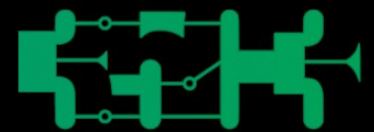


# Die dreckige Empirie

**Aiko**

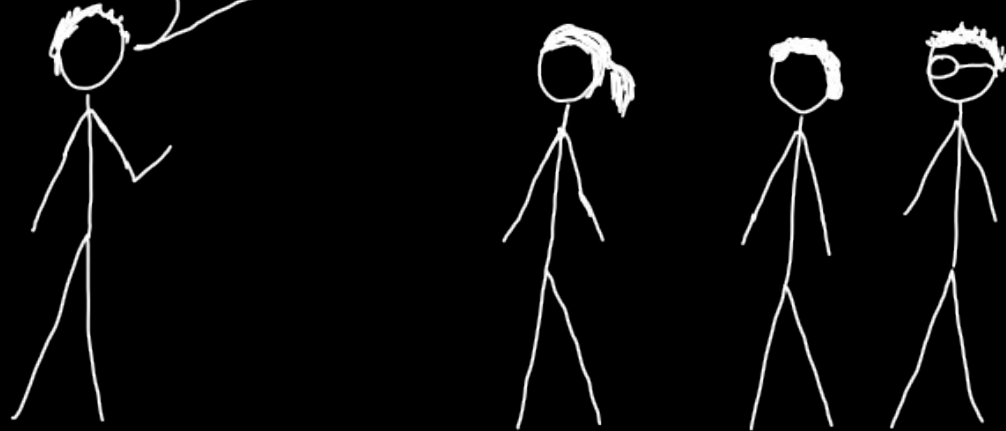
**REFRESHING**



**MEMORIES**

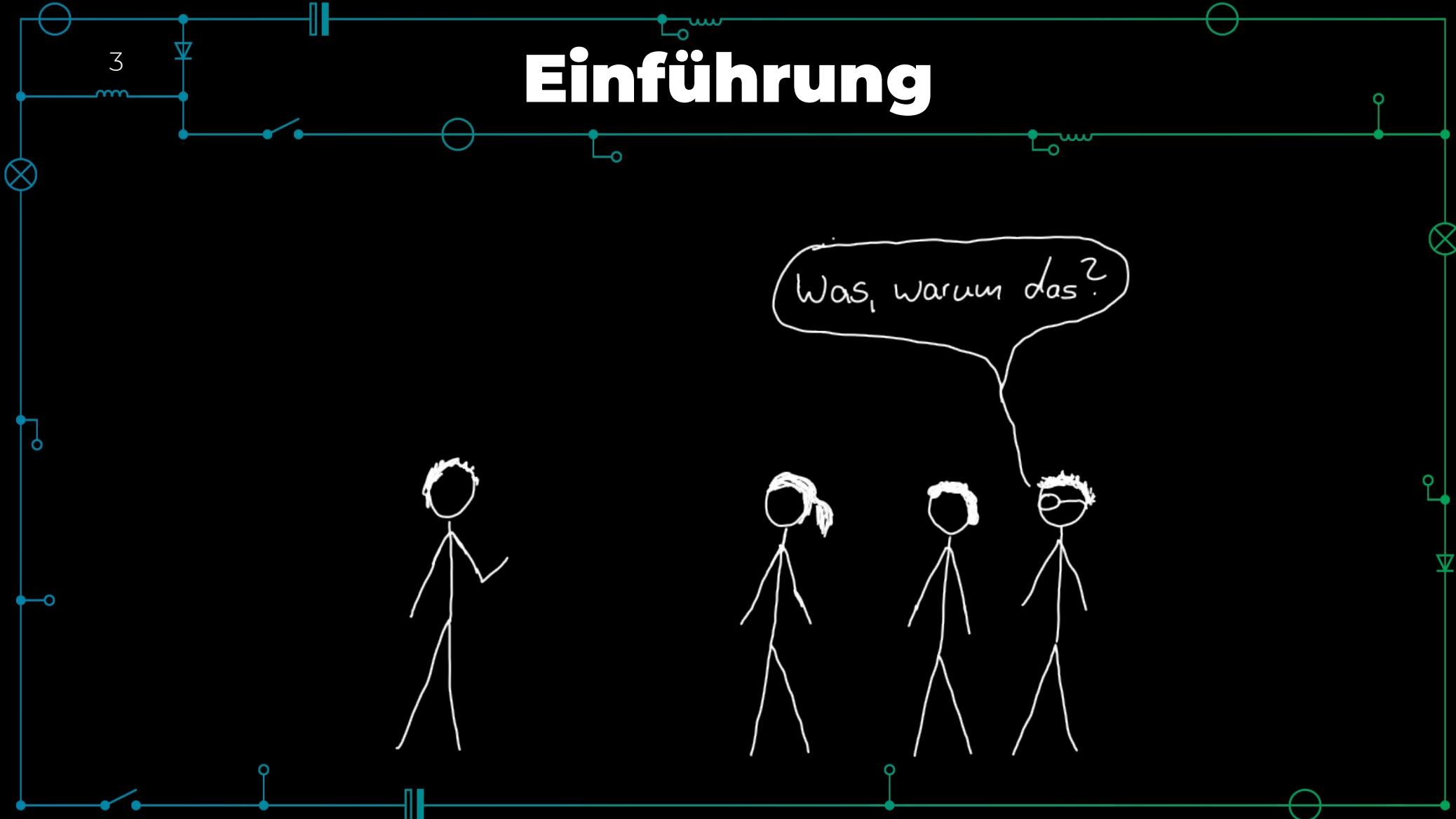
# Einführung

Puh, so viele Leute hier und mir ist schwindelig. Wenn ich umkippe hilft mir sicher niemand!



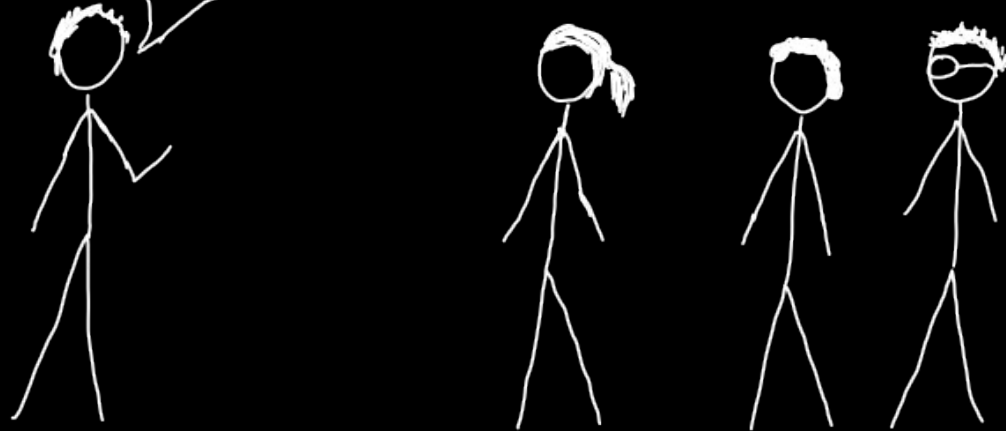
# Einführung

Was, warum das?



# Einführung

Kennst du nicht den Bystander Effect? Wenn viele Leute da sind hilft keiner! Ich hab das in einer Studie gelesen, die das bewiesen hat.



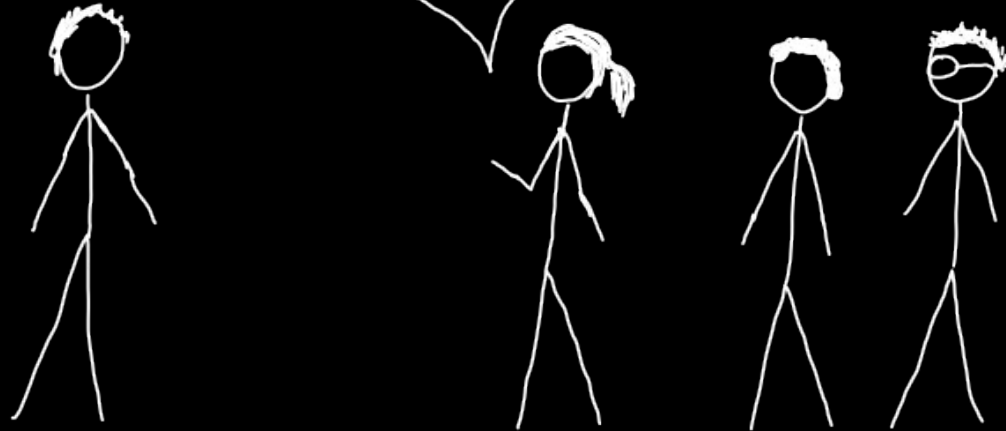
# Einführung

Das kann nicht stimmen!  
Mir hat auf dem Congress  
schon mal jemand geklopft!



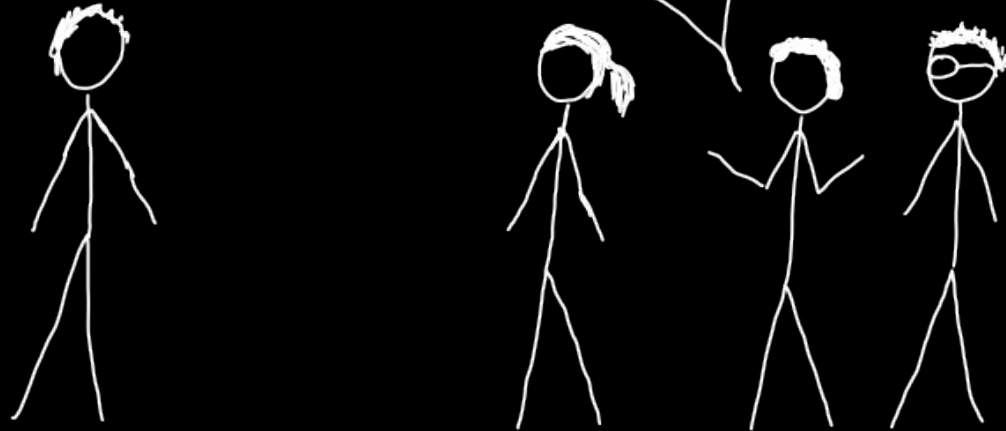
# Einführung

Das kann nicht stimmen! Die Studie von Borges & Pentra hat gefunden dass Menschen in Gruppen nicht seltener helfen.



# Einführung

Tja, sag ich doch immer:  
Traue keiner Statistik,  
die du nicht selbst  
gefälscht hast!





# Leitfragen

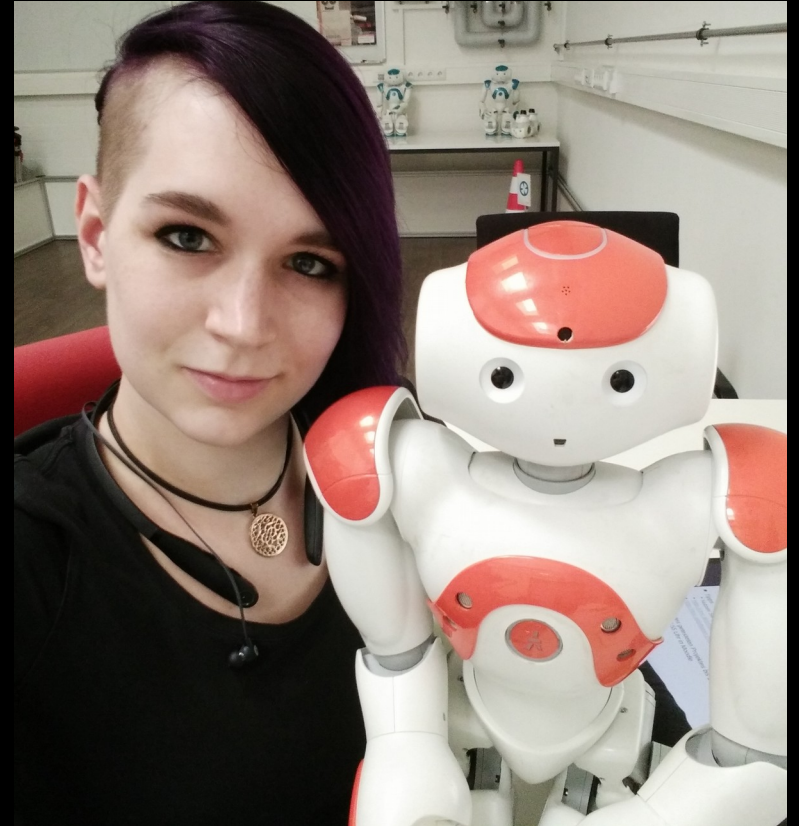
- **Kann man empirischen Studien trauen?**
- Wie funktioniert empirische Forschung?
- Was sagen Studien aus?
- Was ist eine gute Studie?
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?



# Wer ist das?

- Anna Klingauf
- Psychologin (M. Sc.)
- Momentan:  
Uni Kassel  
Mensch-Maschine-  
Systemtechnik

@aiko\_psi





# Beispiel: Facial Feedback

- Idee: Gesichtsausdruck hat Einfluss auf die Emotionen
- Lächeln: Positive Stimmung wird verstärkt
- Nicht lächeln: Positive Stimmung abgeschwächt



# Beispiel: Facial Feedback

- Zwei Erklärungsansätze:
  - Kognitiv: Selbstwahrnehmung – wenn ich lächle bin ich wohl glücklich
  - Physiologisch: Muskeln Lächeln hängen mit Wahrnehmung von Gefühlen direkter zusammen

# Beispiel: Facial Feedback

Journal of Personality and Social Psychology  
1988, Vol. 54, No. 5, 768-777

Copyright 1988 by the American Psychological Association, Inc.  
0022-3514/88/\$00.75

## Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis

Fritz Strack  
Universität Mannheim  
Mannheim, Federal Republic of Germany

Leonard L. Martin  
University of Illinois  
at Urbana-Champaign

Sabine Stepper  
Universität Mannheim  
Mannheim, Federal Republic of Germany

We investigated the hypothesis that people's facial activity influences their affective responses. Two studies were designed to both eliminate methodological problems of earlier experiments and clarify theoretical ambiguities. This was achieved by having subjects hold a pen in their mouth in ways that



# Leitfragen

- **Kann man empirischen Studien trauen?**
- *Wie funktioniert empirische Forschung?*
- *Was sagen Studien aus?*
- Was ist eine gute Studie?
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?



# Wie funktioniert Forschung?

- Idee: Vertrauen vs. Verstehen
- Studien besser verstehen = besser Einschätzen
- Hier: Quantitative Forschung
  - Hypothesen/ Annahmen prüfen
- Gegensatz: Qualitative Forschung
  - Hypothesen entwickeln



# Wie funktioniert Forschung?

- Hypothese
  - Lächeln verstärkt positive Emotionen
- Experiment
  - Versuchspersonen auswählen
  - Manipulieren: Lächeln oder nicht lächeln
  - Messen: positive Emotionen
  - Vergleichen (Statistik)

A faint circuit diagram is visible in the background, featuring a power source at the top left, a resistor labeled '16', a diode, a switch, and various other electronic components connected in a loop.

# Versuchspersonen

- Ideale Welt: Alle Menschen testen (Population)
- Realistisch: Möglichst große Stichprobe
- Ideal: Repräsentative Stichprobe
- Realistisch: Theoriegeleitet & Verfügbarkeit
- Zufällige Auswahl → Zufallseinfluss in den Daten
- Hier: 92 Studierende



# Manipulation

- Effekt wirklich Lächeln?
- Nur Lächeln ändern, sonst alles gleich
- Keine direkte Anweisung zu lächeln
- Verfälschung vermeiden: Sinn verschleiern

# Manipulation

- Hier: Stift in den Mund nehmen
  - Zwischen die Zähne = gleiche Muskeln wie lächeln
  - Zwischen die Lippen = Lächeln nicht mehr möglich
- Andere Erklärung für Lächeln = Ausschließen der kognitiven Theorie

# Messung

- Emotionen nicht direkt messbar!
- Indirekt messen: Befragen / Bewerten
- Emotionen auslösen: Cartoon
- Operationalisierung: Messbar machen
- Messung immer ungenau → zufällige Fehler  
→ Statistik

A circuit diagram is overlaid on the slide, featuring a power source at the top left, a resistor labeled '20', a diode, a switch, and various other components like capacitors and inductors connected in a loop. The diagram is drawn in a light blue color.

# Zusammenfassung Experiment

- Stift halten
  - Zähne = Lächeln
  - Lippen = nicht Lächeln
  - Hand = Kontrollgruppe
- Cartoon schauen und bewerten
- Coverstory: Menschen mit Behinderung

# Vergleichen

- Vergleich: Bewertung der Cartoons zwischen drei Gruppen
- Mehrere Personen, also Gruppendurchschnitt
- Direkt vergleichen?
  - Zufällige Fehler
  - Nur Stichprobe, nicht Population

# Vergleichen

Table 1

*Ratings of Funniness and Difficulty: Study 1*

Cartoon	Position of pen		
	Lip	Hand	Teeth
First	3.90	5.13	5.09
Second	4.00	4.10	4.19
Third	4.47	4.67	5.78
Fourth	4.90	5.17	5.50
Mean funniness	4.32	4.77	5.14
Mean difficulty	4.47	2.72	4.91

*Note.* All ratings were made on a scale from 0 to 9, where a lower value stands for lower funniness and difficulty, a higher value for higher funniness and difficulty.

# Vergleichen

- Statistisches Verfahren auswählen
- Sehr sehr große Auswahl
- T-Test: Vergleich zweier Mittelwerte
- Varianzanalyse: Vergleich mehrerer Gruppen
- Geplante Kontraste: Muster mehrerer Werte vergleichen

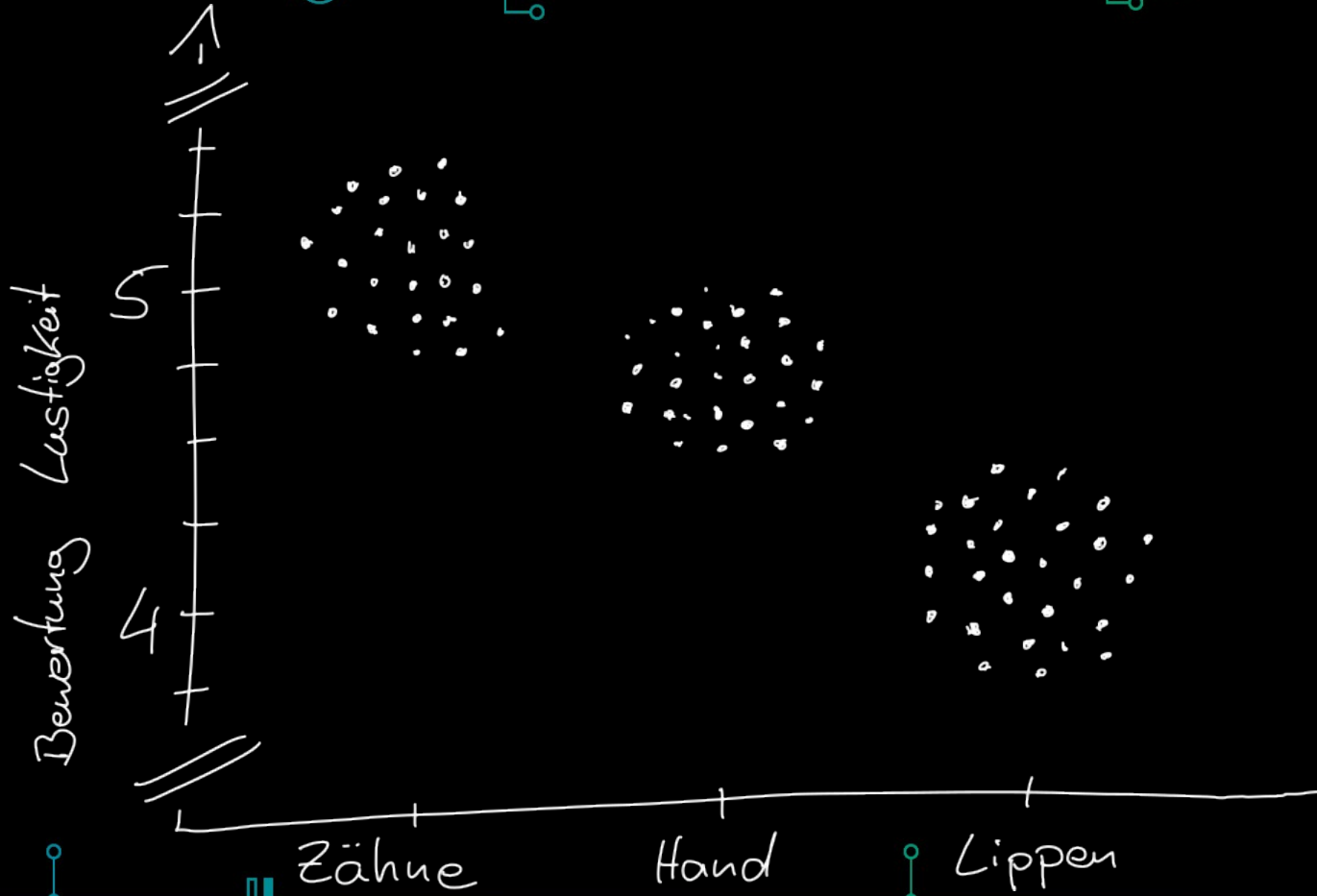


# Varianzanalyse

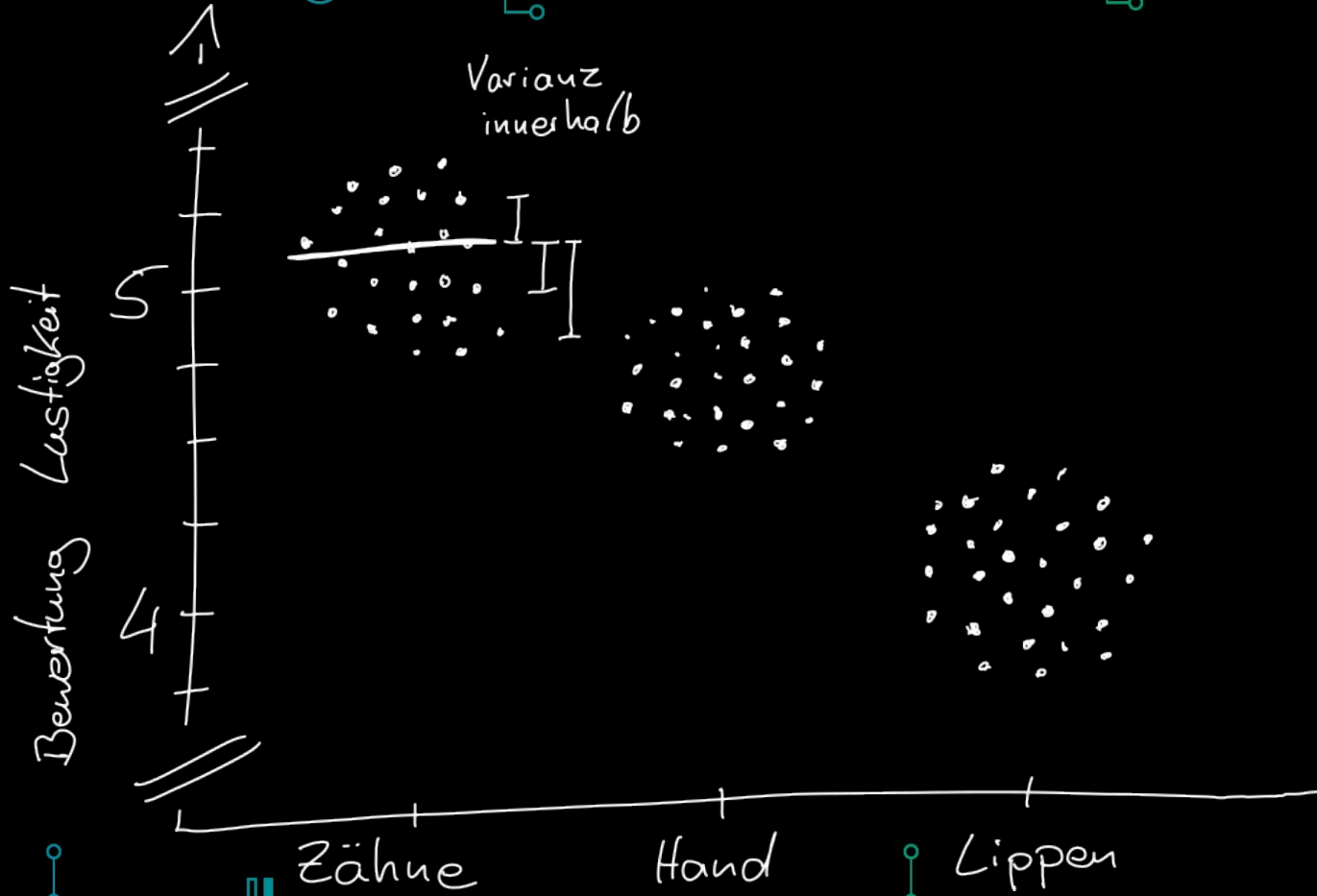
- Grundfrage: Unterscheiden sich die Werte zufällig voneinander oder wegen eines Effekts
- Also: Wie wahrscheinlich ist es, durch Zufall so ein Muster zu bekommen?



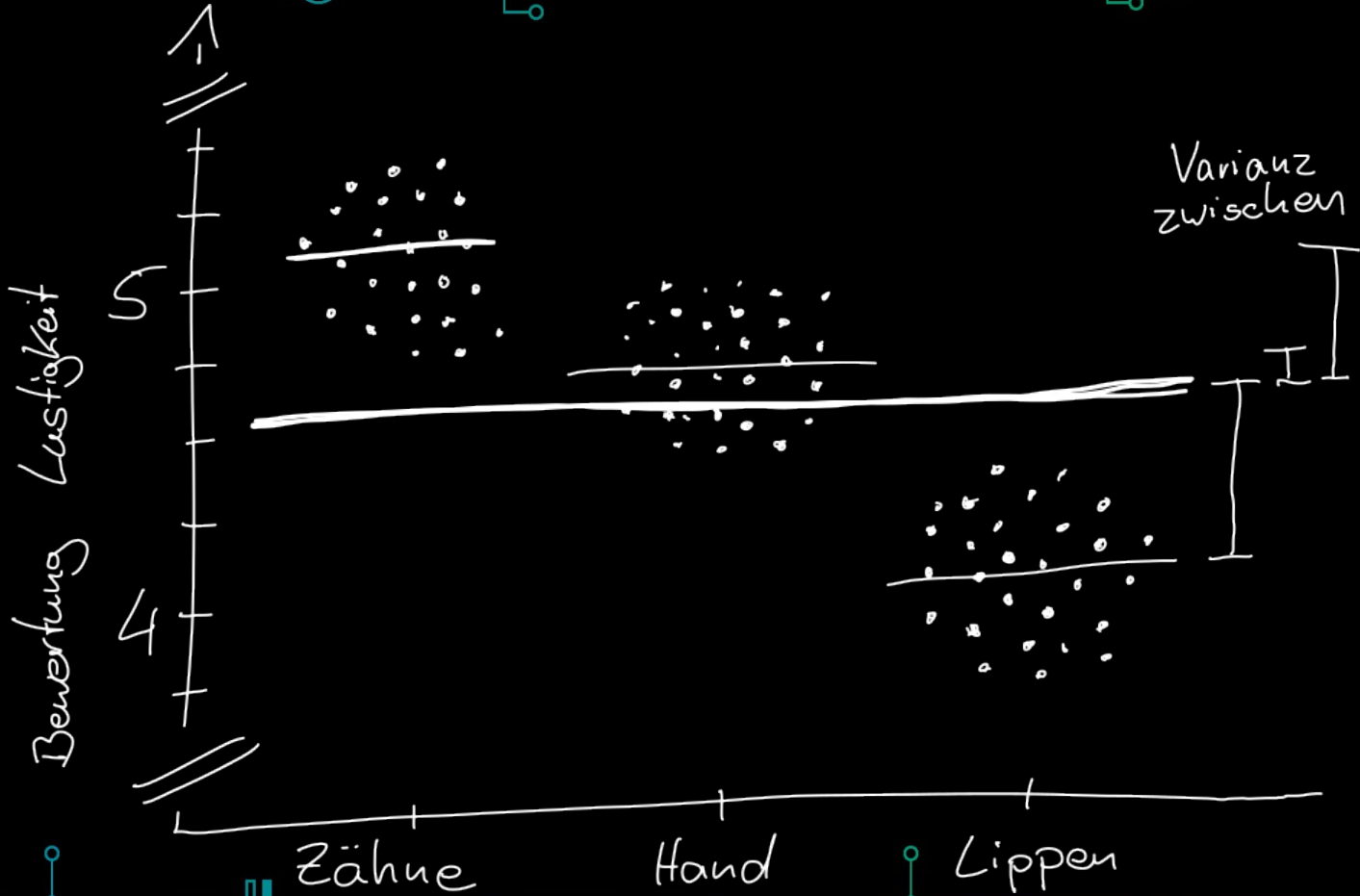
# Varianzanalyse



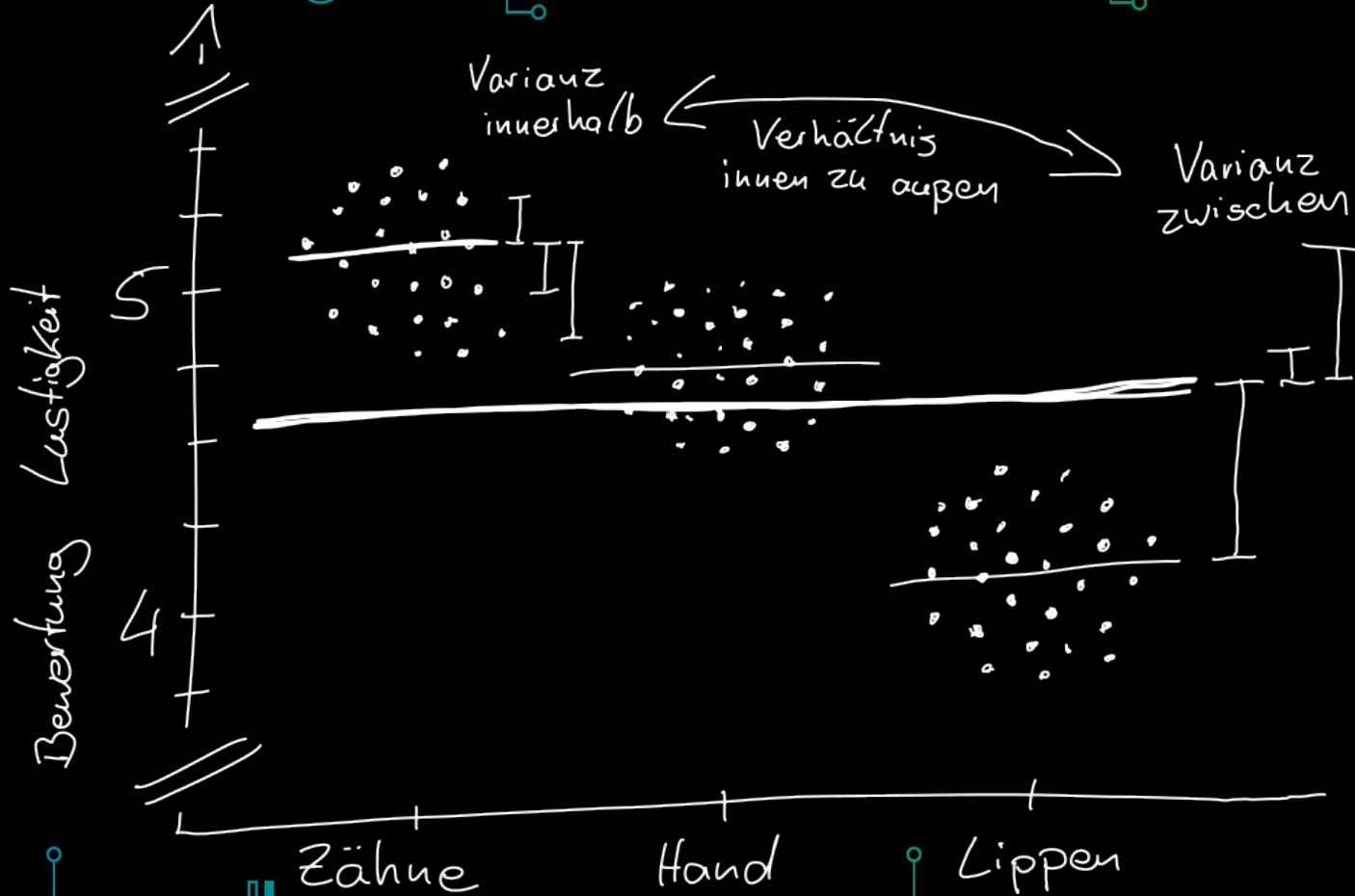
# Varianzanalyse



# Varianzanalyse

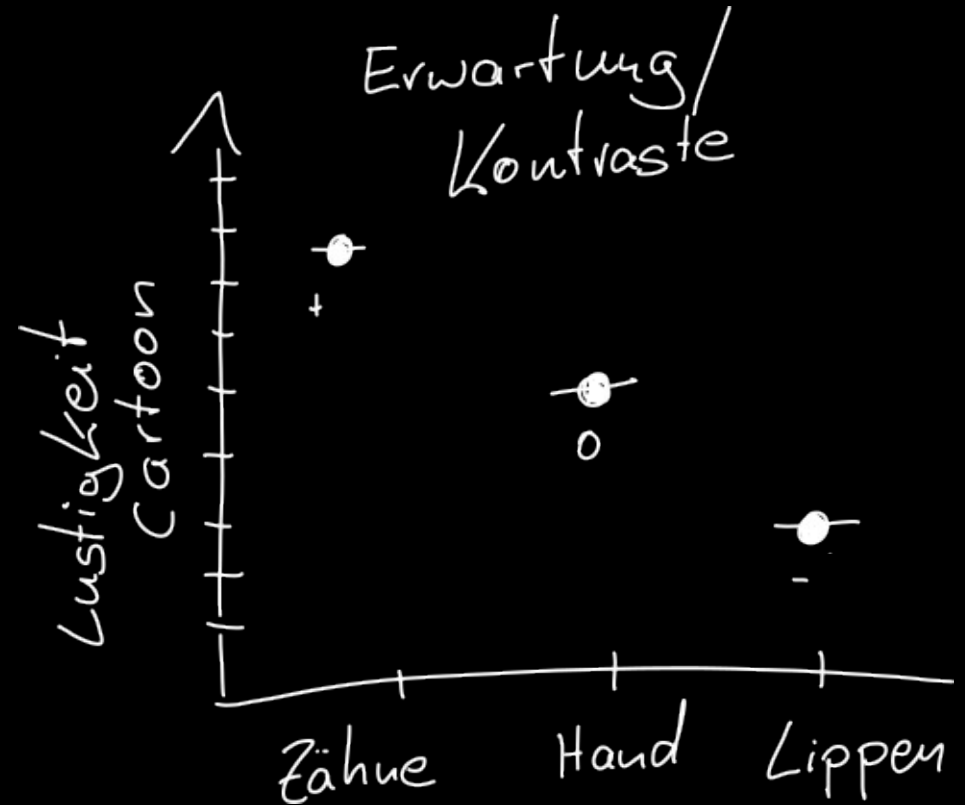


# Varianzanalyse



# Geplante Kontraste

- Spezialfall Varianzanalyse
- Muster erwartet:
- Kontraste



# Signifikanz

- Wahrscheinlichkeit, dass es nicht nur Zufall ist
- 100% Sicherheit = unrealistisches Ziel
- Oft 5% oder 1% Irrtumswahrscheinlichkeit
- Irrtumswahrscheinlichkeit:  $p$
- Wenn  $p < 0.05$  dann signifikant
- Hier:  $p = 0.03$



# Ergebnis

- Test signifikant
- Beobachteter Unterschied geht wahrscheinlich auf Manipulation zurück
- Da keine kognitive Verarbeitung:  
physiologische Theorie wird wahrscheinlicher
- Kein Beweis, aber auch kein Gegenbeweis  
→ Unterstützung der Hypothese



# Leitfragen

- **Kann man empirischen Studien trauen?**
- Wie funktioniert empirische Forschung?
- Was sagen Studien aus?
- *Was ist eine gute Studie?*
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?



A detailed illustration of a green printed circuit board (PCB) with various electronic components. The board features a network of copper traces, several through-hole components including resistors, capacitors, and integrated circuits, and a central area with a white background containing text. The components are arranged in a complex, interconnected pattern, typical of a multi-layer PCB design.

# Was ist eine gute Studie?

- Gute Operationalisierung
- Angemessene Stichprobe
- Ordentliche Statistik
- Korrelation und Kausalität
- Hinweis: Peer Review
- Hinweis: Impact Factor



# Gute Operationalisierung

- z.B. Standardisierte Tests
- Sinnvolle Erklärung der Messung
- Angabe von Qualitätskriterien (Reliabilität, Validität, etc.)
- Gut: Manipulation Check
- Ähnliche Studien: Messverfahren vergleichen



# Angemessene Stichprobe

- Dauemenregel: 30 Personen pro Gruppe
- Ständiges Problem: Studierende als Stichprobe
- Unterschiede zu anderen denkbar?
- Überprüfen: Experiment durchschaut?

# Ordentliche Statistik

- Richtiges statistisches Verfahren
- Beispiel: Viele Einzeltests (nicht gut)
  - Irrtumswahrscheinlichkeiten summieren sich
- Oft schwer einzuschätzen
- Merkmal: Wird genau berichtet?



# Korrelation und Kausalität

- Welche Schlüsse werden gezogen?
- Korrelation: Zusammenhang
- Kausalität: A verändert B
- Manipulation oder zeitliche Abfolge
- Andere Gründe auch beleuchtet?



# Hinweis: Peer Review

- Studien in Zeitschriften veröffentlicht
- Peer-Review: Andere Forscher aus dem Feld bewerten den Beitrag
- Nicht unumstritten
  - Bekanntheit, Zitieren unter Kollegen, etc.
- Aber: Oft guter Hinweis auf Qualität

A green circuit board with various components like resistors, capacitors, and a diode is visible in the background. The text is overlaid on this background.

# Hinweis: Impact Factor

- Kennwert der Zeitschrift
- Wie oft wurde daraus zitiert?
- Nicht überbewerten!
- Idee: Bekanntere Zeitschrift → mehr Auswahl → nimmt bessere Studien
- Sortiert „Predatory Journals“ oft aus



# Leitfragen

- **Kann man empirischen Studien trauen?**
- Wie funktioniert empirische Forschung?
- Was sagen Studien aus?
- Was ist eine gute Studie?
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?





# Beispiel: Bystander Effect

- Mehr Leute = weniger Hilfeleistung
- Grundlage: Kitty Genovese
- Forschung: John M. Darley
- Experiment: Simulierter Unfall, Anzahl der Personen variiert, Zeit bis Hilfeleistung
- Oft gefunden, viel diskutiert

A complex circuit diagram is overlaid on the slide. It features a central loop with a blue wire on the left and a green wire on the right. The top wire is blue and contains a battery, a resistor labeled '42', a diode pointing down, a switch, a resistor, and a diode pointing right. The bottom wire is green and contains a battery, a diode pointing left, a resistor, a switch, a resistor, and a diode pointing right. There are also two light bulbs, one on the left and one on the right, connected to the circuit.

# Unterschiedliche Ergebnisse

- Genau hinschauen: Unterschiede?
- Meta-Analysen
- Reviews



# Unterschiede?

- Ergebnisse scheinen nur widersprüchlich
- Stichprobe: Kinder vs. Erwachsene, etc.
- Operationalisierung: Mehrere echte Personen vs. mehrere eingeweihte
- Genauer herausfinden: Meta-Analyse, Moderatoren-Analyse



# Meta-Analysen

- Zusammenfassung mehrerer Studien
- Umfassende Literaturrecherchen
- Beschreibungen der Unterschiede
- Berechnung von Effektstärken
- Berechnung einer Gesamt-Effektstärke

# Effektstärken

- Kennwert wie groß ein Effekt ist
- Standardisiert = Vergleichbar
- z.B. Cohens d
  - Kleiner Effekt = 0.2
  - Mittlerer Effekt = 0.5
  - Großer Effekt = 0.8

# Meta-Analyse

Psychological Bulletin  
2011, Vol. 137, No. 4, 517–537

© 2011 American Psychological Association  
0033-2909/11/\$12.00 DOI: 10.1037/a0023304

## The Bystander-Effect: A Meta-Analytic Review on Bystander Intervention in Dangerous and Non-Dangerous Emergencies

Peter Fischer  
University of Regensburg

Joachim I. Krueger  
Brown University

Tobias Greitemeyer  
University of Innsbruck

Claudia Vogrincic  
University of Graz

Andreas Kastenmüller  
Liverpool John Moores University

Dieter Frey  
University of Munich

Moritz Heene, Magdalena Wicher, and Martina Kainbacher  
University of Graz



# Reviews

- Zusammenfassung des aktuellen Wissensstandes
- Meist von Forschern aus der entsprechenden Richtung
- Widersprüche mithilfe von Fachwissen aufgelöst



# Leitfragen

- **Kann man empirischen Studien trauen?**
- Wie funktioniert empirische Forschung?
- Was sagen Studien aus?
- Was ist eine gute Studie?
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?





# Probleme der Wissenschaft

- P-Hacking
- Replikationskrise
- Publication Bias
- ...



# p-Hacking

- Daten manipulieren → gewünschtes Ergebnis
- Ziel: p-Wert unter 5% drücken → Signifikanz
- Absichtliches Täuschen!
- Im Grunde kaum zu bemerken
- Ansatz: Replikationsstudien, ...
- Problem nicht gelöst

# Replikationskrise

- Replikation = Studie wiederholen
- Andere Forscher, gleiche Fragestellung
- Krise: Wichtige Befunde nicht replizierbar
- Ansätze: Bessere Studienplanung, Irrtumswahrscheinlichkeit reduzieren, ...
- Problem nicht gelöst



# Publication Bias

- Signifikante Ergebnisse = eher veröffentlicht
- Irrtumswahrscheinlichkeit + Publication Bias = falsches Bild
- Meta-Analysen: Korrekturrechnungen
- Literaturrecherche: Autoren anfragen
- Problem nicht gelöst



# Leitfragen

- **Kann man empirischen Studien trauen?**
- Wie funktioniert empirische Forschung?
- Was sagen Studien aus?
- Was ist eine gute Studie?
- Unterschiedliche Ergebnisse, was jetzt?
- Welche Probleme hat die Wissenschaft?



# Fazit

Kann man empirischen Studien trauen?

Ja, ABER...

The background of the slide is a detailed illustration of a green printed circuit board (PCB) with various electronic components. These include a 555 timer chip, several resistors, capacitors, a diode, a transistor, and various interconnecting traces and pads. The components are rendered in a light, semi-transparent style, allowing the text to be clearly visible.

# Fazit

- Studien: unterschiedliche Qualität
- Eigene Bewertung der Studienqualität
- Wahrscheinlichkeiten verstehen
- Meta-Analysen, Reviews, etc.
- offene Probleme in der Wissenschaft
- „Das beste was wir haben“ & kritisch bleiben

A decorative background circuit diagram in red and blue. It features a battery at the top, a light bulb on the left, a switch on the bottom, and various resistors and capacitors connected in a network. The word 'Danke!' is prominently displayed in the center.

# Danke!

## **Studien (frei verfügbar):**

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of personality and social psychology*, 8(4p1), 377.

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., ... & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological bulletin*, 137(4), 517.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of personality and social psychology*, 54(5), 768