisch	3	0	3
6	2	Klassisch	3	0	3
7	-2	Klassisch	3	0	3
8	-2	Klassisch	4	0	4
9	4	Online	2	4	6
10	4	Online	4	4	8
11	-4	Online	2	-4	-2
12	-4	Online	2	-4	-2
13	2	Online	2	2	4
14	2	Online	3	2	5
15	-2	Online	3	-2	1
16	-2	Online	3	-2	1

> Mittelwert Therapieeffekt Klassisch: 2.88 , Mittelwert BDI Verbesserung Klassisch: 2.88

> Mittelwert Therapieeffekt Online : 2.62 , Mittelwert BDI Verbesserung Online : 2.62

---

Block Designs

## **Quasi-Experimentelle Designs**

Ex-Post-Facto Designs

Selbstkontrollfragen

## Definition

Untersuchung, bei der auf eine Randomisierung verzichtet werden muss, weil natürliche bzw. bereits bestehende Gruppen untersucht werden.

Reiß and Sarris (2012)

## Beispiele

- Untersuchung des Effektes von Psychotherapie bei Frauen vs. Männern
- Untersuchung des Effektes von Psychotherapie bei Jugendlichen vs. Erwachsenen
- Untersuchung des Effektes von Psychotherapie bei Depression vs. Schizophrenie

## Techniken zur Kontrolle personenbezogener Störvariablen

- Block Designs
- Propensity-Score Matching

## Propensity Score Matching

“Paarweise Zuordnung auf Basis von Neigungsscores” nach Rosenbaum and Rubin (1983)

### Ziel

- Proband:innen sollten bis auf das Gruppenmerkmal möglichst “gleich” sein.

### Ansatz

- Erfassung einer großen Anzahl von Kovariaten (Geschlecht, Alter, Komorbiditäten, ...).
- Bilden von Proband:innenpaaren, die für die Kovariaten möglichst gleich sind.
- ⇔ Bilden von “statistische Zwillingen” über Gruppen.

### Problem

- Paare ähnlicher Werte auf allen Kovariaten zu finden kann schwer sein.
- Je mehr Kovariaten, desto schwieriger wird ein exaktes Matching (Curse of Dimensionality).

### Lösung

- Hochdimensionale Kovariatenvektoren werden auf eindimensionalen Propensity Score projiziert.
- Diese Projektion geschieht meist mithilfe der logistische Regression.
- Propensity Scores werden als Proxy für die Kovariaten gematched.

---

Block Designs

Quasi-Experimentelle Designs

**Ex-Post-Facto Designs**

Selbstkontrollfragen

## Definition

Bei der Ex-Post-Facto Versuchsanordnung liegt keine Manipulation seitens der Untersuchenden vor, es wird von korrelativen Datensätzen lediglich auf eine oder mehrere hypothetische “Verursachungen” zurückgeschlossen.

Reiß and Sarris (2012)

Keine Experimente im Sinne der Wundtschen Definition eines Experiments.

Übergang zu korrelativen Designs fließend.

Aus ethischen Erwägungen oftmals nötig, aber notwendigerweise weniger kontrolliert.

Beispiel: Einfluss des Rauchens auf Lungenkrebsentwicklung

- Experimentelles Design mit Bedingungen “Nichtrauchen” vs. “Rauchen” nicht möglich.
- Analyse von Krebsentwicklung bei Rauchern vs. Nichtrauchern muss Störvariablen einbeziehen.
- Man denke zum Beispiel an allgemeine Lebensweise, sozioökonomischer Standard usw.

---

Block Designs

Quasi-Experimentelle Designs

Ex-Post-Facto Designs

**Selbstkontrollfragen**

1. Erläutern Sie die grundlegende Idee von Block Designs.
2. Erläutern Sie das Verfahren zur Blockbildung durch Parallellisierung.
3. Definieren Sie den Begriff des quasi-experimentellen Designs nach Reiß und Sarris (2012)
4. Nennen Sie drei Beispiele für Studien mit quasi-experimentellen Designs.
5. Erläutern Sie die Methode des Propensity Score Matchings bei quasi-experimentellen Designs.
6. Definieren Sie den Begriff des Ex-Post-Facto Designs nach Reiß und Sarris (2012)

## References

---

- Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.
- Rosenbaum, Paul R, and Donald B Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," 15.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (9) Experimentelle Versuchspläne IV

---

Korrelative Designs

Korrelationskoeffizient

Stichprobenauswahleffekte

Selbstkontrollfragen

---

## **Korrelative Designs**

Korrelationskoeffizient

Stichprobenauswahleffekte

Selbstkontrollfragen

## Definition

Nicht-experimentelle Datenanordnungen, die typischerweise nur den wechselseitigen Zusammenhang (Korrelation) zwischen zwei oder mehreren Variablen betreffen.

Reiß and Sarris (2012)

## Bemerkungen

- “Correlation is not causation!”

⇒ Korrelationen können immer durch Drittvariablen kausal bedingt sein.

Ja, aber ...

- Der Begriff der “Kausalität” ist nicht eindeutig definiert.
- Experimentelle Designs werden im Normalfall mit Korrelationen (Regression, ALM) untersucht.
- Kausale Inferenz nutzt auch “nur” probabilistische Modelle, wird wenig angewendet/gelehrt.
- Im Sinne der zeitlichen Präzedenz benutzt kausale Inferenz zum Teil Zeitserienmodelle.
- Das Psychologiestudium sieht keine Auseinandersetzung mit Graphical Models und Differentialgleichungsmodellen vor, die für ein Verständnis zeitgenössischer kausaler Inferenz nötig wäre.

Kausalzusammenhänge sind ein latentes Konstrukt das nur datenanalytisch erschlossen werden kann!

⇒ Kausalzusammenhänge sind sowohl in experimentellen als auch korrelativen Designs latent!

---

Korrelative Designs

**Korrelationskoeffizient**

Stichprobenauswahleffekte

Selbstkontrollfragen

# Korrelationskoeffizient

---

Wir bezeichnen einen bivariaten Datensatz mit

$$x = ((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)). \quad (1)$$

$x_i$  und  $y_i$  bezeichnen das erste und zweite Merkmal der  $i$ ten Einheit, respektive.

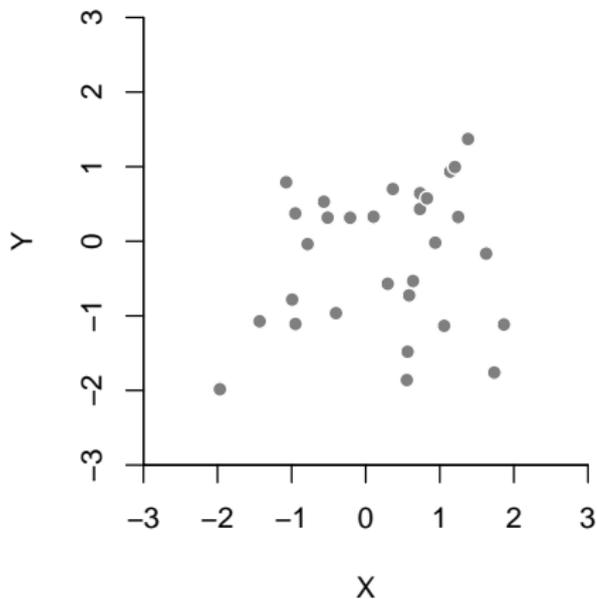
$n$  ist die Anzahl an bivariaten Datenpunkten  $(x_i, y_i)$ .

Untenstehende Tabelle zeigt ein Beispiel.

$i$	1	2	3	4	5	6	7	8	9	10
$x_i$	3.4	1.5	2.7	4.5	6.1	3.8	2.0	2.5	6.2	9.1
$y_i$	5.5	7.3	1.1	1.9	4.5	2.3	8.4	8.6	3.9	1.6

# Korrelationskoeffizient

Die Darstellung der  $(x_i, y_i), i = 1, \dots, n$  in einem Koordinatensystem heißt *Streudiagramm*.



## Definition (Empirische Stichprobenkovarianz)

$x = ((x_1, y_1), \dots, (x_n, y_n))$  sei ein bivariater Datensatz. Dann heißt die Zahl

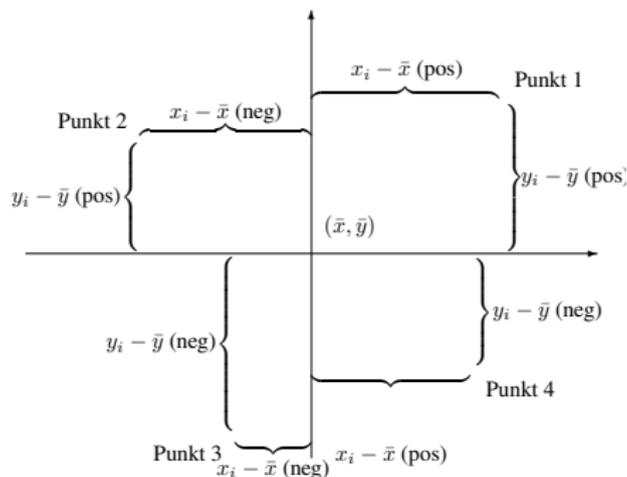
$$C := \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (2)$$

*empirische Stichprobenkovarianz* von  $x_1, \dots, x_n$  und  $y_1, \dots, y_n$ .

### Bemerkungen

- $\bar{x}$  und  $\bar{y}$  bezeichnen die Mittelwerte der  $x_1, \dots, x_n$  und  $y_1, \dots, y_n$ , respektive.
- Der Faktor  $1/n$  normiert für die Stichprobengröße.
- Eine Intuition vermittelt untenstehende Abbildung aus Fahrmeir et al. (2016).

# Korrelationskoeffizient



	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
Punkt 1 (1. Quadrant)	positiv	positiv	positiv
Punkt 2 (2. Quadrant)	negativ	positiv	negativ
Punkt 3 (3. Quadrant)	negativ	negativ	positiv
Punkt 4 (4. Quadrant)	positiv	negativ	negativ

Abbildung 3.10 und Tabelle 3.9 aus Fahrmeir et al. (2016), Kapitel 3.4.

## Definition (Pearson's Stichprobenkorrelationskoeffizient)

$x = ((x_1, y_1), \dots, (x_n, y_n))$  sei ein bivariater Datensatz,  $C$  sei die empirische Stichprobenvarianz von  $x_1, \dots, x_n$  und  $y_1, \dots, y_n$  und  $S_x$  bzw.  $S_y$  seien die empirischen Stichprobenstandardabweichungen von  $x_1, \dots, x_n$  bzw.  $y_1, \dots, y_n$ . Dann heißt die Zahl

$$r := \frac{C}{S_x S_y} \quad (3)$$

*Pearson's Stichprobenkorrelationskoeffizient* oder *empirischer Korrelationskoeffizient*.

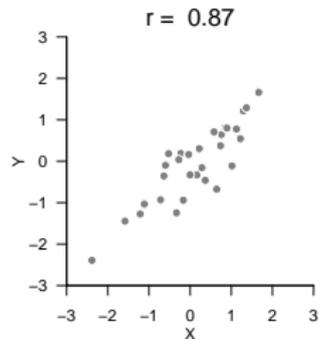
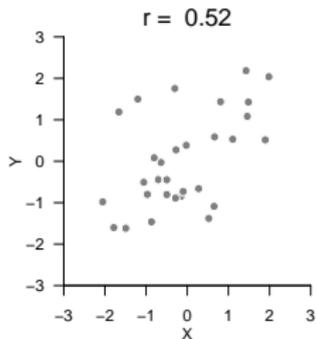
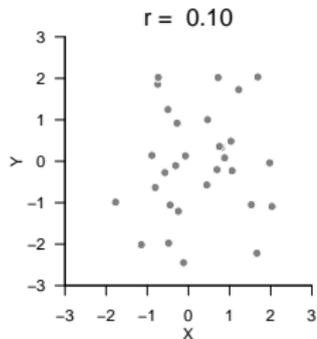
### Bemerkungen

- Ohne Beweis halten wir fest, dass gilt  $-1 \leq r \leq 1$ .
- Ohne Beweis halten wir fest, dass  $r$  die Stärke des linearen Zusammenhangs von  $x$  und  $y$  misst.
- Korrelationsstärken werden in etwa nach

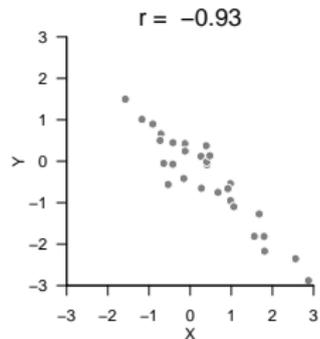
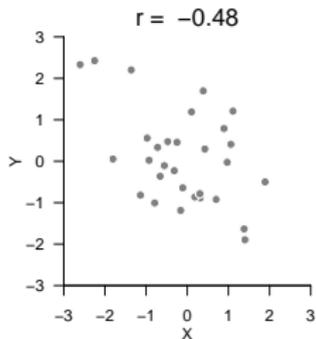
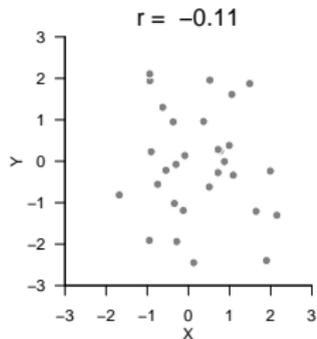
$$|r| < 0.5, 0.5 \leq |r| \leq 0.8, |r| > 0.8 \quad (4)$$

als "schwache," "mittlere," oder "starke" Korrelation bezeichnet

## Korrelationsstärken und Streudiagramme



## Korrelationsstärken und Streudiagramme



---

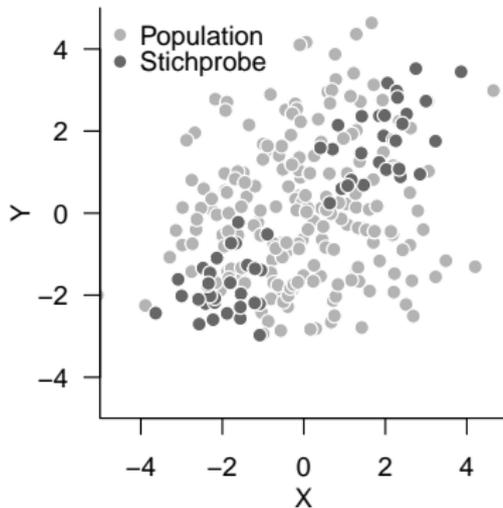
Korrelative Designs

Korrelationskoeffizient

**Stichprobenauswahleffekte**

Selbstkontrollfragen

## Stichprobe aus Extremgruppen

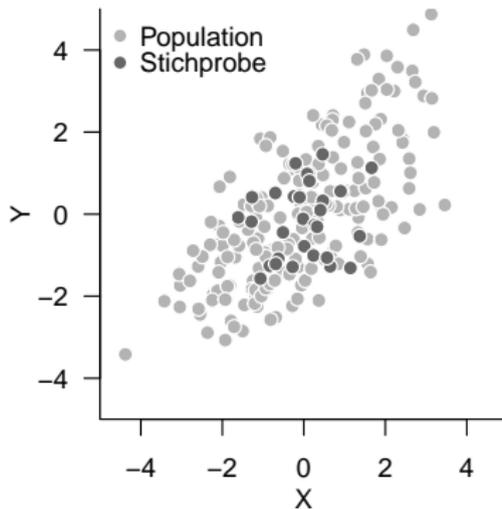


> Korrelation basierend auf Gesamtpopulationsdaten = 0.357

> Korrelation basierend auf Stichprobendaten = 0.919

# Stichprobenauswahleffekte

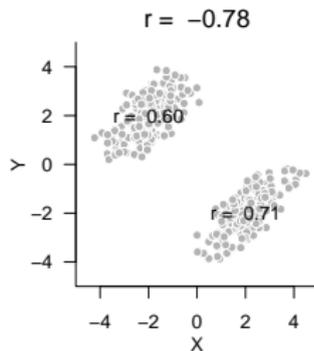
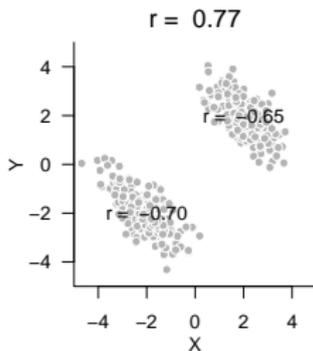
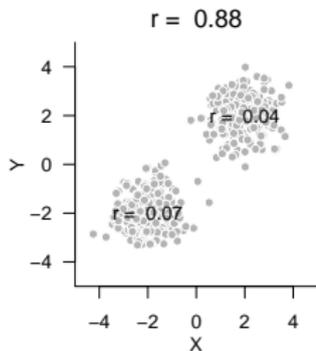
## Stichprobe mit zu kleiner Streubreite



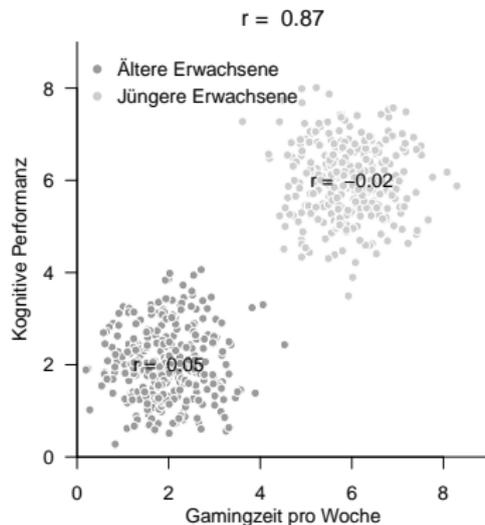
> Korrelation basierend auf Gesamtpopulationsdaten = 0.685

> Korrelation basierend auf Stichprobendaten = 0.186

## Bedingte vs. unbedingte Korrelationen (Simpson's Paradox)



## Bedingte vs. unbedingte Korrelationen (Simpson's Paradox)



⇒ Durch Drittvariable Alter (Jung, Alt) induzierte Korrelation!

---

Korrelative Designs

Korrelationskoeffizient

Stichprobenauswahleffekte

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Definieren Sie den Begriff des korrelativen Designs nach Reiß und Sarris (2012).
2. Definieren Sie die Empirische Stichprobenkovarianz.
3. Definieren Sie Pearson's Stichprobenkorrelationskoeffizient.
4. Beschreiben Sie Streudiagramme bivariater Datensätze bei schwachen und stark positiven bzw. negativen Korrelationen.
5. Erläutern Sie die Implikationen einer Stichprobenauswahl aus Extremgruppen.
6. Erläutern Sie die Implikationen einer Stichprobe mit zu kleiner Streubreite.
7. Erläutern Sie die Begriffe der bedingten und der unbedingten Korrelationen.
8. Diskutieren Sie ein Beispiel für eine durch eine Drittvariable induzierte Korrelation.

## References

---

- Fahrmeir, Ludwig, Christian Heumann, Rita Künstler, Iris Pigeot, and Gerhard Tutz. 2016. *Statistik*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-50372-0>.
- Reiß, Siegbert, and Viktor Sarris. 2012. *Experimentelle Psychologie: von der Theorie zur Praxis*. Pearson Studium Psychologie. München: Pearson.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

(10) Stichproben

Datum	Einheit	Thema
21.10.2021	Einführung	(1) Einführung
28.10.2021	Einführung	(2) Theorien, Hypothesen, Experimente
04.11.2021	Einführung	(3) Operationalisieren, Messen, Skalieren
11.11.2021	Studiendesign	(4) Prinzipien der Versuchsplanung
18.11.2021	Studiendesign	(5) Kriterien der Versuchsplanung
25.11.2021	Studiendesign	(6) Experimentelle Versuchspläne I
02.12.2021	Studiendesign	(7) Experimentelle Versuchspläne II
09.12.2021	Studiendesign	(8) Experimentelle Versuchspläne III
16.12.2021	Studiendesign	(9) Experimentelle Versuchspläne IV
	Weihnachtspause	
<b>13.01.2022</b>	<b>Datenerhebung</b>	<b>(10) Stichproben</b>
20.01.2022	Datenerhebung	(11) Datenerhebungsverfahren I
27.01.2022	Datenerhebung	(12) Datenerhebungsverfahren II
30.03.2022	Klausurtermin	12:00 - 13:00 Uhr G16-H5
Jul 2022	Klausurwiederholungstermin	

---

Grundgesamtheit und Stichprobe

Nichtprobabilistische Stichproben

Probabilistische Stichproben

Selbstkontrollfragen

Die Vorlesung basiert auf Kapitel 9 Stichprobenziehung in Döring and Bortz (2016).

---

## **Grundgesamtheit und Stichprobe**

Nichtprobabilistische Stichproben

Probabilistische Stichproben

Selbstkontrollfragen

## Grundgesamtheit und Vollerhebung

Die Grundgesamtheit (engl. population) ist die Gesamtmenge aller Untersuchungseinheiten, über die in wissenschaftlichen Studien Aussagen getroffen werden sollen. Werden Daten aller Untersuchungseinheiten einer Grundgesamtheit erhoben, so spricht man von einer Vollerhebung.

## Stichprobe und Teilerhebung

Eine ausgewählte Teilmenge von Untersuchungseinheiten aus der Grundgesamtheit wird Stichprobe (engl. sample) genannt. Werden Daten einer Stichprobe aus einer Grundgesamtheit erhoben, so spricht man von einer Teilerhebung.

Grundgesamtheit



Stichprobenziehung



Stichprobe



## Zielpopulation

Die Gesamtmenge aller Untersuchungseinheiten, über die in einer wissenschaftlichen Studie Aussagen getroffen werden sollen, wird Zielpopulation (engl. target population) genannt.

## Auswahlpopulation

Die Gesamtheit aller Untersuchungseinheiten, die prinzipiell die Möglichkeit haben, in eine Stichprobe zu gelangen, wird Auswahlpopulation (engl. frame population) genannt.

## Inferenzpopulation

Die Gesamtheit aller Untersuchungseinheiten, aus der eine Stichprobe tatsächlich stammt, wird Inferenzpopulation (engl. inference population) genannt.

⇒ Ziel-, Auswahl- und Inferenzpopulation müssen nicht identisch sein.

## Beispiel

Zielpopulation	Alle Menschen, die in Magdeburg leben
Auswahlpopulation	Alle Menschen, die in Magdeburg gemeldet sind
Inferenzpopulation	Abhängig von der Stichprobenzusammensetzung

## Undercoverage

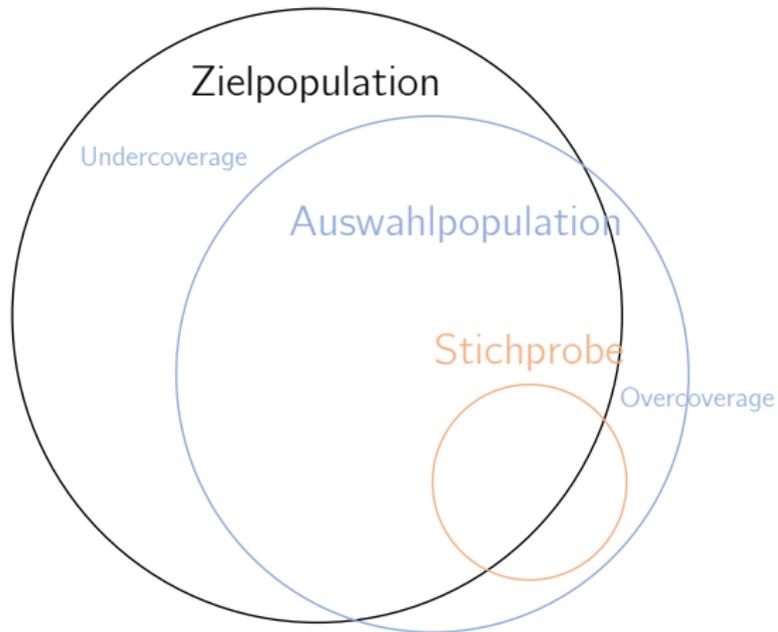
Die Menge der Untersuchungseinheiten der Zielpopulation, die nicht Teil der Auswahlpopulation sind, wird Undercoverage genannt.

- Wartezeit für Bürgeramtstermin in Magdeburg nach Zuzug beträgt etwa zwei Monate

## Overcoverage

Die Menge der Untersuchungseinheiten der Auswahlpopulation, die nicht Teil der Zielpopulation, wird Overcoverage genannt.

- Wartezeit für Bürgeramtstermin in Berlin nach Zuzug beträgt etwa zwei Monate



## Vollerhebung und Teilerhebung

### Vollerhebung

- Sinnvoll und möglich bei kleinen Grundgesamtheiten oder großen Ressourcen.
- Keine verzerrten Resultate durch verzerrte Stichprobenauswahl möglich.
- Eher kostenintensiv.
- Nicht möglich bei destruktiven Tests (z.B. Crashtests in der Autoindustrie).

### Teilerhebung

- Sinnvoll und möglich bei großen Grundgesamtheiten und beschränkten Ressourcen.
- Verzerrte Resultate durch verzerrte Stichprobenauswahl möglich.
- Eher weniger kostenintensiv.
- Möglich bei destruktiven Tests (z.B. Crashtests in der Autoindustrie).

⇒ Teilerhebungen sind in der quantitativen Forschung die Regel.

## Nichtprobabilistische Stichprobenarten

Auswahl der Untersuchungseinheiten basiert nicht explizit auf einem Zufallsprozess

- Gelegenheitsstichprobe
- Quotenstichprobe
- Spezielle Stichprobenverfahren

## Probabilistische Stichprobenarten

Auswahl der Untersuchungseinheiten basiert explizit auf einem Zufallsprozess

- Einfache Zufallsstichprobe
- Stratifizierte Stichprobe
- Clusterstichprobe

---

Grundgesamtheit und Stichprobe

## **Nichtprobabilistische Stichproben**

Probabilistische Stichproben

Selbstkontrollfragen

## Gelegenheitsstichprobe

### Definition

Bei einer Gelegenheitsstichprobe (engl. opportunity sample) werden willkürlich die Untersuchungseinheiten für eine wissenschaftliche Studie ausgewählt, die gerade zur Verfügung stehen oder leicht zugänglich sind (nach Döring and Bortz (2016), S. 306).

### Vorteil

Geringer Aufwand.

### Nachteil

Keine explizite Zielpopulation, nur posthoc Konzeption einer Inferenzpopulation möglich.

### Beispiele

Die meisten Stichproben in der akademischen psychologischen Grundlagenforschung sind Gelegenheitsstichproben (z.B. Versuchspersonenstunden im Psychologiestudium).

## Quotenstichprobe

### Definition

Die Stichprobenziehung nach dem Quotenverfahren (engl. quota sampling) setzt voraus, dass man die Populationszusammensetzung hinsichtlich einiger für die Untersuchungsfrage wichtiger soziodemographischer Merkmale kennt. Diese Merkmalsverteilung wird anhand eines Quotenplans durch eine systematische Auswahl von Untersuchungseinheiten nachgebildet (nach Döring and Bortz (2016), S. 307).

### Vorteil

Quotenstichproben sind hinsichtlich ihrer definierenden Merkmale für die Zielpopulation repräsentativ.

### Nachteil

Hoher Aufwand; interessierende und quotendefinierende Merkmale können divergieren.

### Beispiele

Fernseheinschaltquoten, Wahlumfragen.

## Spezielle Stichprobenverfahren

### Schneeballverfahren

- Einzelne Mitglieder einer Zielpopulation rekrutieren weitere Untersuchungseinheiten.
- Hilfreich, wenn Zielpopulation für Forschende schwer, für Mitglieder leicht erreichbar ist.
- Beispiele: Tabuisierte Zielpopulationen wie Hacker, Drogenkonsumierende, Sexarbeitende.

### Respondent Driven Sampling

- Weiterentwicklung des Schneeballverfahrens zur Steigerung der Repräsentativität.
- → Limitation der Rekrutierung pro Zielpopulationsmitglied.
- → Klar geregelte Incentivierung für Rekrutierung, Messung der Vernetzung.

### Targeted Sampling

- Einzelne Mitglieder einer Zielpopulation werden aktiv gesucht.
- Hilfreich, wenn Zielpopulation für Forschende und Mitglieder schwer erreichbar ist.
- Beispiele: Tabuisierte Zielpopulationen mit geringer Vernetzung wie Gewalttäter:innen.

---

Grundgesamtheit und Stichprobe

Nichtprobabilistische Stichproben

**Probabilistische Stichproben**

Selbstkontrollfragen

## Einfache Stichprobe

### Definition

Man zieht eine einfache Stichprobe (engl. simple random sample), indem man aus einer vollständigen Liste der Untersuchungseinheiten einer finiten Zielpopulation nach einem Zufallsprinzip (z.B. mittels eines Zufallszahlengenerators) eine Anzahl von Untersuchungseinheiten auswählt, wobei die Auswahlwahrscheinlichkeiten aller Untersuchungseinheiten gleich groß sein müssen (nach Döring and Bortz (2016), S. 311).

### Vorteil

Der Aufwand bei Ziehung einer einfachen Stichprobe ist überschaubar.

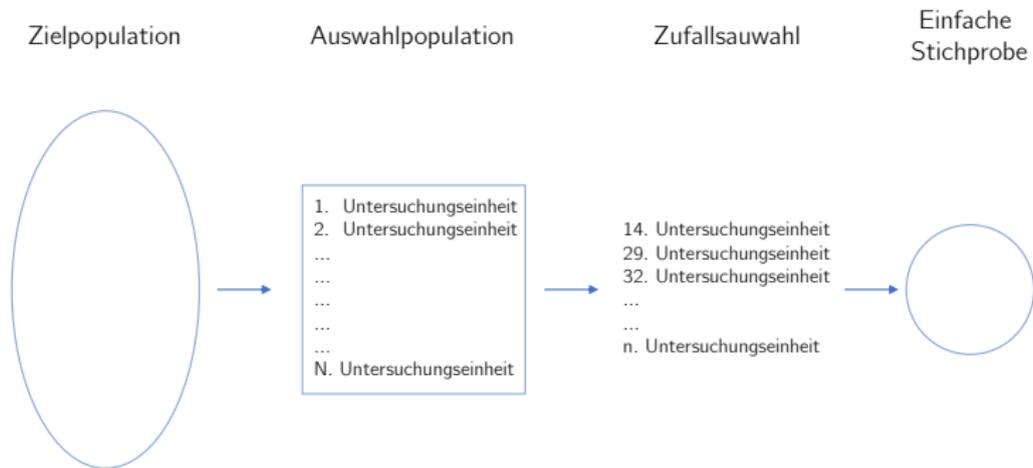
### Nachteil

Die Stichprobe kann über- oder unterproportional viele Untersuchungseinheiten im Vergleich zur Zielpopulation enthalten und somit nicht repräsentativ sein.

### Beispiel

Auswahl jeder  $n$ ten Untersuchungseinheit von einem zufälligen Startpunkt aus.

## Einfache Stichprobe



## Stratifizierte Stichprobe

### Definition

Man zieht eine stratifizierte Stichprobe (engl. stratified random sample), indem man die Zielpopulation auf der Basis einer oder mehrerer Merkmale in Teilpopulationen (Schichten) einteilt, so dass pro Merkmalsausprägung(skombination) eine Teilpopulation entsteht. Aus jeder der Teilpopulation zieht man dann eine einfache Stichprobe (nach Döring and Bortz (2016), S. 313).

### Vorteil

Hinsichtlich des stratifizierenden Merkmals repräsentiert die Stichprobe die Zielpopulation in der Regel besser als eine einfache Stichprobe.

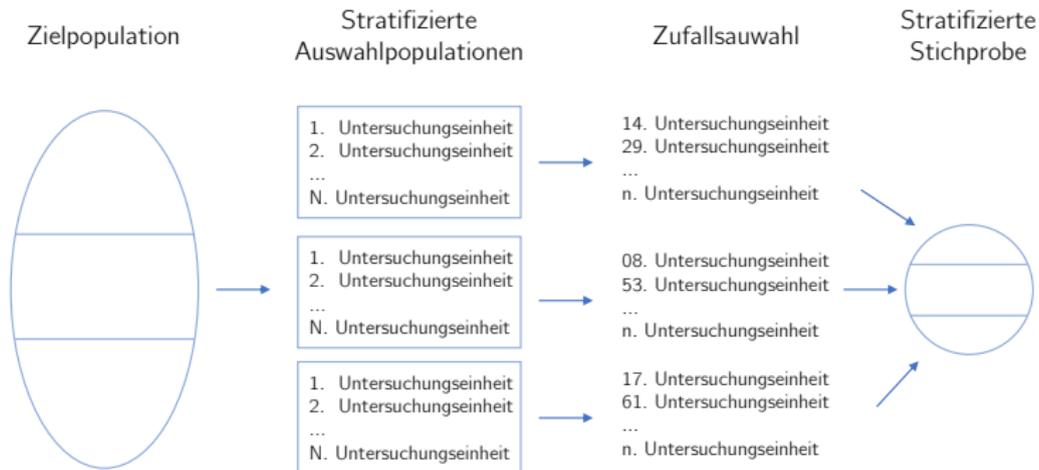
### Nachteil

Das stratifizierende Merkmal muss nicht zwangsläufig mit dem eigentlichen interessierenden Merkmal kovariieren und so die Stichprobe für das interessierende Merkmal für die Zielpopulation nicht repräsentativ sein; der Stratifizierungsaufwand kann groß sein.

### Beispiele

Stratifizierung nach soziodemographischen Merkmalen (Alter, Geschlecht, Einkommen, usw.).

## Stratifizierte Stichprobe



## Clusterstichprobe

### Definition

Man zieht eine Clusterstichprobe (engl. cluster sample), indem man aus einer in natürliche Gruppen (Cluster) gegliederter Zielpopulation nach dem Zufallsprinzip eine ausreichende Anzahl von Clustern auswählt und diese Cluster dann vollständig untersucht (nach Döring and Bortz (2016), S. 314).

### Vorteil

Geringer Aufwand, da nur eine Auflistung aller Cluster, nicht aber aller Untersuchungseinheiten der Zielpopulation nötig ist.

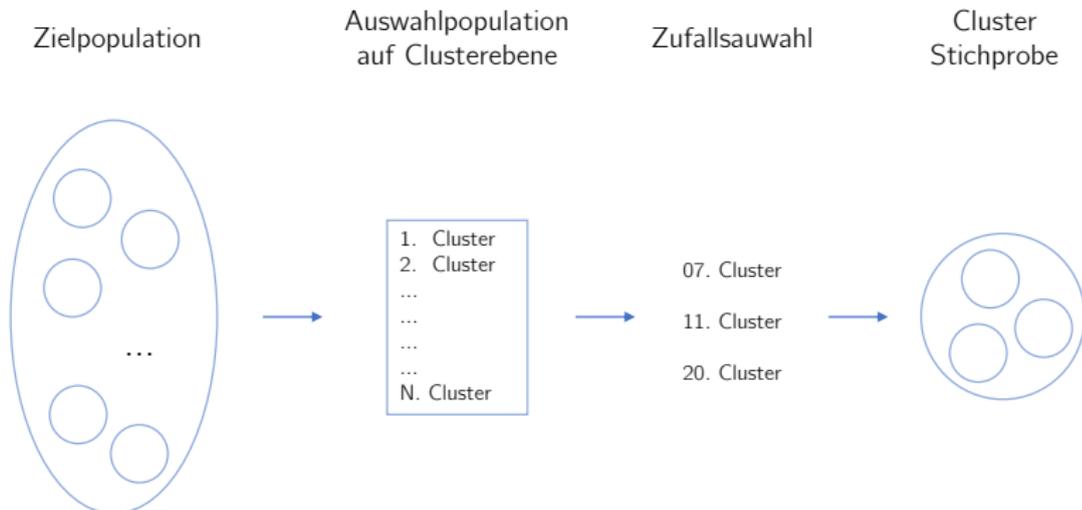
### Nachteil

Bei großer Variabilität von Merkmalsausprägungen zwischen den Clustern können die ausgewählten Cluster ein verzerrtes Bild der Zielpopulation liefern.

### Beispiele

Universitäten als Cluster für Studierende, Krankenhäuser als Cluster für Patient:innen.

## Clusterstichprobe



---

Grundgesamtheit und Stichprobe

Nichtprobabilistische Stichproben

Probabilistische Stichproben

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Erläutern Sie die Begriffe Grundgesamtheit, Stichprobe, Vollerhebung und Teilerhebung.
2. Erläutern Sie die Begriffe Zielpopulation, Auswahlpopulation und Inferenzpopulation.
3. Erläutern Sie die Begriffe Undercoverage und Overcoverage.
4. Nennen Sie drei Arten nichtprobabilistischer Stichproben.
5. Nennen Sie drei Arten probabilistischer Stichproben.
6. Erläutern Sie den Begriff und Vor- und Nachteile der Gelegenheitsstichprobe.
7. Erläutern Sie den Begriff und Vor- und Nachteile der Quotenstichprobe.
8. Erläutern Sie die Begriffe Schneeballverfahren, Respondent Driven Sampling und Targeted Sampling.
9. Erläutern Sie den Begriff und Vor- und Nachteile der einfachen Stichproben.
10. Erläutern Sie den Begriff und Vor- und Nachteile der stratifizierten Stichprobe.
11. Erläutern Sie den Begriff und Vor- und Nachteile der Clusterstichprobe.

Döring, Nicola, and Jürgen Bortz. 2016. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-41089-5>.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (11) Datenerhebungsverfahren I

---

## Psychologie als Datenwissenschaft des Erlebens und Verhaltens

### Befragung

- Mündliche Befragung
- Schriftliche Befragung

### Beobachtung

- Fremdbeobachtung
- Selbstbeobachtung

---

## Psychologie als Datenwissenschaft des Erlebens und Verhaltens

### **Befragung**

- **Mündliche Befragung**
- **Schriftliche Befragung**

### Beobachtung

- Fremdbeobachtung
- Selbstbeobachtung

---

Daten

Fragen

Mündliche Befragung

Schriftliche Befragung

Selbstkontrollfragen

---

## Daten

Fragen

Mündliche Befragung

Schriftliche Befragung

Selbstkontrollfragen

## Häufige Zieldaten bei Befragungen

### Demographische Eigenschaften

- Alter, Geschlecht, Ausbildung, Einkommen

### Dispositionen (Traits)

- Charaktereigenschaften

### Aktuelle Zustände (States)

- Momentane Stimmung

### Motivationen

- Leistungsmotivation

### Interessen

- Beruf, Hobbies

### Meinungen

- Einstellungen, Überzeugungen

### Überzeugungen zum eigenen Verhalten

- Gewohnheiten

---

Daten

## **Fragen**

Mündliche Befragung

Schriftliche Befragung

Selbstkontrollfragen

# Fragen

---

## Offene Frage

Frage ohne Vorgabe von Antwortmöglichkeiten

- Antwort im Referenzsystem der Befragten
- Antwort abhängig von der Artikulationsfähigkeit der Befragten
- Mögliche Intervieweffekte
- Notwendigkeit der Kodierung zur Datenanalyse nach Datenerhebung

## Geschlossene Frage

Frage mit Vorgabe von Antwortmöglichkeiten (Multiple-Choice Frage)

- Antwort im Referenzsystem der Interviewenden
- Antwort abhängig vom Antwortverständnis der Befragten
- Geringere Möglichkeit von Interviewereffekten
- Geringer Aufwand für Datenanalyse nach Datenerhebung

## Hybridfrage

Frage mit Vorgabe von erweiterbaren Antwortmöglichkeiten

## Fragendesign

- Einfache Sprache und eindeutige Begrifflichkeiten
- Beschränkung auf einen Sachverhalt pro Frage
- Neutrale nicht-sugesstive Formulierungen
- Reflektion von Fragenkontexteffekten
- Erschöpfende und disjunkte Antwortkategorien

---

Daten

Fragen

**Mündliche Befragung**

Schriftliche Befragung

Selbstkontrollfragen

## Unstandardisiertes Interview

- Interview ohne standardisierten Fragebogen
- Flexible Anpassung von Fragenformulierung
- Flexible Anpassung von Fragenreihenfolge

Natürliche soziale Interaktion ↑ Verlaufsvariabilität ↑ Frageninterpretationsspielraum ↑

## Teilstandardisiertes Interview

- Interview basierend auf einem Gesprächsleitfaden
- Vorgegebene Fragenformulierungen
- Flexible Anpassung der Fragenreihenfolge

Natürliche soziale Interaktion ↔ Verlaufsvariabilität ↔ Frageninterpretationsspielraum ↔

## Standardisiertes Interview

- Interview basierend auf einem Fragebogen
- Vorgegebene Fragenformulierungen
- Vorgegebene Fragenreihenfolge

Natürliche soziale Interaktion ↓ Verlaufsvariabilität ↓ Frageninterpretationsspielraum ↓

## Standardisiertes Interview

### Einleitungsfragen

- Eröffnung der Interviewsituation
- Leichte Beantwortbarkeit
- Abmildern von Ängsten der Interviewten vor Interviewsituation

### Überleitungsfragen

- Einleitung neuer Fragenkomplexe zu bestimmten Themen

### Filterfragen

- Steuerung des Interviewablaufs
- Beschränkung von Fragenkomplexen auf bestimmte Zielgruppen

## Standardisiertes Interview

### Designkriterien

- Einfachheit in der Handhabung für Interviewende
- Verständlichkeit der Fragen für Interviewte
- Engagement der Interviewten
- Kontinuität des Interviewflusses
- Güte der Filterführung
- Variation der Antwortmöglichkeiten
- Ausmaß von Fragenkontexteffekten
- Dauer und Belastung der Interviewten

## Standardisiertes Interview

### Beispiel: Strukturiertes Klinisches Interview für DSM-IV-I (SKID-I)

Verfahren zur Diagnostik psychischer Störungen nach dem DSM-IV Klassifikationssystem

DSM-IV : Diagnostic and Statistical Manual of Mental Disorders IV (1994 - 2013)

- (A) Affektive Symptome (z. B. Major Depression, Manie, Dysthymie)
- (B) Psychotische Symptome
- (C) Differentialdiagnose psychotischer Störungen (z. B. Schizophrenie, schizoaffektive Störung, Wahn)
- (D) Differentialdiagnose affektiver Störungen
- (E) Missbrauch und Abhängigkeit von psychotropen Substanzen
- (F) Angststörungen (z. B. Panikstörung, Agoraphobie, Soziale Phobie, Zwangsstörung)
- (G) Somatoforme Störungen (z. B. Somatisierungsstörung, Hypochondrie, Schmerzstörung)
- (H) Essstörungen (z. B. Anorexia Nervosa, Bulimia Nervosa)
- (I) Anpassungsstörungen (z. B. depressive Verstimmung, ängstliche Gehemmtheit, Verhaltensstörung)
- (J) Optionale Störungen (z. B. Stresstörungen, manische Episode)

### Kodierung durch Interviewenden

- ? = unsicher / zu wenig Informationen
- 1 = nein / nicht vorhanden
- 2 = vorhanden, aber nicht kriteriumsgemäß
- 3 = sicher vorhanden und kriteriumsgemäß

## Standardisiertes Interview

### Beispiel: Strukturiertes Klinisches Interview für DSM-IV-I (SKID-I)

Sektion A: Affektive Syndrome (A1-A4) 1

**SEKTION A: AFFEKTIVE SYNDROME**

In der Sektion A werden depressive, manische sowie andere affektive Syndrome beurteilt. Diagnosen werden jedoch z.T. erst in Sektion D abgeleitet. Wenn kein Hinweis auf eine derzeitige depressive Stimmungslage besteht, kreuzen Sie hier an und gehen Sie zu A38 (frühere Episode eines depressiven Syndroms)!

**Derzeitige Major Depression Episode**

**MDE-Kriterien**

Ich möchte Ihnen zunächst einige Fragen zu Ihrer Stimmung stellen.

**Während der letzten 4 Wochen ...**

A1) ... gab es da eine Zeitspanne, in der Sie sich fast jeden Tag nahezu durchgängig niedergeschlagen oder traurig fühlten? (Können Sie das genauer beschreiben?) ? 1 2 3

Wenn ja: Wie lange hielt dies insgesamt an? (2 Wochen lang?)

A2) ... haben Sie das Interesse oder die Freude an fast allen Aktivitäten verloren, die Ihnen gewöhnlich Freude machten? ? 1 2 3

Wenn ja: War dies fast jeden Tag der Fall? Wie lange hielt das an? (2 Wochen lang?)

Wenn keine 3 in 1) oder 2) → A38 (frühere MDE)

## Standardisiertes Interview

### Beispiel: Strukturiertes Klinisches Interview für DSM-IV-I (SKID-I)

Beziehen Sie sich bei den folgenden Fragen auf die schlimmsten 2 Wochen während des vergangenen Monats. (Kodieren Sie 1, wenn es sich eindeutig um einen allgemeinen medizinischen Krankheitsfaktor, stimmungsinkongruente Wahnvorstellungen oder Halluzinationen handelt.)

**Während dieser Zeit...**

A3 ...haben Sie da ab- oder zugenommen? (Wieviel? Haben Sie versucht, abzunehmen?) ? 1 2 3

Wenn nein: Wie war Ihr Appetit? (Im Vergleich zu sonst? Mußten Sie sich zum Essen zwingen? Haben Sie mehr/weniger als sonst gegessen?)

3) Erheblicher Gewichtsverlust ohne Diät oder Gewichtszunahme (mehr als 5% des Körpergewichts/Monat) oder verminderter/gesteigerter Appetit fast täglich. Bei Kindern die zu erwartende Gewichtszunahme beachten! ? 1 2 3

**Kreuzen Sie an**

Gewichtsverlust/verminderter Appetit

Gewichtszunahme/gesteigerter Appetit

A4 ... hatten Sie irgendwelche Schlafprobleme? (Ein- oder Durchschlafprobleme, häufiges oder zu frühes Erwachen, vermehrter Schlaf? Wieviele Stunden im Vergleich zu sonst? War dies fast täglich der Fall?) ? 1 2 3

4) Schlaflosigkeit oder vermehrter Schlaf fast täglich

Insomnia

Hypersomnia

## Standardisiertes Interview

### Interviewerschulung

- Erläuterung des Untersuchungskontextes
- Erläuterung des Fragebogens
- Erläuterung der Dokumentation von Antworten (“Kodierung”)
- Verhaltensregeln für Interviewsituation
- Probeinterviews und Rollenspiele
- Videoanalyse

# Mündliche Befragung

## Standardisiertes Interview

### Beispiel: Strukturiertes Klinisches Interview für DSM-IV-I (SKID-I)



## Standardisiertes Interview

### Beispiel: Strukturiertes Klinisches Interview für DSM-IV-I (SKID-I)

#### SKID-Grundregeln (DO's AND DONT's)

- DO:** Geben Sie dem Betroffenen vor Beginn eine kurze Erklärung über den Zweck des Interviews.

**DONT:** Entschuldigen Sie sich nicht dafür, ein strukturiertes Interview durchzuführen. (Sagen Sie etwa nicht: "Ich muß diese Fragen vorlesen. Die meisten werden nicht auf Sie zutreffen. Ertragen Sie es gemeinsam mit mir". "Ich muß dieses standardisierte Interview durchführen"). Wenn das SKID richtig angewendet wird, ist es ein klinisches Interview und bedarf keiner Entschuldigung. Die sehr kurze Einleitung im SKID-Überblick reicht in der Regel aus.
- DO:** Verwenden Sie den "Überblick" dazu, die Problemeinschätzung des Probanden zu erfahren. (Der Überblick kann auch zum Sammeln von spezifischen Informationen verwendet werden, die für bestimmte Untersuchungen notwendig sind und nicht durch das SKID abgedeckt werden z.B. Familiengeschichte.

**DONT:** Fragen Sie im Überblick nicht nach spezifischen Symptomen, die in späteren Abschnitten des SKID erfragt werden.
- DO:** Verschaffen Sie sich am Anfang des Interviews einen ausreichend großen Überblick über die derzeitige Störung, um den Kontext zu verstehen, in dem sich die Störung entwickelt hat.

**DONT:** Stürzen Sie sich bei dem oberflächlichen Überblick über die derzeitige Störung nicht in spezifische Fragen über Symptome. Dies verlängert das Interview und macht bei den strukturierten Fragen später Probleme!

## Standardisiertes Interview

### Methodische Probleme

#### Reaktivität

- Reaktion der Untersuchten auf die Untersuchungssituation
- Verzerrung des Antwortverhaltens möglich

#### Anwortverzerrungen

- Bewusst oder unbewusst
- Antwortverweigerung oder Pseudoantwort
- Sozial erwünschte Antworten
- Erinnerungslücken
- Interviewereffekte, Frageneffekte, Fragenpositionseffekte

---

Daten

Fragen

Mündliche Befragung

**Schriftliche Befragung**

Selbstkontrollfragen

## Definition

- Schriftliche Befragung in Abwesenheit der Interviewenden
- Online oder postalisch

## Potentielle Vorteile

- Geringer personeller Aufwand
- Keine Interviewendenfehler
- Keine Verzerrungen aufgrund persönlicher sozialer Interaktion
- Freie Wahl des Beantwortungszeitpunkts mit potentiell höherer Konzentration
- Höheres Maß an Anonymität

## Potentielle Nachteile

- Höhere Ausfallquoten
- Stichprobenverzerrungen durch Selbstrekrutierung text-affiner Befragter
- Keine Kontrolle der Erhebungssituation
- Keine Möglichkeit zur Erfassung alternativer Antworten
- Befragte haben einen Gesamtüberblick über Fragen vor dem Ausfüllen

# Schriftliche Befragung

---

## Mündliche Befragung

### Vorteil

- Erhebungsbedingungen kontrollierbar
- Direkte Erläuterung bei Fragenverständnisproblemen möglich

### Nachteile

- Verzerrung durch direkte soziale Interaktion möglich
- Instruktionsvarianz durch verschiedene Interviewende möglich

## Schriftliche Befragung

### Vorteile

- Keine Verzerrung durch direkte soziale Interaktion möglich
- Keine Instruktionsvarianz durch verschiedene Interviewende möglich

### Nachteile

- Erhebungsbedingungen nicht kontrollierbar
- Keine direkte Erläuterung bei Fragenverständnisproblemen möglich

---

Daten

Fragen

Mündliche Befragung

Schriftliche Befragung

**Selbstkontrollfragen**

1. Nennen Sie drei typische Datentypen, die mithilfe von Befragungen erhoben werden können.
2. Erläutern Sie die Begriffe der offenen Frage und der geschlossenen Frage.
3. Erläutern Sie die Begriffe des unstandardisierten und des standardisierten Interviews.
4. Erläutern Sie den Begriff der Reaktivität.
5. Diskutieren Sie Vor- und Nachteile mündlicher und schriftlicher Befragungen.



# Psychologische Forschungsmethoden

BSc Philosophie-Neurowissenschaften-Kognition WiSe 2021/22

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

## (12) Datenerhebungsverfahren II

---

## Psychologie als Datenwissenschaft des Erlebens und Verhaltens

### Befragung

- Mündliche Befragung
- Schriftliche Befragung

### Beobachtung

- Fremdbeobachtung
- Selbstbeobachtung

---

## Psychologie als Datenwissenschaft des Erlebens und Verhaltens

### Befragung

- Mündliche Befragung
- Schriftliche Befragung

### Beobachtung

- **Fremdbeobachtung**
- **Selbstbeobachtung**

---

Klassifikationsformen

Fremdbeobachtung

Selbstbeobachtung

Selbstkontrollfragen

---

## **Klassifikationsformen**

Fremdbeobachtung

Selbstbeobachtung

Selbstkontrollfragen

# Klassifikationsformen

---

## Direkt vs. Indirekt

- Direkte Beobachtung ist die Verhaltensbeobachtung im engeren Sinne.
- Indirekte Beobachtung ist die Beobachtung von Verhaltensartefakten.

## Systematisch vs. unsystematisch

- Systematische Beobachtungen basieren auf einem ausführlichen Beobachtungsschema.
- Unsystematische Beobachtungen basieren auf groben Beobachtungsanweisungen.

## Natürliche Umgebung vs. Laborumgebung

- Beobachtungen in der natürlichen Umgebung des Verhaltens ("im Feld").
- Beobachtungen im Labor.

## Offen vs. verdeckt

- Bei der offenen Beobachtung haben die Beobachteten Kenntnis von der Beobachtung.
- Bei der verdeckten Beobachtung haben die Beobachteten keine Kenntnis von der Beobachtung.

## Teilnehmend vs. nicht-teilnehmend

- Bei der teilnehmenden Beobachtung nehmen die Beobachtenden an Interaktionen teil.
- Bei der nicht-teilnehmenden Beobachtung ist dies nicht der Fall.

## Fremdbeobachtung vs. Selbstbeobachtung

- Bei der Fremdbeobachtung sind Beobachtende und Beobachtete verschieden.
- Bei der Selbstbeobachtung sind Beobachtende und Beobachtete identisch.

# Typische Beobachtungsverfahren in der Psychologie

---

## Direkt vs. Indirekt

- Direkte Beobachtung ist die Verhaltensbeobachtung im engeren Sinne.
- Indirekte Beobachtung ist die Beobachtung von Verhaltensartefakten.

## Systematisch vs. unsystematisch

- Systematische Beobachtungen basieren auf einem ausführlichen Beobachtungsschema.
- Unsystematische Beobachtungen basieren auf groben Beobachtungsanweisungen.

## Natürliche Umgebung vs. Laborumgebung

- Beobachtungen in der natürlichen Umgebung des Verhaltens ("im Feld").
- Beobachtungen im Labor.

## Offen vs. verdeckt

- Bei der offenen Beobachtung haben die Beobachteten Kenntnis von der Beobachtung.
- Bei der verdeckten Beobachtung haben die Beobachteten keine Kenntnis von der Beobachtung.

## Teilnehmend vs. nicht-teilnehmend

- Bei der teilnehmenden Beobachtung nehmen die Beobachtenden an Interaktionen teil.
- Bei der nicht-teilnehmenden Beobachtung ist dies nicht der Fall.

## Fremdbeobachtung vs. Selbstbeobachtung

- Bei der Fremdbeobachtung sind Beobachtende und Beobachtete verschieden.
- Bei der Selbstbeobachtung sind Beobachtende und Beobachtete identisch.

---

Klassifikationsformen

**Fremdbeobachtung**

Selbstbeobachtung

Selbstkontrollfragen

## Methoden der systematischen Fremdbeobachtung

### Event-sampling Verfahren (Ereignisgesteuerte Verfahren)

- Das Auftreten bzw. Nicht-Auftreten eines bestimmten Verhaltensmerkmals wird im natürlichen Verhaltensstrom zeitkontinuierlich kodiert und/oder quantifiziert.
- Meist zur Analyse von Videoaufzeichnungen eingesetzt.

### Time-sampling Verfahren (Zeitgesteuerte Verfahren)

- Das Auftreten bzw. Nicht-Auftreten eines bestimmten Verhaltensmerkmals wird nur innerhalb vorher definierter Zeitintervalle kodiert und/oder quantifiziert.
- Meist zur Beobachtung in natürlichen Umgebungen eingesetzt.

⇒ Automatisierung videobasierter systematischer Fremdbeobachtung ist ein aktives Forschungsfeld.

### Methodische Probleme der systematischen Fremdbeobachtung

- Unvollständigkeit des Verhaltensmerkmalschemas.
- Selektive Aufmerksamkeit der Beobachtenden.

## Beispiel: Time-sampling-basierte Spielverhaltenskodierung bei Kindern

4 Stunden, 15 Minutenintervalle, All-or-none Kodierung (grau)

	1	2	3	4
1 Besitz von Spielzeug				
2 Fragt nach Spielzeug	■	■	■	■
3 Befiehlt einem Kind, Spielzeug zu geben				
4 Schlägt einen Handel vor				
5 Appelliert an Fairness				
6 Macht mit				
7 Reicht nach einem Spielzeug				
8 Versucht gegen Widerstand, Spielzeug zu nehmen				
9 Stößt, schlägt, um Spielzeug zu erlangen	■	■	■	■
10 Ist erfolgreich, im Erlangen von Spielzeug				

Abbildung 9: Beispiel einer Checkliste. Aus Jack (1934, S. 23).<sup>75</sup>

# Fremdbeobachtung

## Beispiel: Time-sampling-basierte Mimikkodierung

3 Minuten, 15 Sekundenintervalle, Rating

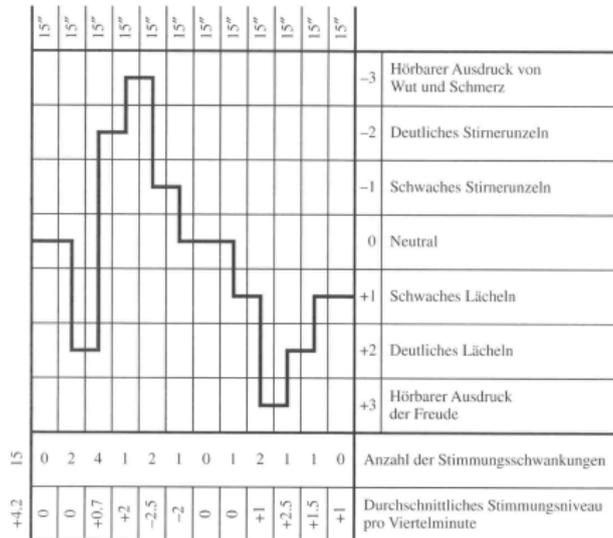


Abbildung 11: Rating-Skala zur Einschätzung der emotionalen Instabilität nach Lee (1932, S. 143). Die Skala wurde in einem 15-Sekunden-Rhythmus appliziert.<sup>78</sup>

---

Klassifikationsformen

Fremdbeobachtung

**Selbstbeobachtung**

Selbstkontrollfragen

# Selbstbeobachtung

## Ambulantes Assessment

Erfassung von Selbstberichten, Verhalten, und/oder physiologischen Maßen mithilfe eines schriftlichen Tagebuchs oder digitaler Endgeräte im Alltag in kontinuierlicher, zeit- oder ereignisabhängiger Form mit dem Ziele der Minimierung von Verzerrungen von posthoc Berichten oder Befragungen.



Quelle: [www.technabob.com](http://www.technabob.com)



Quelle: [www.osxdaily.com](http://www.osxdaily.com)

## Ambulantes Assessment

Aspekte des Ambulanten Assessment mithilfe digitaler Endgeräte

- Automatisierte Aufforderung zur Selbstbeobachtung
- Personalisierte Antwortmöglichkeiten
- Zuverlässige Ablaufsteuerung inklusive Antwortlatenzen
- Möglichkeit der Sperrung zur Revision von Einträgen
- Vermeidung von Retrospektionseffekten
- Automatische Datensicherung und Datentransfer
- Sicherstellung DSGVO-konformer digitaler Datenverarbeitung

## Ambulantes Assessment

Beispiel: <https://www.psymate.eu/>

English ▾

PsyMate™  
Insight into daily life

Welcome Vision FAQ Portfolio Basic Module My data

Home—info—PsyMate™ for myself

— Disable 'size' on Android

### PsyMate™ for myself

An app to really get a grip on your mood? A way to work on your recovery yourself? That's possible! We'd like to present the PsyMate™: a practical and scientifically based tool to get a grip on psychological and physical problems.

The PsyMate™ beeps at different times during the day and asks how you are doing at that moment. Where are you? With whom? What are you doing? To what extent do you experience feelings of stress, anger or sadness? Are you hungry, sleepy or craving for a cigarette? A tool that can be adapted to your personal or medical situation. You can use the PsyMate™ alone, but also with your doctor, therapist or a family member. By filling out the questions you get an insight in how you feel. When are you feeling good and when are you not? This way you can adjust yourself better or give your therapist a call before it's too late. Moreover, it feels nice to be actively involved in your treatment this way.

Using the PsyMate™ is easy: an app on your phone is all you need. Download the app now from the App Store or Play Store and discuss with your clinician how the two of you can get started with the PsyMate™.

— Disable 'size' on Android

PsyMate™ in research —

Maastricht University

smarteHealth  
smart solutions for personal health

Health Foundation  
Umburg

# Selbstbeobachtung

## Ambulantes Assessment

Beispiel: <https://www.psymate.eu/>

### Question types

I feel I'm being **active**

1  2  3  4  5  6  7

Not at all                      Moderate                      Very

Likert scale

Denk aan de voor jou **belangrijkste gebeurtenis** sinds de vorige piep.

**Deze gebeurtenis was:**

-3  -2  -1  0  1  2  3

heel onplezierig                      heel plezierig

Bipolar scale

Hvor **INTERESSERT** har du vært i det du har holdt på med?

ingenting                      veldig mye

OK

Visual analog scale

### What am I doing (just before the beep)?

resting

work, school

housekeeping, shopping

hygiene

eating, drinking

active relaxation

passive relaxation

something else

OK

Multiple choice

---

Klassifikationsformen

Fremdbeobachtung

Selbstbeobachtung

**Selbstkontrollfragen**

1. Nennen und erläutern Sie drei Klassifikationsformen von Beobachtungsverfahren.
2. Nennen und erläutern Sie zwei Verfahren zur systematischen Fremdbeobachtung.
3. Erläutern Sie den Begriff des Ambulanten Assessments.