

24.10.2023

Universität Regensburg

Marlena Baldauf

OPEN SCIENCE II: Lösungsansätze

RECAP: REPLIKATIONSKRISE

EINE SUBSTANTIELLE PROPORTION VON REPLIKATIONSSTUDIEN HAT SCHWÄCHERE EVIDENZ FÜR DEN UNTERSUCHTEN EFFEKT HERVORGEBRACHT ALS DIE ORIGINALSTUDIE
(OPEN SCIENCE COLLABORATION, 2015)

- Unsachgemäß verwendete Inferenzstatistik
 - Null-Hypothesen-Signifikanztesten
 - Multiples Testen
 - Wissenschaftsoziologische Faktoren
 - Publish or Perish
 - Publikationsbias?
- > Questionable Research Practices

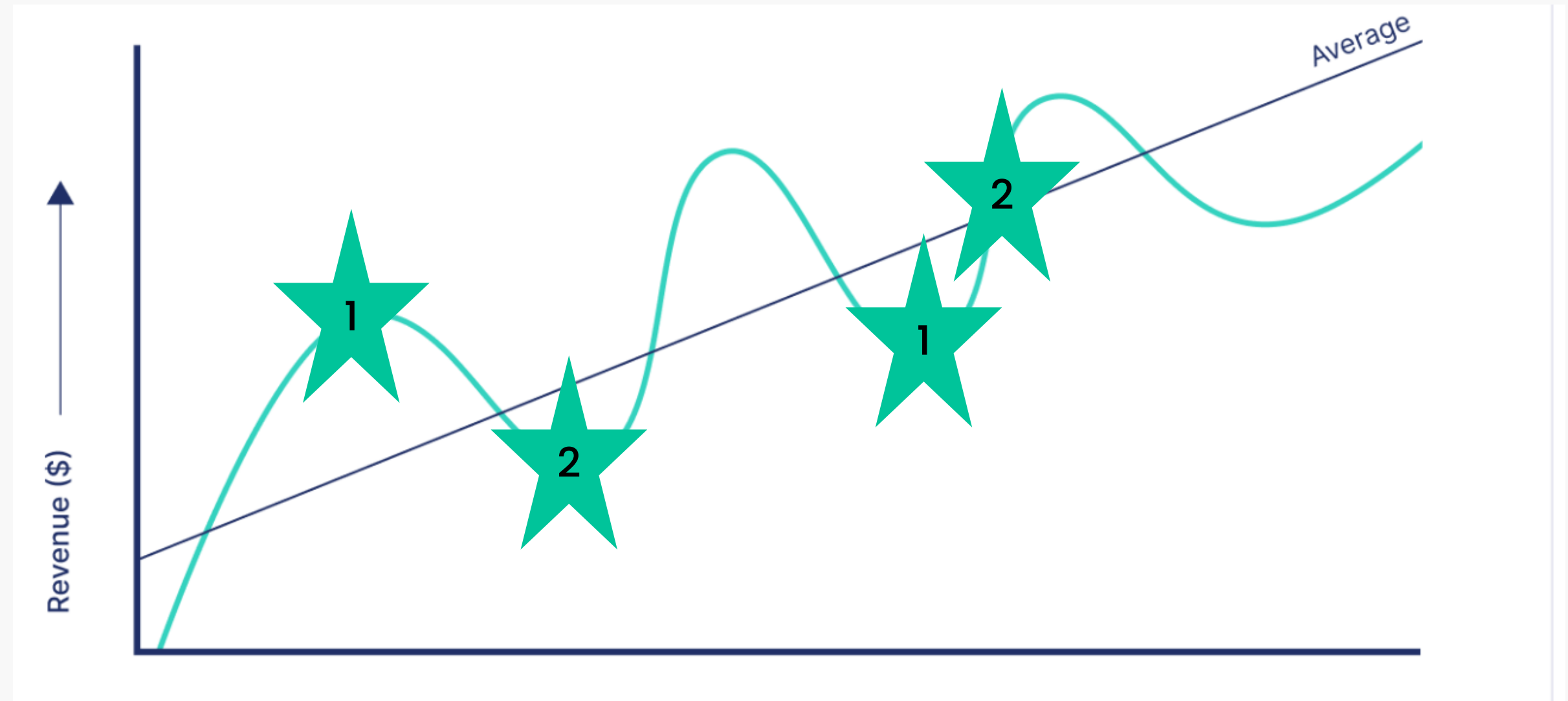
Recap

RECAP: Publication Bias

ZUSAMMENHANG ZWISCHEN PUBLIKATIONS BIAS UND NIEDRIGER REPRODUZIERBARKEIT?

REGRESSION TO THE MEAN!

= statistisches Phänomen, das die Tendenz von extremen Messungen beschreibt, sich bei nachfolgenden Messungen dem Durchschnitt (Mittelwert) anzunähern



Recap

ALTERNATIVE STATISTISCHE ANSÄTZE

- Unsachgemäß verwendete Inferenzstatistik
 - Null-Hypothesen-Signifikanztesten
 - Multiples Testen

WISSENSCHAFTS- PHILOSOPHISCHE ANSÄTZE

- Wissenschaftsoziologische Faktoren
 - Publish or Perish
- > Questionnable Research Practices

PREREGISTRATION/ OPEN SCIENCE

Heutige Sitzung

- I Preregistration
- II Open Science
- III Alternative Statistische Ansätze
- IV Wissenschaftsphilosophische Ansätze
- V Ausblick: nächste Woche

Was ist 'Preregistration'?

= Registrierung der Hypothesen, Methoden und/oder Analysen einer wissenschaftlichen Studie vor deren Durchführung

e.g. OSF Preregistration Template

Aspredicted.com



Preregistered Reports



Comparing dream to reality: an assessment of adherence of the first generation of preregistered studies

Aline Claesen, Sara Gomes, Francis Tuerlinckx and
Wolf Vanpaemel

II PREREGISTRATION

METHODE



papers with preregistered badge in
Psychological Science
from February 2015 to November 2017
 $n = 38$ plans (23 papers)

template

- Aspredicted.org: 1 plan
- van 't Veer and Giner-Sorolla: 1 plan
- COS prereg challenge registration from: 5 plans
- no template: 31 plans

repository

- Aspredicted.org: 1 plan
- OSF registration from: 2 plans
- OSF file: 22 plans
- OSF registration from and file: 4 plans
- OSF wiki: 8 plans
- not on repository: 1 plan

- time-stamped: 1 plan
 - non-ambiguously available: 4 plans
 - on a third-party repository: 1 plan
- excluded: $n = 4$ plans (2 papers)

$n = 34$ plans (21 papers) are accessible.

- variables: 2 plans
 - operationalization of variables: 1 plan
 - sample size: 3 plans
 - statistical model: 6 plans
- excluded: $n = 7$ plans (5 papers)

$n = 27$ plans (16 papers) contain minimal detail and are included for adherence assessment:

- research question and/or hypothesis
- variables
- sample size
- exclusion criteria
- procedure
- analysis

Figure 1. Flowchart of assessment of preregistered studies.

METHODE: AV

Primäre AV

Adhärenz zwischen Studie und dem vorregistrierten Plan (6 Dimensionen)

Sekundäre AVs

- Nutzung von Templates
- genutzte Plattform, um die Preregistration zu speichern

ERGEBNISSE: Adhärenz

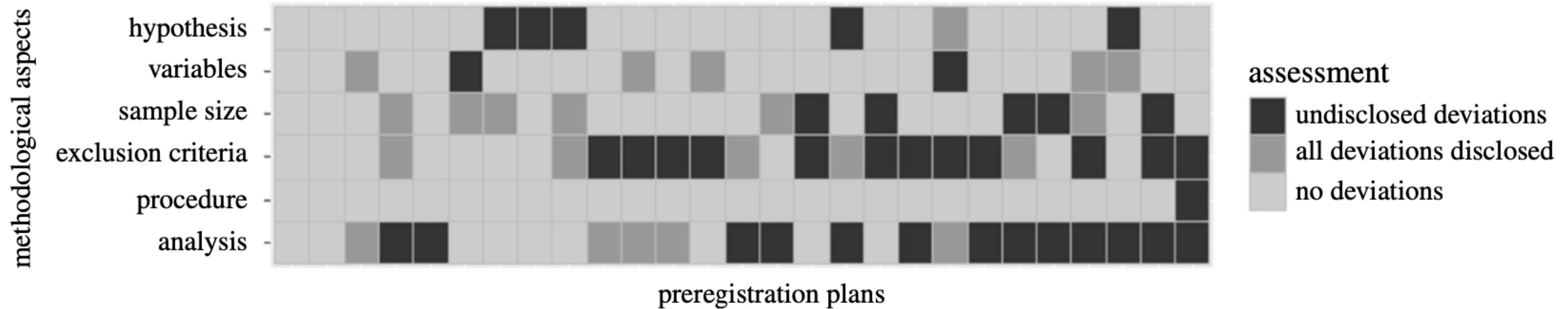


Figure 2. Tile plot of the assessment of each methodological aspect per preregistration plan. Only the 27 studies that were accessible and included the minimal number of methodological details required for our adherence assessment are shown.

DISKUSSION

Abweichung \neq Exploited Researcher Degrees of Freedom

- Abweichungen dem Forschungsprozess inhärent
- DISCLOSURE!

Fehlende Standardisierung

- Detaillierungsgrad
- welche Aspekte der Studie

DISKUSSION: Empfehlungen

1. Zugänglichkeit
2. Angemessener Detaillierungsgrad (Templates)
3. Review der Preregistration (seitens der Journals)
4. Explizite Nennung von Abweichungen
5. Erklärung der genannten Abweichungen
6. Richtlinien zur Berichterstattung

DISKUSSION: Limitationen

- kleine Sample Size
- Vorteil für vage Studien
- eher oberflächliche Betrachtung der Studien
- keine Wertung von Abweichungen
- Zeitrahmen

OPEN SCIENCE

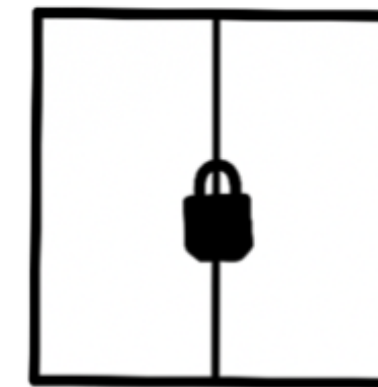
What's good for research

Careful research
Transparency
Collaboration



How research careers work

High output
Secrecy
Competitiveness



Was ist 'Open Science'?

"Open Science aims to make all parts of the research cycle as accessible as possible for as many people as possible. It applies to all disciplines, the sciences and the humanities, from archaeology to zoology."

- transparente & reproduzierbare Forschung
- Zugang zu Publikationen
- Austausch von Forschungsdaten
- Freier Zugang zu Forschungssoftware
- etc...

Open Data

Findable

Metadata and data should be findable for both humans and computers

F

A

Interoperable

Data needs to work with applications or workflows for analysis, storage and processing

I

R

Accessible

Once found, users need to know how the data can be accessed

Reusable

The goal of FAIR is to optimise data reuse via comprehensive well-described metadata

Open Data: Datenschutz?

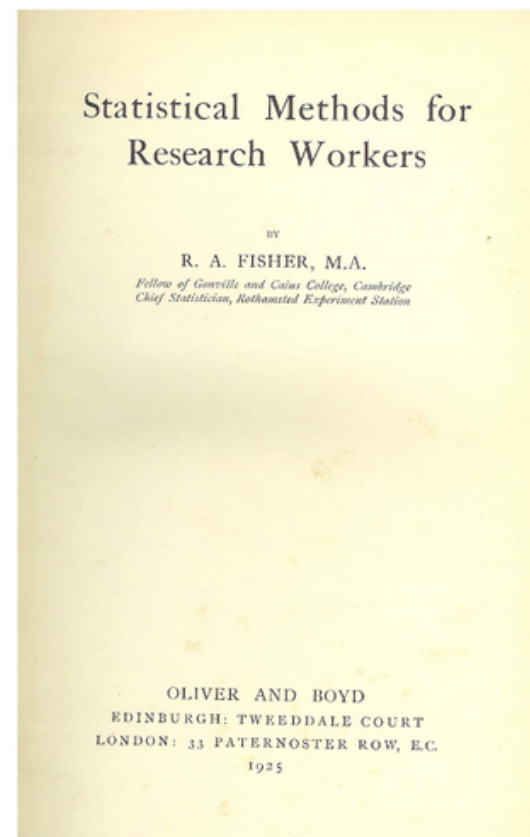
- systematische Einschätzung des Risikos der Identifizierung in 2.169 psychologischen Datensätzen (Psychological Science, Judgment and Decision Making, PLOS ONE)
- in 5,3% konnten Einzelpersonen anhand von Namen, IP-Adressen, Web-Identifikatoren, Geburtsdaten, Postleitzahlen & Initialen in Kombination mit anderen demografischen Variablen identifiziert werden
- 4,3 % enthielten (oft unnötige) demografische Informationen, die gewisses Risiko der Re-Identifizierung darstellten
- 53 % erhielten (potenziell) identifizierenden Daten, die gemäß der Datenschutz-Grundverordnung als sensibel gelten

IV FREQUENTIST VERSUS BAYESIAN STATISTICS

FREQUENTIST STATISTICS

Null-Hypothesen-Signifikanz-Testen

- Bestimmung von $P(D|H_0)$
- Anwendung einer Entscheidungsregel bzgl. H_0



ON THE USE AND INTERPRETATION OF CERTAIN TEST CRITERIA FOR PURPOSES OF STATISTICAL INFERENCE. PART I.

By J. NEYMAN, Ph.D. AND E. S. PEARSON, D.Sc.

CONTENTS.

	PAGE
I. Introductory	175
II. Sampling from a Normal Population:	
(1) Description of the Fundamental Space	178
(2) The Use of the Contour Systems in testing Hypothesis A	182
(3) The Criterion of Likelihood	184
(4) The Application of the Criterion of Likelihood in testing Hypothesis A	187
(5) Student's Test and Hypothesis B	189
(6) Alternative Method of examining Hypothesis B	191
(7) Solutions obtained by the Inverse Method	192
(8) Analysis of Charch's samples from a Skew Population	197
(9) Illustrations of the use of the Tables of F_1	202
III. Sampling from a Rectangular Population:	
(10) Sampling Distributions of the Frequency Constants	208
(11) Solutions obtained by the Inverse Method	213
(12) Illustrations from Experimental Sampling	214
IV. Sampling from an Exponential Population:	
(13) Sampling Distributions of Frequency Constants	221
(14) Test for Random Intervals	226
V. Conclusion	230
VI. Appendix with F_1 Tables	233

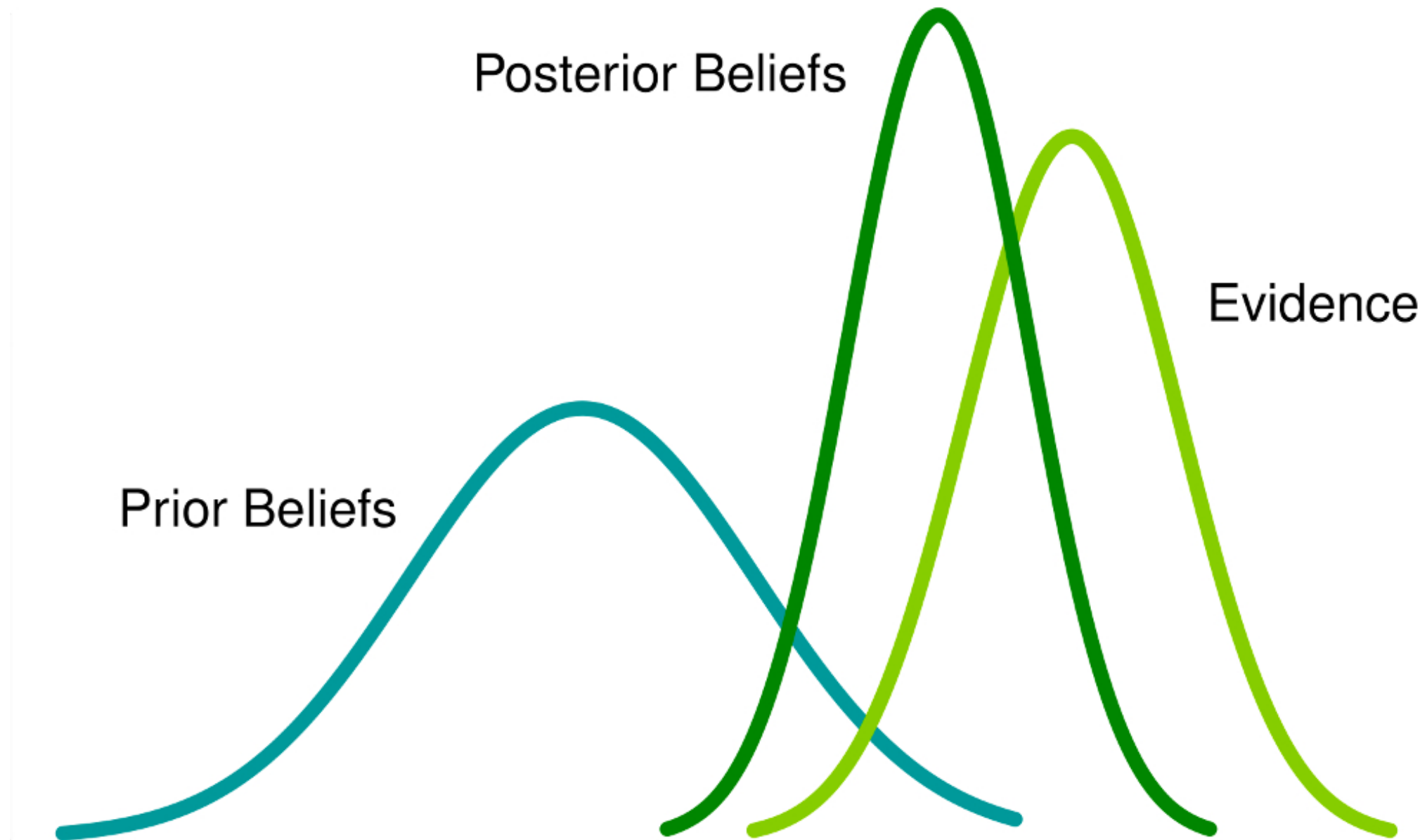
I. INTRODUCTORY.

ONE of the most common as well as most important problems which arise in the interpretation of statistical results, is that of deciding whether or not a particular sample may be judged as likely to have been randomly drawn from a certain population, whose form may be either completely or only partially specified. We may term Hypothesis A the hypothesis that the population from which the sample Σ has been randomly drawn is that specified, namely II. In general the method of procedure is to apply certain tests or criteria, the results of which will enable the investigator to decide with a greater or less degree of confidence whether to accept or reject Hypothesis A, or, as is often the case, will show him that further data are required before a decision can be reached. At first sight the problem may be thought to be a simple one, but upon fuller examination one

KRITIK: FREQUENTIST STATISTICS

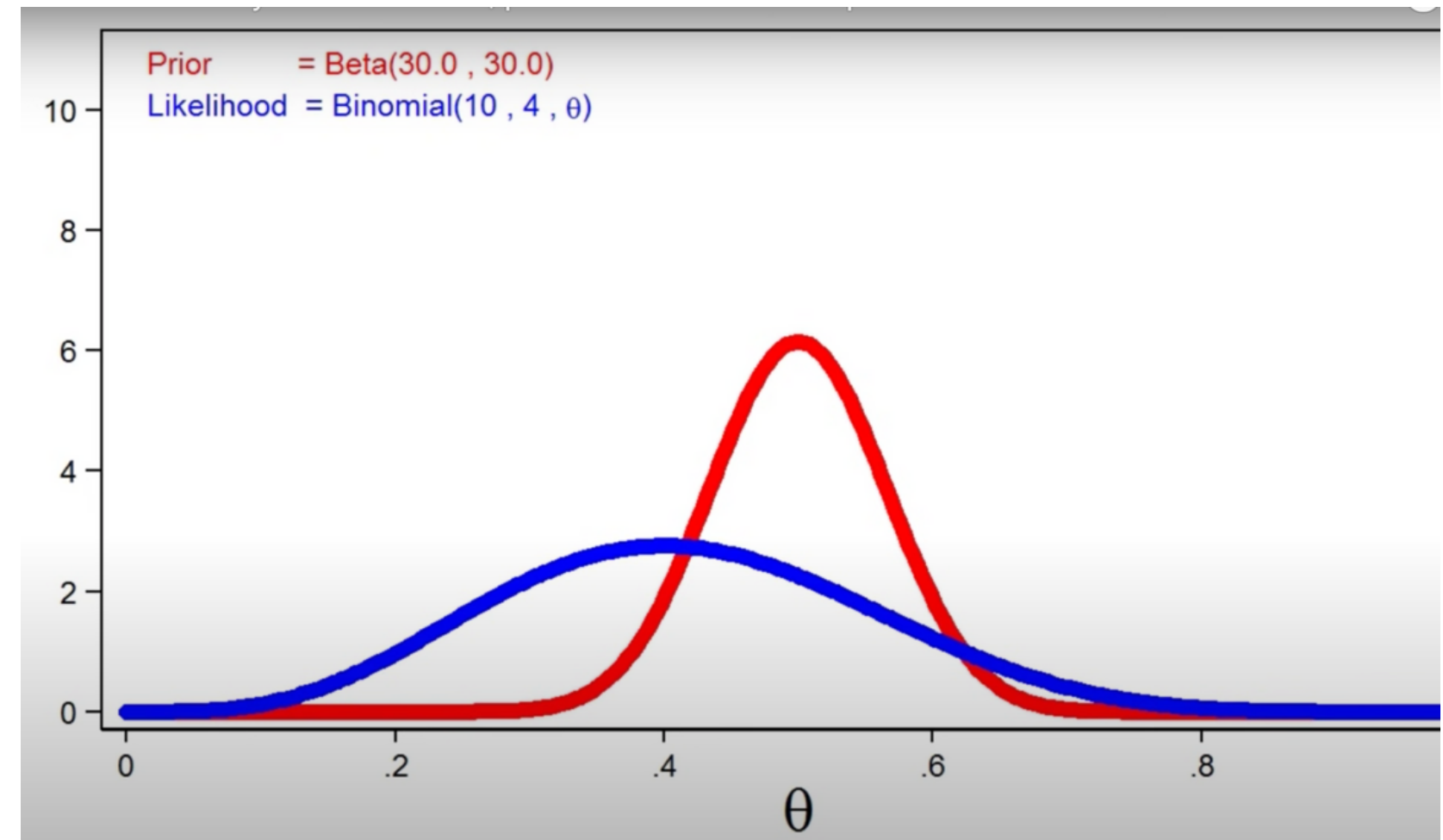
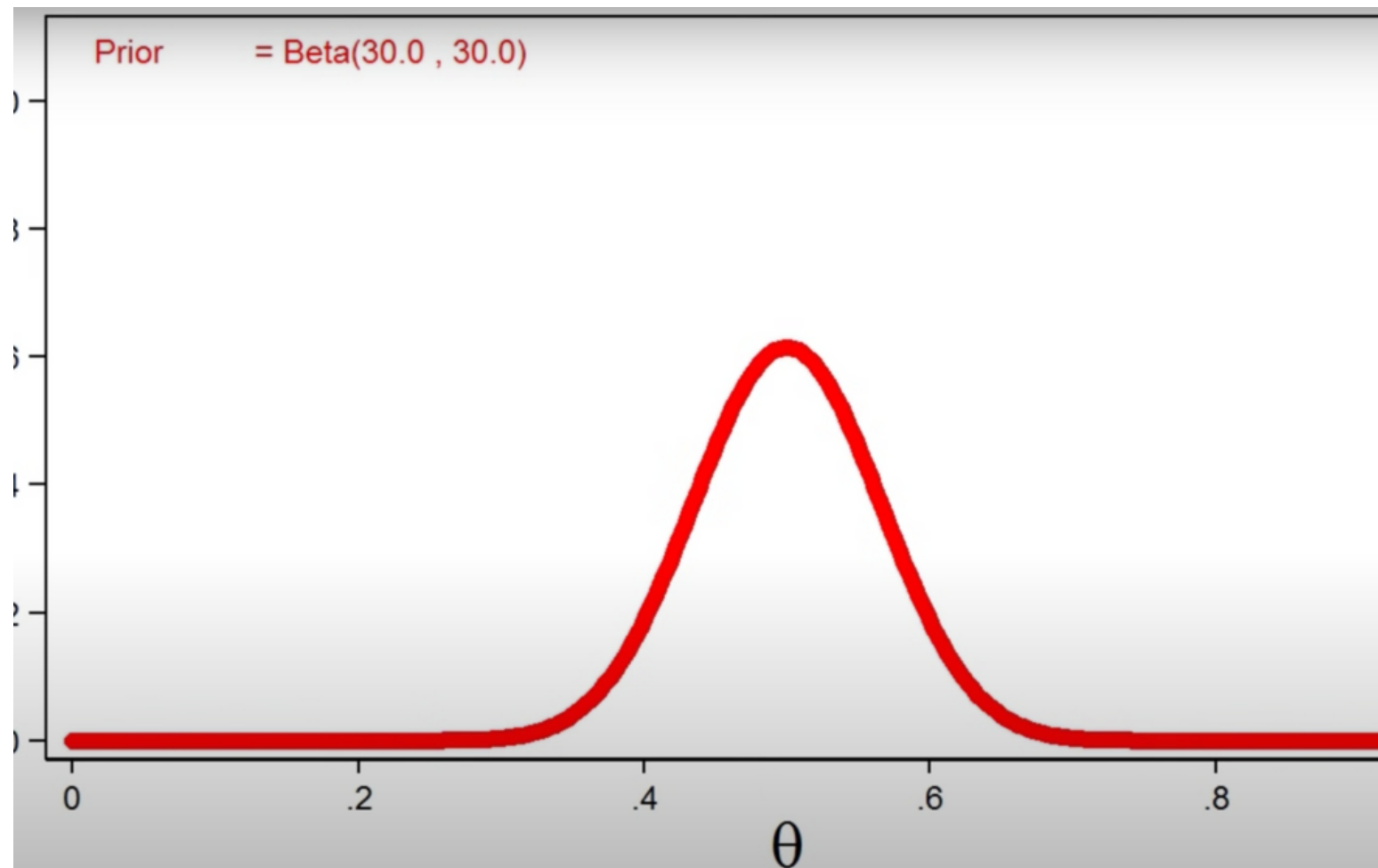
- 0.05 als arbiträrer Wert, künstliche Dichotomisierung
 - $p = 0.051 \neq p = 0.049$?
- großes Sample notwendig
- Typ I Fehler

BAYESIAN STATISTICS

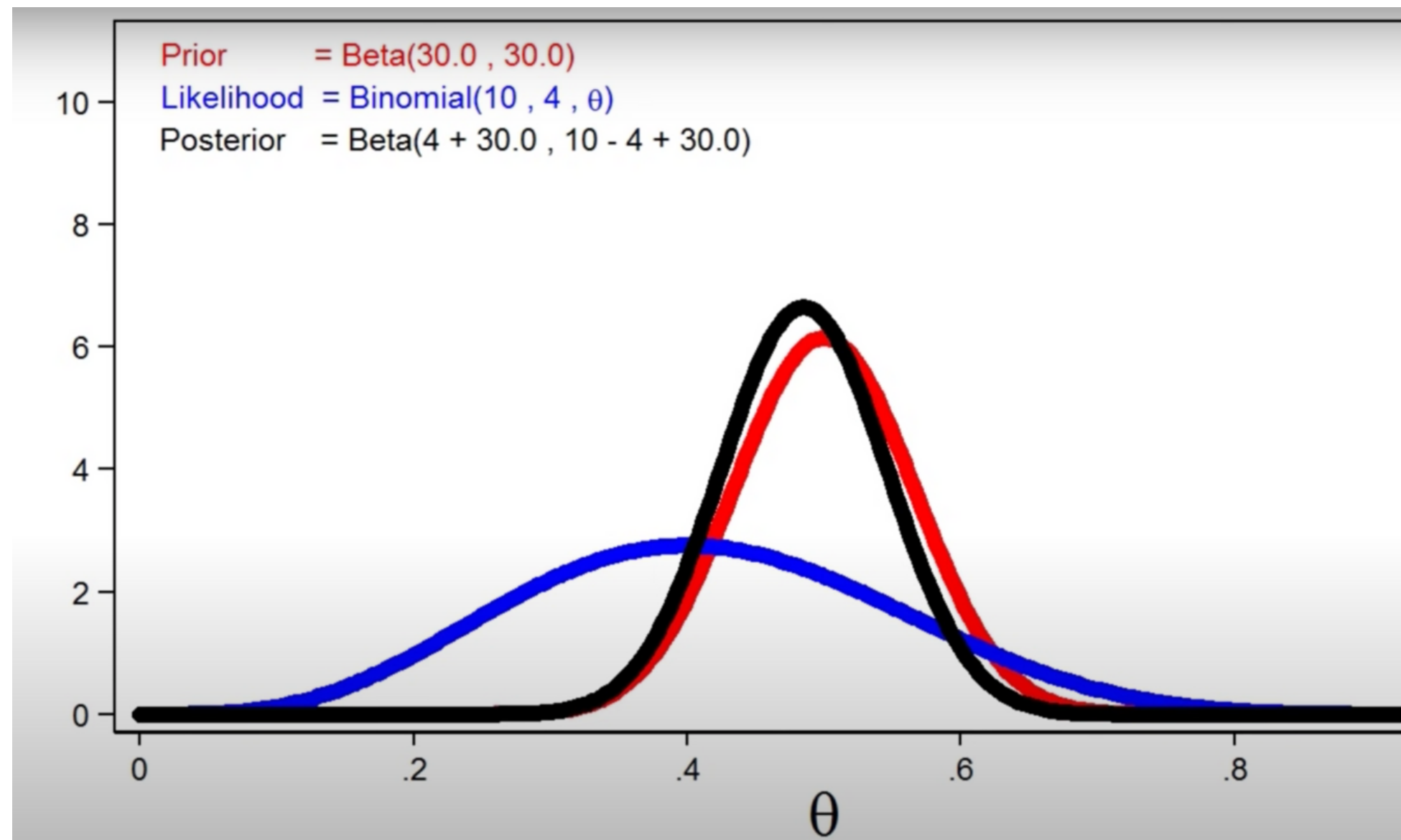


IV FREQUENTIST VERSUS BAYESIAN STATISTICS

BAYESIAN STATISTICS: BEISPIEL MÜNZWURF

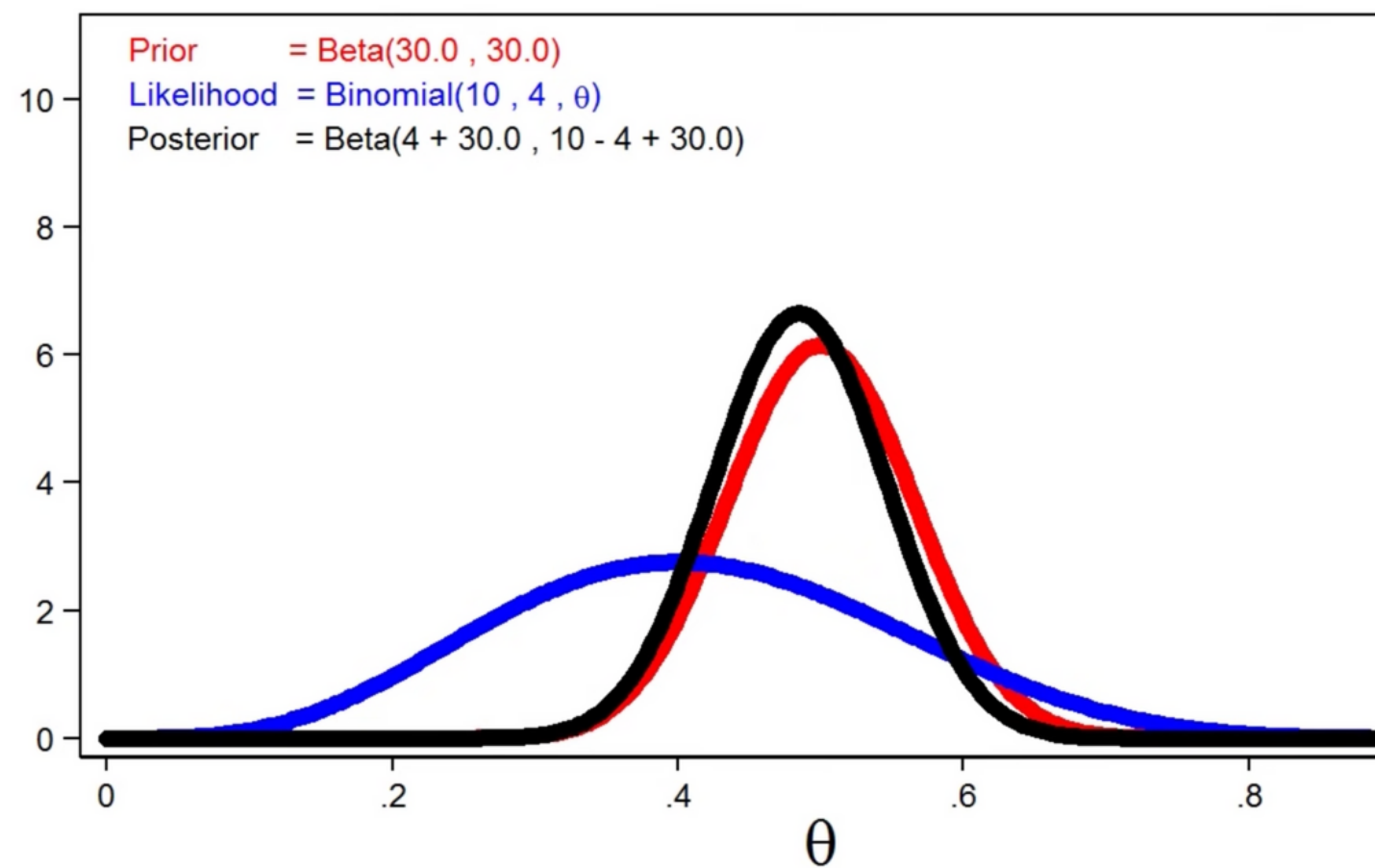


BAYESIAN STATISTICS: BEISPIEL MÜNZWURF

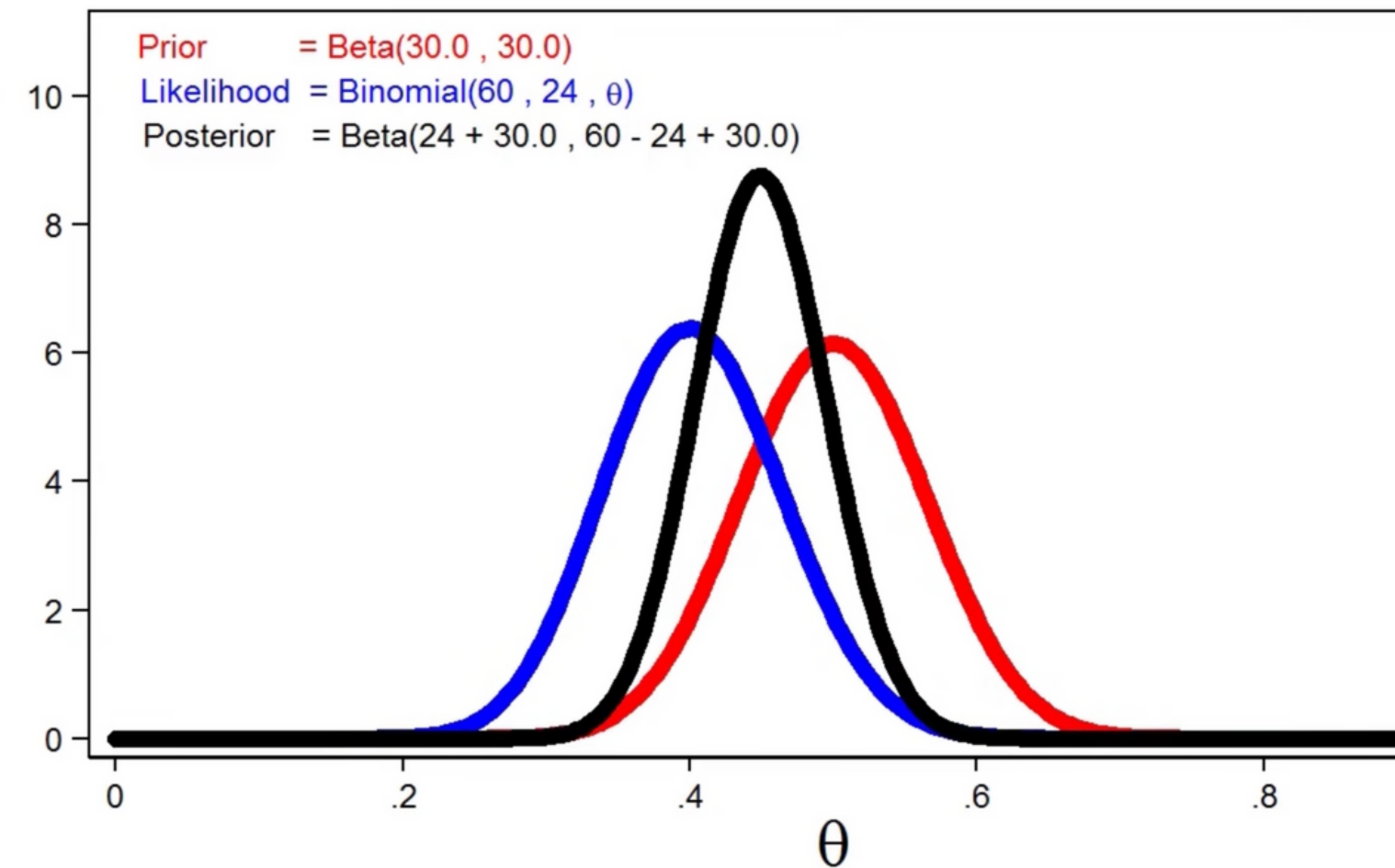


IV FREQUENTIST VERSUS BAYESIAN STATISTICS

BAYESIAN STATISTICS: BEISPIEL MÜNZWURF



N = 10



N = 60

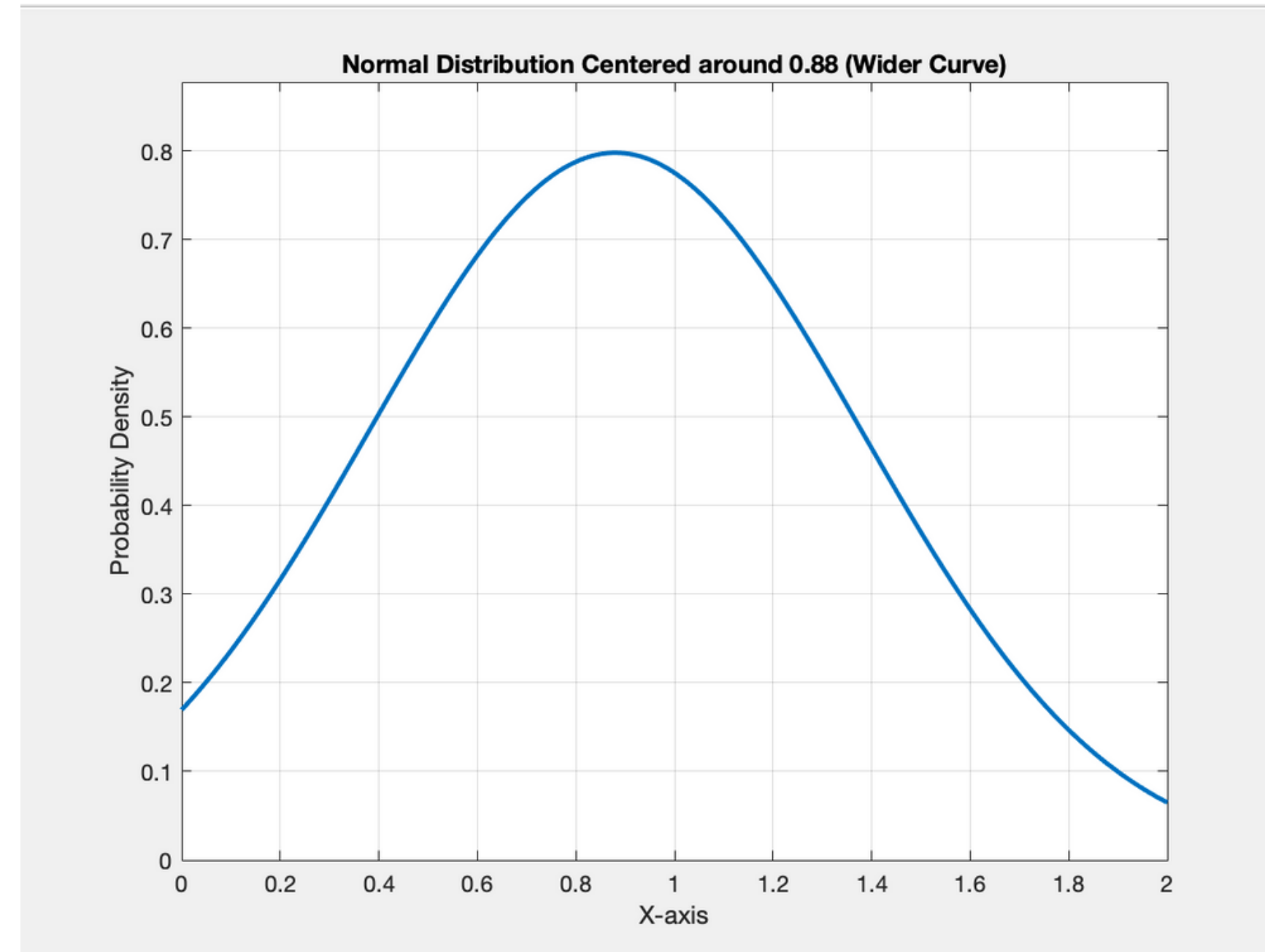
BAYESIAN STATISTICS: BSP. Vohs & Schooler

BEISPIEL: VOHS & SCHOOLER 2008

- Experiment 136
- Manipulation 'Glaube an Freien Willen'
- Lösen von Arithmetische Gleichungen --> Möglichkeit zum Betrug
- N = 30
- ANCOVA: $F(1, 27) = 7.81, p < 0.01, d = 0.88$

<https://osf.io/fgjvw>

Vohs & Schooler, 2008

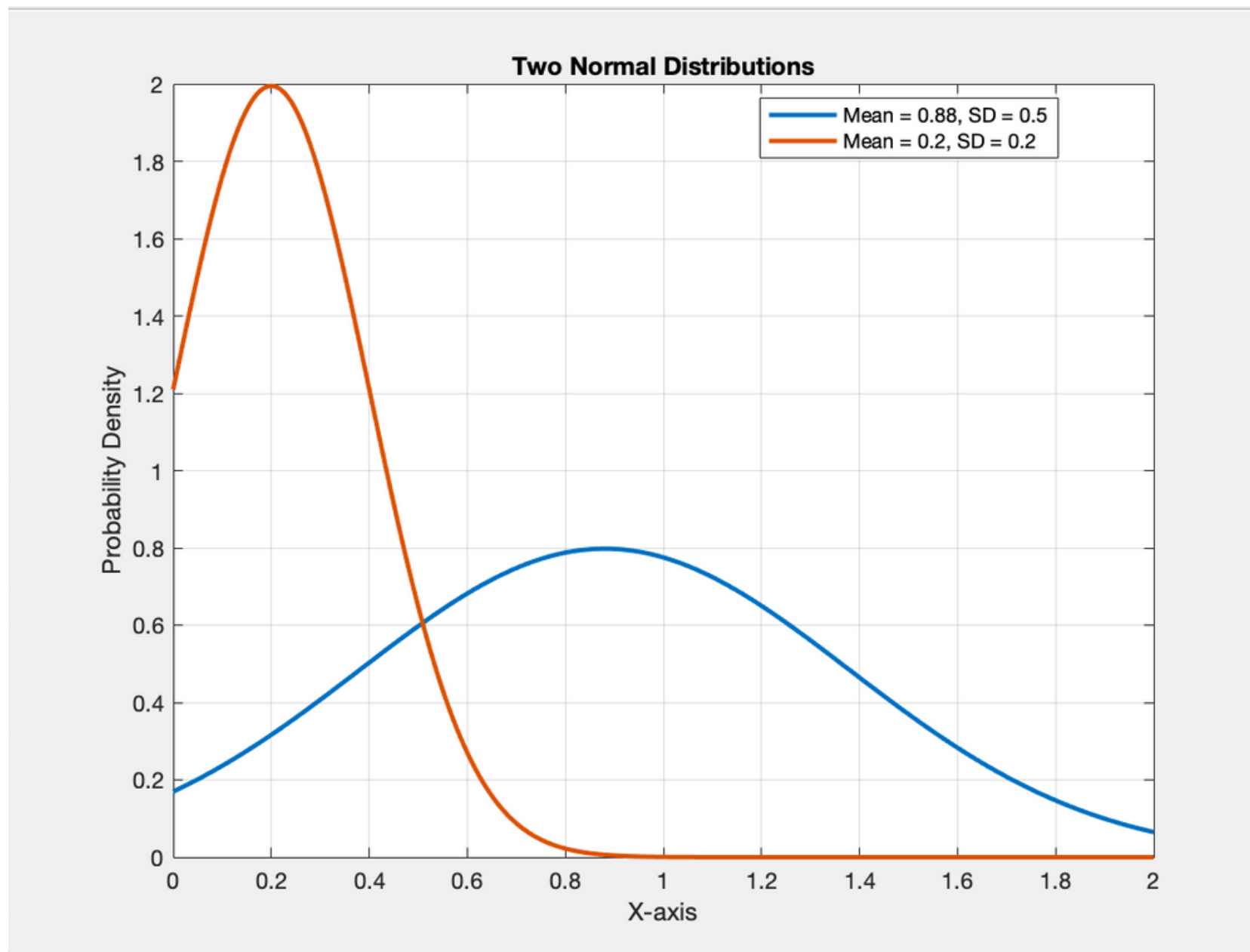


Original : N = 30

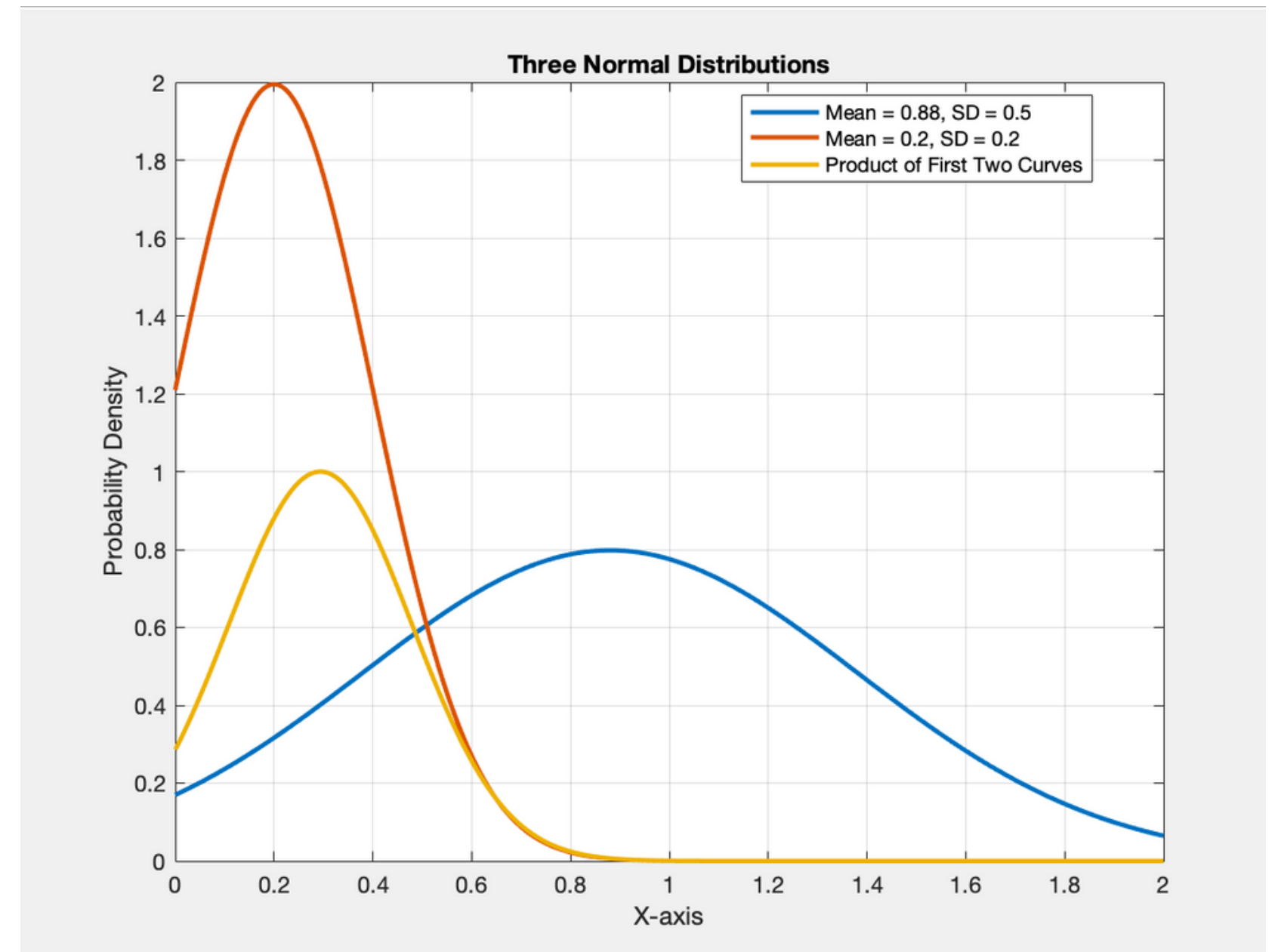
Baldauf, 2023 nach Vohs & Schooler, 2008

IV FREQUENTIST VERSUS BAYESIAN STATISTICS

BAYESIAN STATISTICS: BSP. Vohs & Schooler



Replikation : N = 60



Posterior

IV FREQUENTIST VERSUS BAYESIAN STATISTICS

IM DIREKTEN VERGLEICH

	FREQUENTIST	BAYESIAN
ZENTRALE IDEE	Wahrscheinlichkeit als Häufigkeit	Wahrscheinlichkeit als Grad der Überzeugung
OUTCOME	Ablehnung/ Akzeptanz der Nullhypothese	Wahrscheinlichkeitsverteilung
Parameter	Fix	Varibel; eigene Wahrscheinlichkeitsverteilung

Multiverse Analysen

= stat. Analyse von alle möglichen
(dabei: sinnvollen) Versionen des
Datensatzes

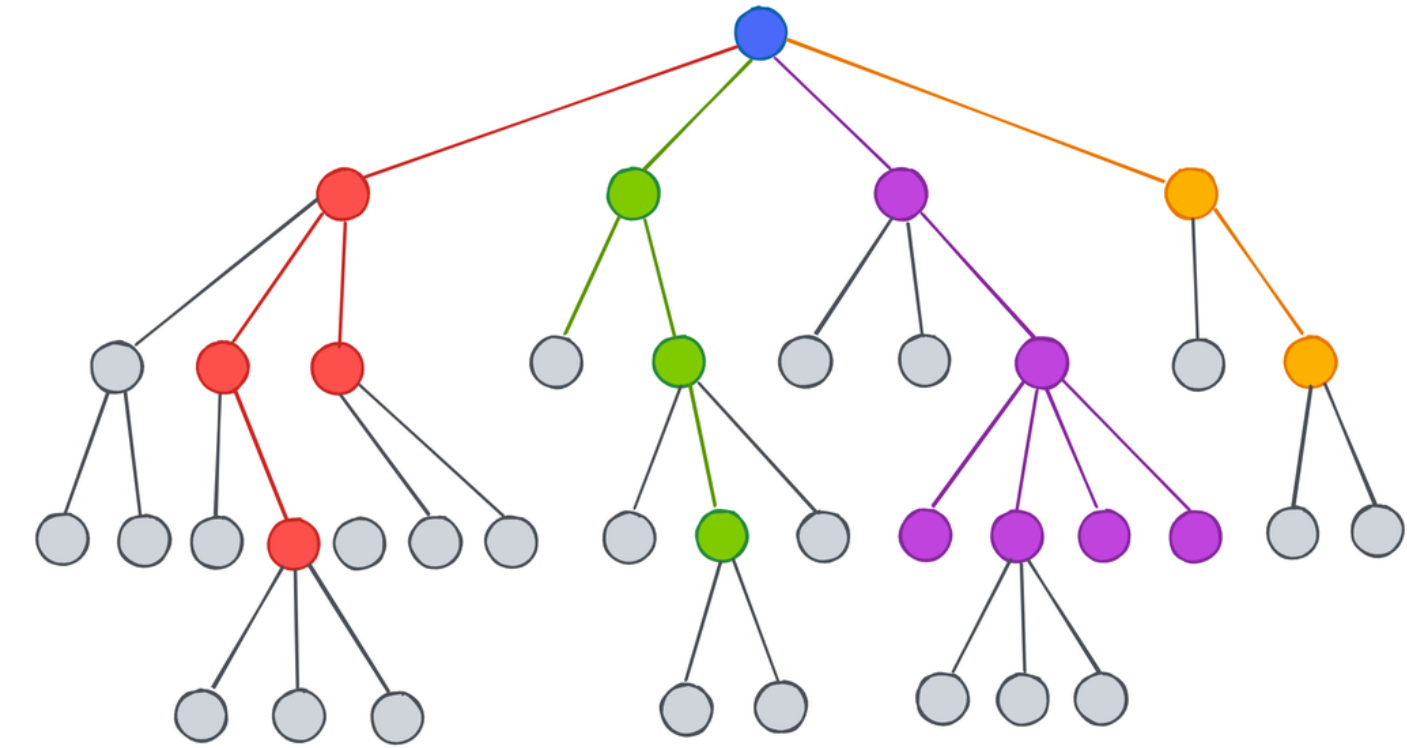
Ziele:

- Stabilität des Effekts
- Identifiziert die zentralen Entscheidungen, von denen der Effekt abhängt



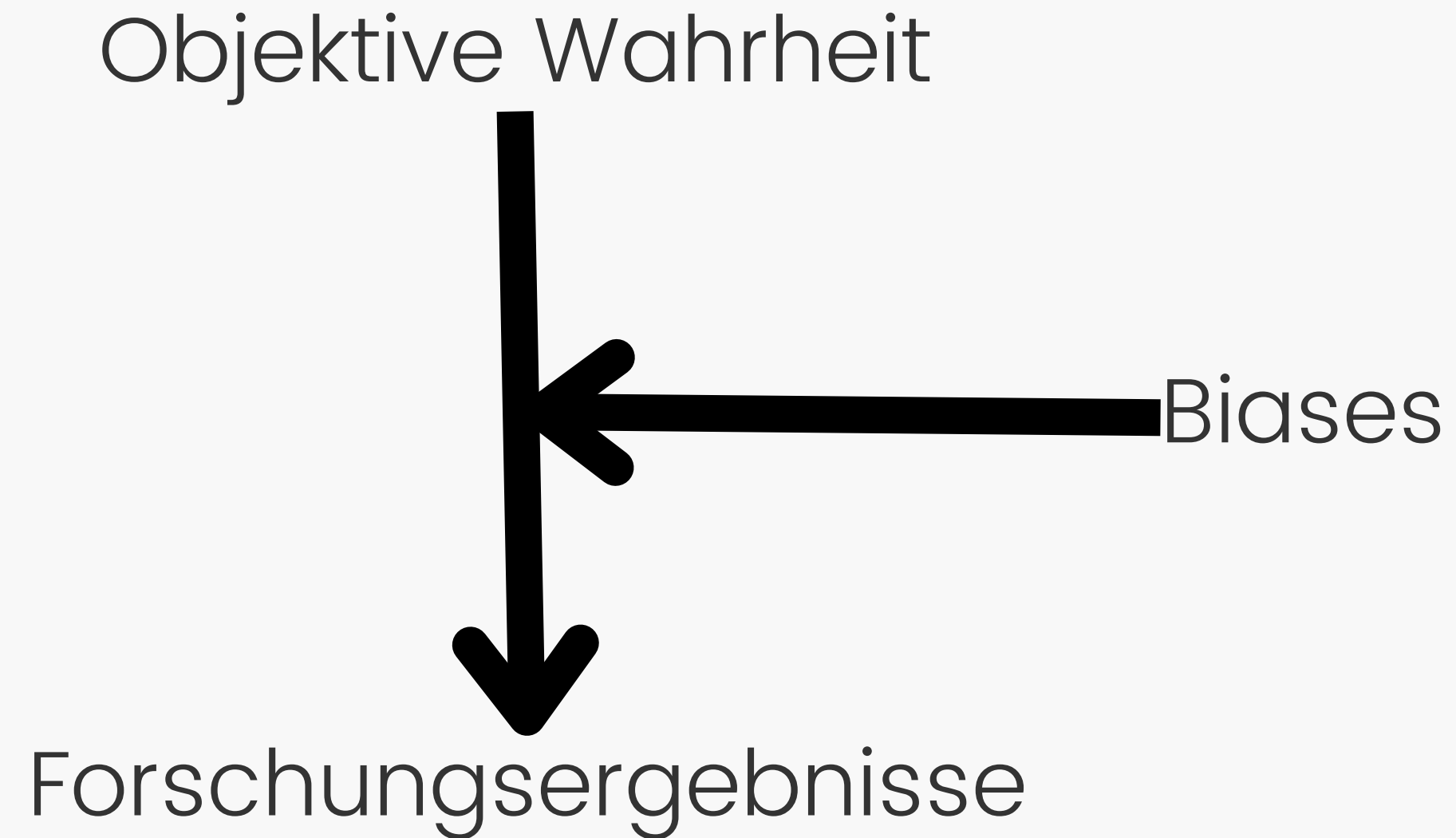
Multiverse Analysen II

1. Konstruktion des Multiversums der Datensätze
2. Statistische Analyse von den verschiedenen Datensätzen

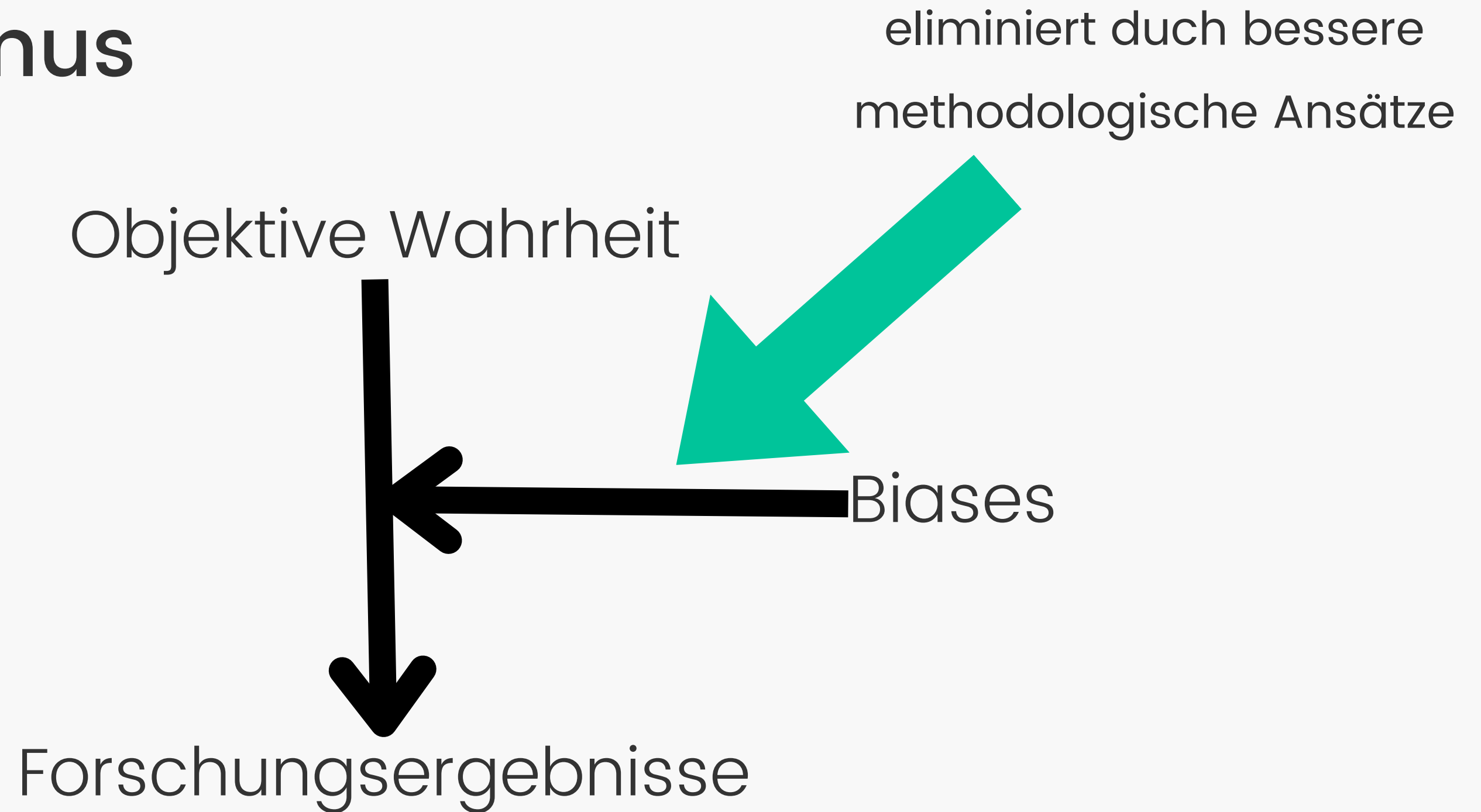


		R1					R2					R3						
		F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	EC1	EC2
		0	0	0.03	0.02	0.08	0.02	0.02	0.03	0.05	0.06	0	0	0.02	0.01	0.04	EC1	ECL1
		0.01	0	0.08	0.03	0.16	0.07	0.05	0.18	0.24	0.41	0.01	0	0.09	0.06	0.23	EC2	ECL2
		0	0	0.06	0.04	0.37	0.02	0.03	0.07	0.08	0.21	0	0	0.05	0.03	0.23	EC1	ECL1
		0.01	0	0.13	0.08	0.44	0.06	0.03	0.22	0.24	0.52	0.01	0	0.14	0.09	0.43	EC2	ECL2
		0	0	0.03	0.01	0.08	0.15	0.07	0.17	0.07	0.14	0.02	0.01	0.06	0.02	0.07	EC1	ECL1
		0	0	0.02	0.01	0.06	0.2	0.05	0.42	0.23	0.44	0.03	0	0.14	0.05	0.19	EC2	ECL2
		0.01	0.01	0.05	0.01	0.1	0.39	0.2	0.45	0.11	0.26	0.08	0.04	0.17	0.03	0.13	EC1	ECL3
		0.01	0	0.05	0.02	0.11	0.33	0.09	0.59	0.26	0.55	0.09	0.02	0.26	0.08	0.27	EC2	ECL1
		0.01	0.01	0.02	0.1	0.28	0.11	0.09	0.43	0.26	0.85	0.02	0.02	0.12	0.12	0.51	EC1	ECL1
		0.01	0.01	0	0.07	0.06	0.07	0.1	0.11	0.14	0.23	0.01	0.02	0.02	0.06	0.08	EC2	ECL2
		0.02	0.01	0.06	0.11	0.36	0.06	0.04	0.3	0.13	0.66	0.02	0.01	0.13	0.07	0.46	EC1	ECL3
		0.02	0.01	0.02	0.15	0.13	0.04	0.05	0.07	0.07	0.16	0.01	0.02	0.03	0.05	0.09	EC2	ECL1
		0.07	0.04	0.12	0.09	0.16	0.16	0.11	0.54	0.32	0.77	0.07	0.04	0.25	0.13	0.39	EC1	ECL2
		0.02	0.02	0.01	0.06	0.02	0.06	0.1	0.07	0.16	0.17	0.02	0.03	0.02	0.07	0.05	EC2	ECL3

Objektivismus

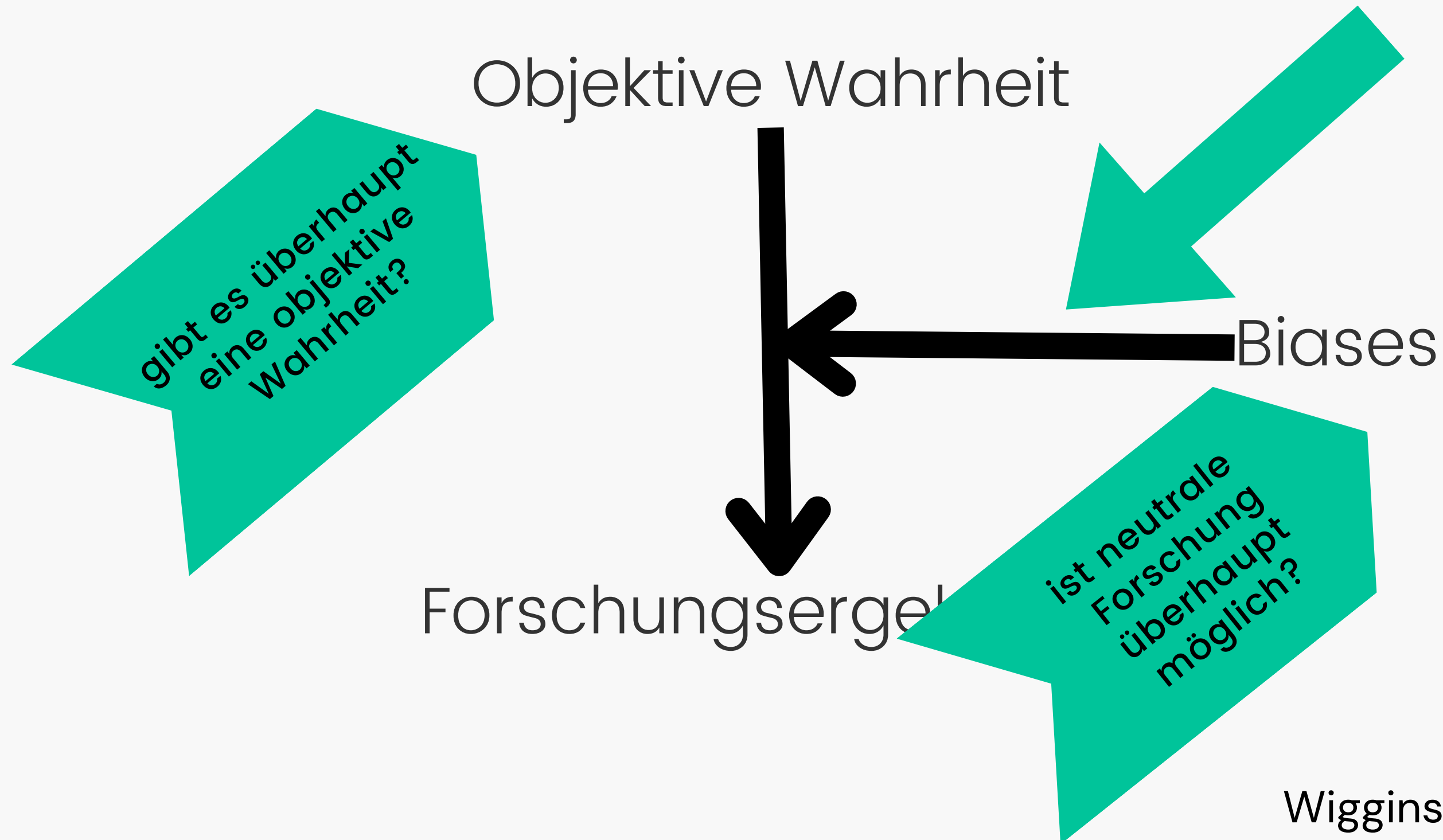


Objektivismus



Objektivismus

eliminiert durch bessere
methodologische Ansätze



Universalismus

= fundamentale (universale) Prinzipien, die über verschiedene Situationen hinweg halten

--> Ziel: universalen Prinzipien mit Theorien festzuhalten

fehlende Replikation aufgrund von Ignoranz gegenüber bestimmten

universalen Prinzipien

Replikation variiert zwangsläufig von Originalstudie --> Unterscheidung zw direkter und konzeptueller Replikation hinfällig?

Experiment as 'Case' = bestimmtes Event in einem ganz bestimmten Kontext

Di, 24.10

Universität Regensburg

OPEN SCIENCE II

Marlena Baldauf

FRAGEN?

QUELLEN

Center for Open Science. (n.d.). Registered reports. <https://www.cos.io/initiatives/registered-reports>

Claesen, A., Gomes, S., Tuerlinckx, F., & Vanpaemel, W. (2021). Comparing dream to reality: An assessment of adherence of the first generation of preregistered studies. *Royal Society Open Science*, 8(10). <https://doi.org/10.1098/rsos.211037>

Edanz-Learning-Team. (2022, September 22). Frequentist vs. bayesian statistics – which should you use?. Edanz Learning Lab. <https://learning.edanz.com/frequentist-bayesian-statistics/>

Fair principles. GO FAIR. (2022, January 21). <https://www.go-fair.org/fair-principles/>

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The Preregistration Revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>

Open Science Working Group at Freie Universität Berlin. (2019, May 16). Open science. Open Science. <https://www.fu-berlin.de/sites/open-science/about-openscience/>

Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016a). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>

Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016b). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>

Volberg, G. (n.d.). Sitzung 05: Replicability crisis und Hypothesentesten. *Geschichte und Methoden der Psychologie*. Regensburg; Universität Regensburg.

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, 39(4), 202–217. <https://doi.org/10.1037/teo0000137>