



Analyse und Dokumentation

BSc Psychologie SoSe 2025

Belinda Fleischmann und Dirk Ostwald

	Gruppe 1/2	Gruppe 3	Format	Thema
1	Do, 10.04.	Fr, 11.04.	Seminar	(1) Quarto, Zotero, Tidyverse
2	Do, 17.04.	Fr, 25.04.	Seminar	(2) Forschungsethik
3	Do, 24.04.	Fr, 02.05.	Seminar	(3) Wissenschaftliche Berichte
4	Mi, 30.04.	Fr, 09.05.	Seminar	(4) Offenheit und Transparenz
5	Do, 08.05.	Fr, 16.05.	Praxisseminar	Offene Übung
6	Do, 15.05.	Fr, 23.05.	Präsentationen	Einfache Lineare Regression
7	Mi, 21.05.	Fr, 30.05.	Präsentationen	Korrelation
8	Do, 05.06.	Fr, 06.06.	Präsentationen	Einstichproben-T-Test
9	Do, 12.06.	Fr, 13.06.	Präsentationen	Zweistichproben-T-Test
10	Do, 19.06.	Fr, 20.06.	Präsentationen	Einfaktorielle Varianzanalyse
11	Do, 26.06.	Fr, 27.06.	Präsentationen	Zweifaktorielle Varianzanalyse
12	Do, 03.07.	Fr, 04.07.	Präsentationen	Multipl. Regression
13	Do, 10.07.	Fr, 11.07.	Präsentationen	Kovarianzanalyse
Juli			Klausurtermin	

(4) Offenheit und Transparenz

Motivation

Methodentransparenz und Datentransparenz

Selbstkontrollfragen

Motivation

Methodentransparenz und Datentransparenz

Selbstkontrollfragen

Motivation

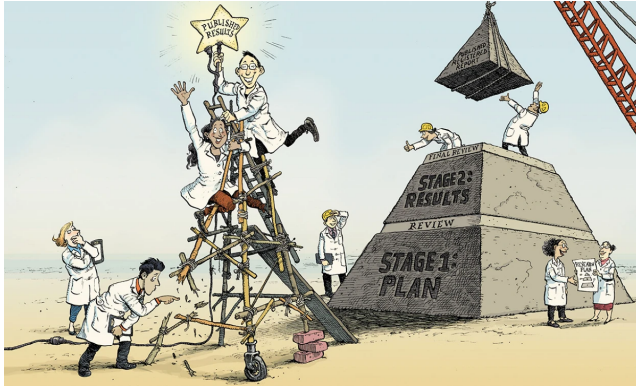


Illustration by David Parkins

Chambers (2019)

Psychologie – 64 von 100 Studien nicht replizierbar

- Open Science (2015)

Nulleffekte in preregistrierten Replizierbarkeitsstudien

- Hagger u. a. (2016), Wagenmakers u. a. (2017), O'Donnell u. a. (2018)

Vehaltensökonomie – 7 von 18 Studien nicht replizierbar

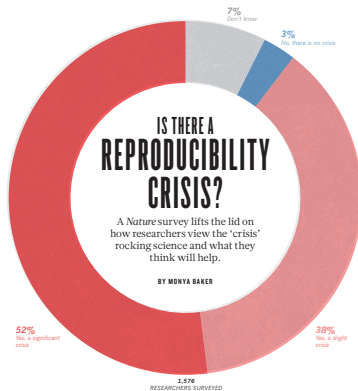
- Camerer u. a. (2016)

Biologie – 47 von 53 Studien nicht replizierbar

- Begley und Ellis (2012)

Reproduzierbarkeitskrise in Machine-Learning-basierter Wissenschaft

- Kapoor und Narayanan (2022)



Baker (2016)

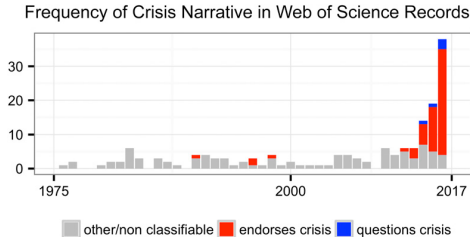
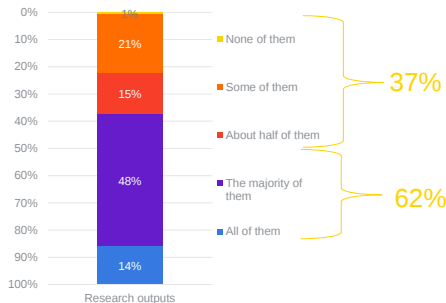


Fig. 1. Number of Web of Science records that in the title, abstract, or keywords contain one of the following phrases: "reproducibility crisis," "scientific crisis," "science in crisis," "crisis in science," "replication crisis," "replicability crisis." Records were classified by the author according to whether, based on title and abstracts, they implicitly or explicitly endorsed the crisis narrative described in the text (red), or alternatively questioned the existence of such a crisis (blue), or discussed "scientific crises" of other kinds or could not be classified due to insufficient information (gray). The complete dataset, which includes all titles and abstracts and dates back to the year 1933, is available in [Dataset S1](#). This sample is merely illustrative, and does not include the numerous recent research articles and opinion articles that discuss the "science is in crisis" narrative without including any of the above sentences in the title, abstract, or keywords.

Das Replizierbarkeitskrisennarrativ

Thinking about the various research outputs that you interacted with (or encountered) last week what proportion of the outputs would you consider trustworthy?



Base: All respondents (n=3133)

<https://senseaboutscience.org/wp-content/uploads/2019/09/Quality-trust-peer-review.pdf>

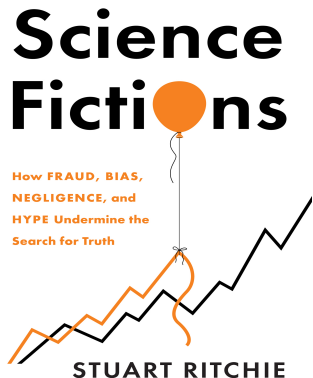
Das Replizierbarkeitskrisennarrativ

Which of the following mechanisms do you employ to compensate for any lack of confidence you have in the content you are considering reading/accessing?



Base: All respondents that do not think all research outputs are trustworthy (n=2715)

<https://senseaboutscience.org/wp-content/uploads/2019/09/Quality-trust-peer-review.pdf>



Ritchie (2020)

Motivation

Methodentransparenz und Datentransparenz

Selbstkontrollfragen

Ausgangslage

- Academia ist mehr an spannenden Resultaten als an sorgfältiger Arbeit interessiert.
- Academia interessiert sich wenig für die eigentliche Arbeit, sondern die Resultate.
- Academia liebt komplexe Analysen ohne den Anspruch, sie verstehen zu wollen.
- Academia mag keine Nullergebnisse.

Konsequenzen

- Publierte Resultate sind oft nicht reproduzierbar und aufbaufähig.
- Ressourcenverschwendung.
- Geringes innerakademisches Vertrauen in publizierte Resultate.
- Gefahr des geringen extraakademischen Vertrauens in publizierte Resultate.

Lösung

- Gewährleistung von Reproduzierbarkeit durch Methoden- und Datentransparenz.

Replizierbarkeit

Andere Wissenschaftler:innen nutzen andere (neue) Daten und potentiell andere Methoden und kommen zur gleichen Schlußfolgerung wie eine Originalstudie.

⇒ Abhängig von den wahren, aber unbekannten, Parametern.

⇒ Unbeeinflussbar

Reproduzierbarkeit

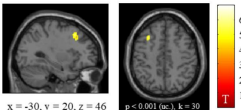
Andere Wissenschaftler:innen nutzen die gleichen Daten und die gleichen Datenanalysemethoden und erhalten die gleichen Resultate (cf. Peng (2011)).

⇒ Abhängig von der Sorgfältigkeit der wissenschaftlichen Arbeit

⇒ Beeinflussbar

⇒ Menschen machen Fehler

Wissenschaftlicher Bericht



Left medial frontal cortex

T = 6.32, MNI: x = -30, y = 20, z = 46

Voxels-level: $p_{\text{FWE-corr}} < 0.001$, $p_{\text{FWE-corr}} = 0.991$

Cluster-level: $p_{\text{FWE-corr}} = 0.003$, $p_{\text{FWE-corr}} = 0.062$

Figure 4. Exemplary application of the EFT-based power, PPV, and sample size calculation framework. (A) The upper panel depicts the results of a perceptual decision-making pilot study with $n = 10$ participants for contrasting perceptual clusters based on high and low visual sensory evidence. The T-values from the identified cluster in the left medial frontal gyrus were averaged to obtain a raw effect size estimate, which was then adjusted based on the effect size bias estimate reported in Figure 7 of *Geiser et al. (2018)* and reproduced in the lower subpanel of panel (A). (B) Sample size calculations for voxel-level minimal and maximal power and PPV based on the effect size estimate of the pilot fMRI study. (C) Sample size calculations for cluster-level minimal and maximal power and PPV based on the effect size estimate of the pilot fMRI study. For implementation details, please see *efpy_figure_4.m*.

revealed a cluster of activity in the left medial frontal gyrus, as shown in the upper panel of Figure 4A (for further details about the experimental and data-analytical procedures, please see Supplement S.3). Our aim was to use the effect size estimate derived from this cluster to calculate the sample sizes necessary to achieve minimal and maximal power and PPV levels of 0.8 for corrected voxels, and cluster-level inference at a significance level of $\alpha_{\text{FWE}} = 0.05$, a partial alternative hypothesis parameter of $\lambda = 0.1$, and a prior hypothesis parameter of $\tau = 0.2$. To this end, we evaluated the average T-values of the cluster, yielding $T = 4.63$, which translates into an effect size estimate of $d = 4.63/\sqrt{10} = 1.47$. However, it is well known that effect size estimates resulting from the thresholding of mass-invariant statistical parameters maps exhibit biases (e.g., Vil et al., 2009; Pollock et al., 2017). To correct our effect size estimate for this bias, we capitalized on recent results by *Geiser et al. (2018)*, which are depicted in the lower panel of Figure 4A. Specifically, using task-related fMRI data from the Human Connectome Project 500 (Van Essen et al., 2013), *Geiser et al. (2018)* estimated the effect size bias exhibited by activations detected in random data subsets of 50 to 100 participants from the approximately 500 participants. As presented in Figure 7A of *Geiser et al. (2018)* and visualized in the lower

Online Repository

Roh- und Metadaten
Datenanalyseskripte
Datenakquisitionskripte

...

```
function rft2
% This function implements the sampling of Bayesian random effects with
% Gaussian covariance functions as visualizer in Figure 1 of Oswald et al.
% (2018) for parameter values
%
% Author: Dirk Ostwald
%
% -----
% Initialization
%
% CSD
% cluster size
%
% random number generator setting for reproducible results
rng(1)
%
% utilities
%
% sub-005_con_0001.nii
%
% figure management
%
% figure = fullfile('figure_dir', sprintf('%04d', sub_id))
%
% discrete domain D = [x1, x2, ..., xN] x [y1, y2, ..., yN]
%
% N = 10
%
% M = 10
%
% sub_id = 5
%
% D_sub = [x1, x2, ..., xN] x [y1, y2, ..., yN]
%
% D = [x1, x2, ..., xN] x [y1, y2, ..., yN]
%
% D = [x1, x2, ..., xN] x [y1, y2, ..., yN]
%
% -----
% sub-005_con_0001.nii 2019-03-26 01:51 PM
% sub-006_con_0001.nii 2019-03-26 01:51 PM
% sub-007_con_0001.nii 2019-03-26 01:51 PM
% sub-008_con_0001.nii 2019-03-26 01:51 PM
% sub-009_con_0001.nii 2019-03-26 01:51 PM
% sub-010_con_0001.nii 2019-03-26 01:51 PM
```

Hyperlink



Reproduktion

Wissenschaftliche Qualitätssicherung

- Pre- und Postpublikations-Mehraugenprinzip
- Reproduzierbarkeit bei identischen Daten und Analysen gesichert
- Replikationspotential für neue Daten und Analysen erhöht

Nachhaltigkeit

- Möglichkeit der Datenwiederverwendung in neuen Kontexten
- Maximierung der Anzahl experimenteller Einheiten
- Ermöglichung automatisierter Metaanalysen

Ethischer Imperativ

- Respekt im Umgang mit öffentlichen Geldern
- Respekt für die Beiträge von Forschungspersonen
- Minimierung von Forschungsrisiken

Maßnahmen

- Computercode als ausführlichste Form der Dokumentation
- Referenzierung datenanalytischer Skripte in Methods und Results Sections
- Verfügbarkeit zum Zeitpunkt der Veröffentlichung (Preprint/Journal Submission)

Voraussetzungen

- Automatisierte Datenanalyse von Rohdaten zu Abbildungen und Tabellen
- Lesbarer, gut-strukturierter und extensiv kommentierter Code
- Verzicht auf Variablennamen mit persönlichkeitsrelevanten Merkmalen

Maßnahmen

- Digitale Rohdaten als ausführlichste Form der Datengrundlage
- Formatierung anhand akzeptierter Datenstandards
- Verfügbarkeit zum Zeitpunkt der Veröffentlichung (Preprint/Journal Submission)

Voraussetzungen

- Vorliegen digitaler Rohdaten
- Verfügbarkeit von Speicherplatz
- Respekt datenschutzrechtlicher Belange

Open Science Framework

- Nichtkommerzielles Datenrepositorium
- Projektarchivierung bei Preprint/Journal Submission/Revision
- Seit 2020 beschränkter Speicherplatz pro Projekt

GitHub

- Kommerzieller Dienst zur Versionsverwaltung für Software-Entwicklungsprojekte.
- Kollaborative Softwareentwicklung (Track Changes)
- Oft irrtümlich zur Analysecodearchivierung benutzt

Forschungsdatenzentrum des ZPID

- Digitales Forschungsdatenzentrum für die psychologische Forschung
- Teil der öffentlich finanzierten Leibniz-Gemeinschaft
- Unterstützung bei Dokumentation und Archivierung als kostenpflichtige Services

Brain Imaging Data Structure

- Datenstandard für Neuroimaging- und Verhaltensdaten
- Stetige Weiterentwicklung auf weitere Datenmodalitäten
- Gorgolewski u. a. (2016)

PsyCuraDat

- Datenstandard für psychologische Daten
- Zur Zeit noch in der Entwicklung
- Blask, Gerhards, und Jalynskij (2021)

Psychologische Daten sind in aller Regel sensible humane Daten

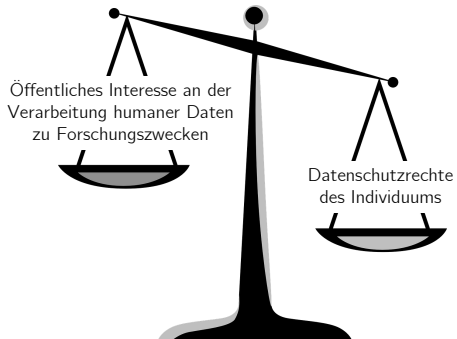
Datenschutzrechtliche Belange müssen respektiert werden

Das Datenschutzrecht gliedert sich in

- EU Datenschutzgrundverordnung ([DSGVO](#)) [EU2016/679]
- Bundesdatenschutzgesetz ([BDSG](#))
- Datenschutzgesetze der Bundesländer

Psychologische Daten fallen in aller Regel unter personenbezogene Daten

“Alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann”



Möglichkeiten der Datentransparenz beim Umgang mit humanen Daten

Public sharing

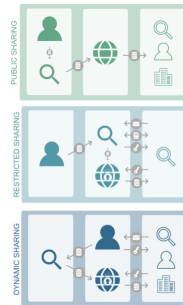
- Open data platforms

Restricted sharing

- Peer-to-peer data sharing

Dynamic sharing

- Participant-centred platforms



Public Sharing

Datenbereitstellung auf weltweit frei zugänglichen Datenplattformen

- Allgemeiner Datenzugang für alle Formen der Datennutzung

Vorteile

- Einfach
- Etabliert
- Ressourcen sensibel

Nachteile

- Unwiderruflich
- Gefahr des Datenmissbrauchs
- Unklarer datenschutzrechtlicher Stand (DSGVO)

Restricted sharing

- Datenzugangsregelung mithilfe von Data Use Agreements

Vorteile

- Risikominimierung
- Klare Verantwortlichkeiten
- Teilweise etabliert

Nachteile

- Restriktiv
- Verantwortlichkeit auf Seiten der Wissenschaftler:innen
- Paternalistisch

Dynamic sharing

- Proband:innenzentrierte Datenplattformen
- Datenzugangsregulation und Monetarisierung durch Proband:innen

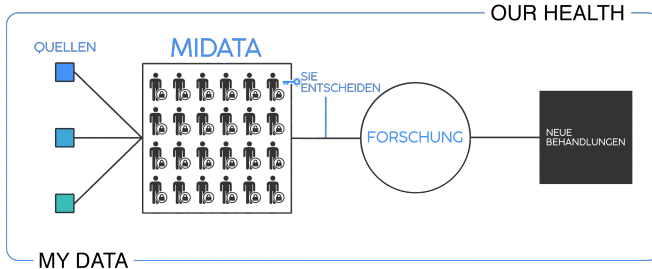
Vorteile

- Digitales Empowerment
- Klare Verantwortlichkeiten
- Zukunftsorientiert

Nachteile

- Drop-out
- Nicht etabliert
- Ressourceintensiv

Dynamic sharing





Pillars of Open Science, UNESCO (2021)

Motivation

Methodentransparenz und Datentransparenz

Selbstkontrollfragen

Selbstkontrollfragen

1. Begründen Sie die Wichtigkeit, die Reproduzierbarkeit von Forschung durch Methoden- und Datentransparenz zu fördern.
2. Erläutern und differenzieren Sie die Begriffe Replizierbarkeit und Reproduzierbarkeit.
3. Nennen und erläutern Sie den Nutzen von Methoden- und Datentransparenz.
4. Nennen Sie die Kernaspekte und Voraussetzungen für Methodentransparenz.
5. Nennen Sie die Kernaspekte und Voraussetzungen für Datentransparenz.
6. Nennen Sie drei Online-Repositorien, die in der Forschung verwendet werden.
7. Nennen Sie zwei Datenstandards, die für die psychologische Forschungs relevant sind.
8. Diskutieren Sie die Vor- und Nachteile verschiedener Formen von Data Sharing.

Referenzen

- Baker, Monya. 2016. „1,500 Scientists Lift the Lid on Reproducibility". *Nature* 533 (7604): 452–54. <https://doi.org/10.1038/533452a>.
- Begley, C. Glenn, und Lee M. Ellis. 2012. „Raise Standards for Preclinical Cancer Research". *Nature* 483 (7391): 531–33. <https://doi.org/10.1038/483531a>.
- Blask, Katarina, Lea Gerhards, und Maria Jalynskij. 2021. „PsyCuraDat: Designing a User-Oriented Curation Standard for Behavioral Psychological Research Data". *Frontiers in Psychology* 11 (Januar): 579397. <https://doi.org/10.3389/fpsyg.2020.579397>.
- Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, u. a. 2016. „Evaluating Replicability of Laboratory Experiments in Economics". *Science* 351 (6280): 1433–36. <https://doi.org/10.1126/science.aaf0918>.
- Chambers, Chris. 2019. „What's Next for Registered Reports?". *Nature* 573 (7773): 187–89. <https://doi.org/10.1038/d41586-019-02674-6>.
- Fanelli, Daniele. 2018. „Is Science Really Facing a Reproducibility Crisis, and Do We Need It To?". *Proceedings of the National Academy of Sciences* 115 (11): 2628–31. <https://doi.org/10.1073/pnas.1708272114>.
- Gorgolewski, Krzysztof J., Tibor Auer, Vince D. Calhoun, R. Cameron Craddock, Samir Das, Eugene P. Duff, Guillaume Flandin, u. a. 2016. „The Brain Imaging Data Structure, a Format for Organizing and Describing Outputs of Neuroimaging Experiments". *Scientific Data* 3 (Juni): 160044. <https://doi.org/10.1038/sdata.2016.44>.
- Hagger, M. S., N. L. D. Chatzisarantis, H. Alberts, C. O. Anggono, C. Batailler, A. R. Birt, R. Brand, u. a. 2016. „A Multilab Preregistered Replication of the Ego-Depletion Effect". *Perspectives on Psychological Science* 11 (4): 546–73. <https://doi.org/10.1177/1745691616652873>.
- Kapoor, Sayash, und Arvind Narayanan. 2022. „Leakage and the Reproducibility Crisis in ML-Based Science". arXiv. <https://arxiv.org/abs/2207.07048>.
- O'Donnell, Michael, Leif D. Nelson, Evi Ackermann, Balazs Aczel, Athfah Akhtar, Silvio Aldrovandi, Nasseem Alshaif, u. a. 2018. „Registered Replication Report: Dijksterhuis and van Knippenberg (1998)". *Perspectives on Psychological Science* 13 (2): 268–94. <https://doi.org/10.1177/1745691618755704>.
- Open Science, Collaboration. 2015. „Estimating the Reproducibility of Psychological Science". *Science* 349 (6251): aac4716. <https://doi.org/10.1126/science.aac4716>.
- Peng, Roger D. 2011. „Reproducible Research in Computational Science". *Science* 334 (6060): 1226–27. <https://doi.org/10.1126/science.1213847>.
- Ritchie, Stuart. 2020. *Science Fictions - Exposing Fraud, Bias, Negligence and Hype in Science*. Penguin.
- Wagenmakers, Eric-Jan, Josine Verhagen, Alexander Ly, Dora Matzke, Helen Steingrover, Jeffrey N. Rouder, und Richard D. Morey. 2017. „The Need for Bayesian Hypothesis Testing in Psychological Science". In *Psychological Science Under Scrutiny*, herausgegeben von Scott O. Lilienfeld und Irwin D. Waldman, 123–38. Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119095910.ch8>.