

Reproduzierbarkeit: Das Fundament der Wissenschaft erklärt

Autor: Wolfram Klatt

Veröffentlicht am: 10.08.2026

Dokument generiert über Gesundheitsmythen.de

ZUSAMMENFASSUNG (ABSTRACT)

Reproduzierbarkeit beschreibt das Prinzip, dass wissenschaftliche Experimente von unabhängigen Forschern unter gleichen Bedingungen wiederholt werden können und identische Ergebnisse liefern. Ohne Reproduzierbarkeit basieren Behauptungen auf Zufall oder Irrtum. Die Replikationskrise zeigt die Notwendigkeit von Open Science und Vorabregistrierung.

Reproduzierbarkeit: Das Fundament der wissenschaftlichen Wahrheit

1. Definition und Grundlagen: Das Gerüst des Wissens

In der Architektur des menschlichen Wissens ist die **Reproduzierbarkeit** (auch Nachvollziehbarkeit oder Wiederholbarkeit genannt) das absolute Fundament. Ohne sie ist die Naturwissenschaft kein systematisiertes Wissen, sondern lediglich eine Sammlung von Anekdoten, Behauptungen und Zufällen. Reproduzierbarkeit beschreibt das Prinzip, dass die Ergebnisse eines wissenschaftlichen Experiments oder einer Studie unter den gleichen

Bedingungen von unabhängigen Forschern erneut erzielt werden können.

Historisch gesehen wurde dieses Prinzip im 17. Jahrhundert von Pionieren wie Robert Boyle etabliert. Als er seine berühmten Experimente mit der Vakuumpumpe durchführte, bestand er darauf, dass diese vor Zeugen stattfinden und dass die Apparaturen exakt dokumentiert werden. Sein Ziel war es, dass jeder Forscher in Europa dasselbe Experiment nachbauen und den Effekt selbst beobachten konnte.

Um präzise zu sein, unterscheidet die moderne Wissenschaftstheorie (insbesondere in den Natur- und Datenwissenschaften) oft zwischen drei verwandten Konzepten:

- **Wiederholbarkeit (Repeatability):** Ein und dasselbe Forscherteam führt dasselbe Experiment im selben Labor mit demselben Setup mehrfach durch und erhält konsistente Ergebnisse. Dies schließt bloße Messfehler oder Einmaleffekte aus.
- **Reproduzierbarkeit im engeren Sinn (Computational Reproducibility):** Ein unabhängiges Forscherteam nimmt die exakten Rohdaten und den Auswertungscode der Originalstudie und kommt zu denselben statistischen Ergebnissen. Dies prüft, ob bei der Datenanalyse Fehler oder unzulässige Manipulationen stattgefunden haben.
- **Replizierbarkeit (Replicability):** Das ist der „Goldstandard“. Ein völlig unabhängiges Forscherteam baut das Experiment basierend auf der publizierten Methodik (dem „Rezept“) in einem anderen Labor mit neuen Probanden oder Proben nach. Wenn die neuen Daten die ursprüngliche Schlussfolgerung stützen, gilt der Effekt als verlässlich repliziert.

Die Metapher des Kompasses und der Landkarte

Ein Experiment ist wie ein Kompass, der uns in einem unbekannten, dunklen Gebiet eine Richtung anzeigt. Wenn ein einziger Forscher behauptet, der Kompass zeige nach Norden, könnte das Instrument defekt sein oder durch ein lokales Magnetfeld (einen Störfaktor) abgelenkt werden. Reproduzierbarkeit bedeutet, dass unabhängige Forscher mit ihren eigenen Kompassen an exakt dieselbe Stelle treten. Erst wenn alle Nadeln unabhängig voneinander in dieselbe Richtung weisen, können wir sicher sein, dass wir auf Basis dieser Daten eine verlässliche Landkarte zeichnen dürfen.

Das Leitmotiv der ältesten wissenschaftlichen Gesellschaft der Welt, der Royal Society, lautet „*Nullius in verba*“ – nach den Worten keines Einzigen. Wissenschaftliche Erkenntnis basiert nicht auf der Autorität, dem Charisma oder dem Prestige der Person, die etwas behauptet,

sondern einzig und allein auf dem objektiv messbaren, durch unabhängige Dritte überprüfbaren Beweis. Ein einmaliges, nicht reproduzierbares Ergebnis, sei es noch so spektakulär, ist für die Wissenschaft im besten Fall eine interessante Hypothese – und im schlechtesten Fall ein Messfehler.

2. Der naturwissenschaftliche Mechanismus und die Methodik

Wie stellt die Wissenschaft sicher, dass Ergebnisse kein Zufall sind? Der Mechanismus der Reproduzierbarkeit basiert auf harter Mathematik, strikter Methodik und radikaler Transparenz.

In der Naturwissenschaft messen wir nie mit absoluter, unendlicher Präzision. Jedes Experiment beinhaltet sogenanntes "Rauschen" – biologische Varianz in menschlichen Zellen, winzige Temperaturschwankungen, Ungenauigkeiten der Geräte oder das grundlegende Rauschen der Statistik. Um ein echtes Signal von diesem Rauschen zu unterscheiden, nutzen Forscher die Inferenzstatistik.

Ein zentrales Problem der Reproduzierbarkeit liegt im sogenannten **Alpha-Fehler (Fehler 1. Art oder falsch-positives Ergebnis)**. In den meisten naturwissenschaftlichen und medizinischen Studien gilt ein Ergebnis traditionell als „statistisch signifikant“, wenn der p-Wert unter 0,05 liegt. Das bedeutet vereinfacht gesagt: Wenn es den behaupteten Effekt in Wirklichkeit gar nicht gäbe (z. B. wenn ein Medikament völlig wirkungslos ist), wäre die Wahrscheinlichkeit, ein solches oder noch extremeres positives Ergebnis durch reinen Zufall zu erhalten, kleiner als 5 Prozent.

Das bedeutet aber im Umkehrschluss: Wenn Forscher weltweit 100 Experimente zu Behandlungen durchführen, die in Wahrheit absolut wirkungslos sind, werden rein durch den statistischen Zufall etwa 5 dieser Experimente ein scheinbar hochwirksames Ergebnis („Signifikanz“). Dies ist kein Betrug, sondern ein unabänderliches Gesetz der Stochastik.

In den letzten Jahrzehnten geriet die Wissenschaft in die sogenannte **Replikationskrise**. Forscher stellten fest, dass sich viele berühmte Studien aus der Psychologie, der Biomedizin und der Krebsforschung nicht reproduzieren ließen. Die Mechanismen dahinter waren meist nicht bewusster Betrug, sondern systemische methodische Fehler, die oft aus einem immensen Publikationsdruck ("Publish or Perish") resultierten:

1. **P-Hacking:** Forscher filtern ihre Daten so lange (z. B. durch das Weglassen unliebsamer Datensätze, das Hinzufügen weiterer Probanden nach Zwischenanalysen oder das Testen unzähliger Untergruppen), bis der p-Wert magischerweise unter die geforderte Grenze von 0,05 rutscht.
2. **HARKing (Hypothesizing After the Results are Known):** Forscher suchen in großen, unübersichtlichen Datensätzen nach zufälligen Mustern. Finden sie eines, behaupten sie im Nachhinein, sie hätten genau dieses Muster von Anfang an erwartet und theoretisch begründet.
3. **Publication Bias (Publikationsverzerrung):** Fachjournale veröffentlichen bevorzugt Studien, die einen aufregenden, neuen Effekt zeigen (positive Resultate). Studien, die belegen, dass etwas *nicht* funktioniert (Null-Resultate), landen oft unveröffentlicht in der Schublade. Wenn 20 Teams ein unwirksames Medikament testen, erhalten 19 ein negatives Ergebnis (das nie publiziert wird) und 1 Team durch Zufall ein falsch-positives Ergebnis, das sogleich mit großem Trommelwirbel gedruckt wird.

VISUALISIERUNG: DAS PROBLEM BEIM HARKing (Der texanische Scharfschütze)

1. Daten wahllos erheben



2. Zufälliges Muster finden



3. Zielscheibe

3. Lösungsansätze: Open Science und Vorabregistrierung

Um diesen Makeln der menschlichen Kognition zu begegnen, hat die Wissenschaft Abwehrmechanismen entwickelt, die heute unter dem Begriff **Open Science** gebündelt werden.

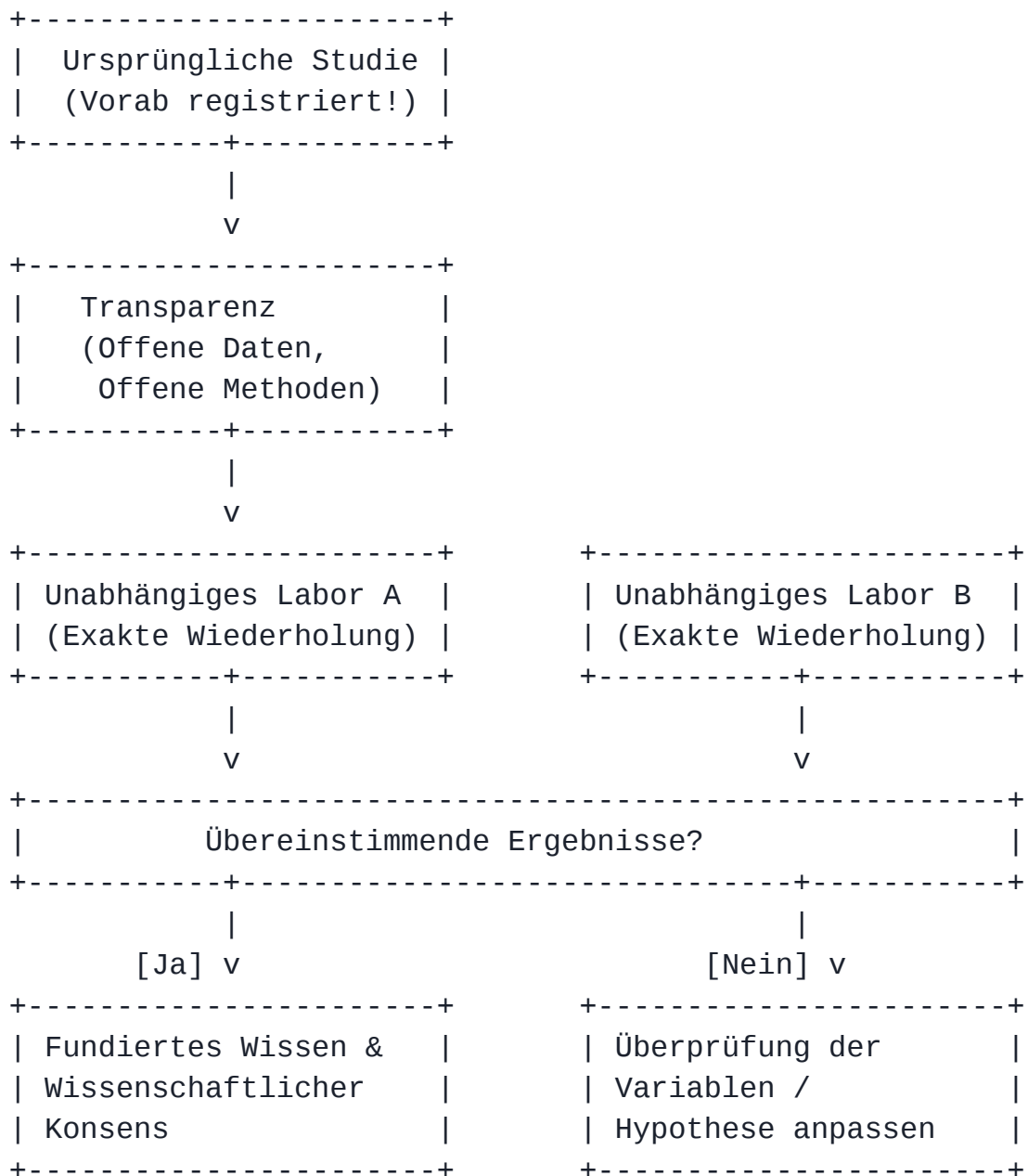
Die wichtigste Maßnahme ist die **Präregistrierung** (Preregistration).

Die Metapher des versiegelten Umschlags: Stellen Sie sich vor, ein Forscher schreibt sein genaues Vorgehen, seine Hypothese und die geplanten statistischen Tests auf ein Blatt Papier, steckt es in einen Umschlag und versiegelt ihn. Erst

danach beginnt er mit dem Experiment. Wenn die Daten später ausgewertet sind, wird der Umschlag öffentlich geöffnet. So wird verhindert, dass die Methodik so lange angepasst wird, bis das Wunschergebnis erreicht ist.

In der modernen Wissenschaft wird dieser "Umschlag" digital auf unabhängigen Plattformen hinterlegt, bevor die erste Speichelprobe im Labor entnommen oder der erste Fragebogen ausgeteilt wird. Zudem fordert Open Science, dass Forschungsdaten (anonymisiert), Codes und Protokolle für die Welt frei zugänglich gemacht werden.

DAS BAU-GERÜST DER ROBUSTEN WISSENSCHAFT



4. Bedeutung für Natur, Körper und Alltag: Schutz durch Methodik

Reproduzierbarkeit ist kein abstraktes, rein akademisches Problem – sie entscheidet buchstäblich über Gesundheit, Lebensqualität und technischen Fortschritt.

In der **Medizin und Pharmakologie** ist die Replizierbarkeit von Studien die wichtigste Firewall, die uns vor ineffektiven oder gefährlichen Therapien schützt. Besonders für schwer erkrankte Menschen, die verzweifelt nach Hilfe suchen, ist diese Absicherung überlebenswichtig. Wenn ein einzelnes Labor behauptet, ein neuer Wirkstoff stoppe das Wachstum von Tumorzellen in der Petrischale, ist das eine wunderbare erste Nachricht. Würden wir diese Substanz jedoch sofort als Medikament freigeben, könnten Patienten falsche Hoffnungen schöpfen, wirksame Alternativen abrechnen oder schweren Nebenwirkungen ausgesetzt sein, falls es sich nur um ein falsch-positives Ergebnis handelte.

Deshalb ist das Prinzip der Reproduzierbarkeit in der Medizin institutionalisiert, und zwar in Form der klinischen Studienphasen:

KLINISCHE PRÜFUNG: INSTITUTIONALISIERTE REPRODUZIERBARKEIT

Phase I	(Labor/Wenige Probanden)	--> Fokus: Sicherheit. Erste Orientierung
Phase II	(Hunderte Probanden)	--> Fokus: Dosierung & erste Wirksamkeit
Phase III	(Tausende Probanden)	--> Fokus: Reproduzierbarkeit im großen Maßstab (Findet an unabhängigen Zentren weltweit statt)

Erst wenn unabhängige Kliniken in doppelblinden, randomisierten Phase-III-Studien denselben therapeutischen Effekt reproduzieren können, wird aus einer Laborbehauptung ein zugelassenes Medikament.

Auch in der **Technik und Physik** basiert unser Alltag auf dieser harten Nachvollziehbarkeit. Dass Halbleiter in Computern zuverlässig arbeiten, dass Satelliten die Relativitätstheorie zur Navigation nutzen und dass Flugzeuge sicher fliegen, beruht darauf, dass die zugrunde liegenden physikalischen Prinzipien millionenfach reproduziert wurden.

5. Häufige Missverständnisse und Halbwissen

Das öffentliche Verständnis von Wissenschaft ist mitunter von absoluten Erwartungen geprägt, was zu Missverständnissen über den Forscheralltag führt.

Ein klassisches **Missverständnis** ist der Glaube: *"Wenn eine Studie nicht repliziert werden kann, haben die Forscher gelogen oder grob fehlerhaft gearbeitet."* Tatsächlich ist wissenschaftlicher Betrug (wie das bewusste Fälschen von Daten) ein ernstes, aber sehr seltenes Problem. Die meisten gescheiterten Replikationen sind schlichtweg die normale Konsequenz eines hochkomplexen Universums. Zellkulturen könnten unbemerkt mutiert sein, oder die Probanden einer psychologischen Studie in den USA im Jahr 2005 unterschieden sich kulturell von Probanden in Deutschland im Jahr 2020. Eine gescheiterte Replikation ist oft der Startschuss, um versteckte, bisher unbekannte Einflussfaktoren (sogenannte Moderatoren) zu entdecken. Ein Medikament könnte beispielsweise in einer US-Studie wirksam sein, in einer japanischen Replikationsstudie aber scheitern, weil sich die genetische Verteilung bestimmter Stoffwechselenzyme unterscheidet. Dies ist keine "schlechte Wissenschaft", sondern ein wertvoller Erkenntnisgewinn für die personalisierte Medizin.

Ein weiteres häufiges **Halbwissen** lautet: *"Wissenschaft behauptet immer dogmatisch, die absolute Wahrheit zu kennen."* Wahr ist das Gegenteil. Die wissenschaftliche Methode ist radikal skeptisch, besonders sich selbst gegenüber. Kein einzelnes Paper, egal wie berühmt das Journal ist, „beweist“ etwas endgültig. Medizinisches Wissen baut auf einem Konsens auf, der sich aus der Summe dutzender, fachlich reproduzierter Studien in Meta-Analysen ergibt.

Zuletzt hört man oft: *"Die Replikationskrise zeigt, dass die Wissenschaft kaputt ist."* Dass Wissenschaftler heute so schonungslos über die Replikationskrise sprechen und schwache Studien aussortieren, zeigt nicht, dass das System kaputt ist – es zeigt, dass es **funktioniert**. Welches andere System wendet so viel Energie und Geld auf, um die eigenen Irrtümer der Vergangenheit aufzudecken und öffentlich zu korrigieren?

6. Fehlinterpretationen in der Gesundheitskommunikation

In der öffentlichen Debatte um Gesundheitsthemen und komplementäre Heilmethoden kommt es gelegentlich zu Missverständnissen bezüglich der wissenschaftlichen Methodik. Besonders das Konzept der Reproduzierbarkeit wird hierbei manchmal fehlinterpretiert – oft aus einem zutiefst nachvollziehbaren Bedürfnis von Patienten nach schneller Linderung, Gewissheit und neuen

Hoffnungsschimmern.

Diese Fehlinterpretationen treten meist in zwei Argumentationsmustern auf:

1. Fehlinterpretation der methodischen Selbstkritik: Mitunter wird argumentiert, dass eine hohe Nicht-Reproduzierbarkeitsquote in bestimmten experimentellen Bereichen bedeute, dass auch etablierte, evidenzbasierte Erkenntnisse der Medizin generell unzuverlässig seien. Daraus wird in manchen Debatten abgeleitet, dass nicht-evidenzbasierte Methoden eine vollkommen gleichwertige Alternative seien. Hierbei wird jedoch der Kern der naturwissenschaftlichen Fehlerkultur übersehen: Nur weil die Wissenschaft extrem strenge Qualitätskontrollen anwendet und systematisch Fehler in den eigenen Reihen korrigiert, führt dies im Umkehrschluss nicht dazu, dass eine andere Methode ohne Wirksamkeitsnachweis verlässlich wird. Die evidenzbasierte Medizin sortiert kontinuierlich aus, was nicht funktioniert, um Patienten immer besser schützen zu können.

2. Rosinenpickerei (Cherry-Picking) von Einzelstudien: Ein weiteres Phänomen ist das selektive Herausgreifen isolierter Forschungsergebnisse. In manchen Publikationen oder bei Werbeversprechen wird eine kleine, methodisch eingeschränkte Einzelstudie präsentiert, in der zufällig (aufgrund des zuvor erklärten Alpha-Fehlers) ein positiver Effekt eines bestimmten Präparats beobachtet wurde. Für Betroffene, die nach jedem Strohalm greifen, klingt dies oft wie ein zweifelsfreier medizinischer Durchbruch. Wenn jedoch in den darauffolgenden Jahren zahlreiche große, methodisch einwandfreie Replikationsstudien durchgeführt wurden, die diesen Verdacht nicht bestätigen konnten, wird dieses Gesamtbild in werblichen Zusammenhängen leider oft unerwähnt gelassen. Eine verantwortungsvolle und empathische Patientenkommunikation stützt sich daher immer auf die Gesamtheit der reproduzierten Befunde (Evidenzsynthese), um sicherzustellen, dass Betroffene bestmöglich informiert sind und keine wertvolle Zeit verlieren.

7. Faszination Wissenschaft (Der Wow-Effekt)

Die wirkliche Faszination der Reproduzierbarkeit liegt in dem atemberaubenden Vertrauen, das sie uns in unsere Erkenntnis über den Kosmos schenkt.

Stellen Sie sich das vor: Ein Forscherteam in Tokio, ein anderes in Boston und ein drittes in Kapstadt entschlüsseln die atomare Struktur eines Virus oder messen die Schwingungen von Schwarzen Löchern. Diese Menschen sprechen verschiedene Sprachen, haben unterschiedliche

persönliche Überzeugungen und nutzen völlig unterschiedliche Maschinen. Und dennoch, am Ende des Tages, wenn sie das Rauschen der Daten bereinigen, blicken alle drei Gruppen exakt auf dieselben Baupläne der Natur.

Dieses Phänomen offenbart eine tiefgreifende, wunderschöne Wahrheit über unser Universum: Die Naturgesetze sind keine sozialen Konstrukte oder bloße Meinungen. Sie sind eine objektive Realität. Das groß angelegte "Reproducibility Project", bei dem Forscher weltweit versuchten, Hunderte alte Befunde systematisch neu zu testen, war daher kein Akt der Zerstörung. Es war einer der heroischsten Akte intellektueller Bescheidenheit in der Menschheitsgeschichte.

Die Wissenschaft ist das einzige Wissenssystem, das institutionell darauf ausgelegt ist, sich selbst permanent herauszufordern. Gerade weil Forscher wissen, wie leicht der menschliche Verstand Muster sieht, wo keine sind, haben sie mit der Reproduzierbarkeit ein Instrument erschaffen, das langsam, manchmal fehleranfällig und mühsam, aber am Ende unaufhaltsam und verlässlich die tatsächlichen Gegebenheiten unserer Welt ans Licht bringt.