

Stellungnahme



Replizierbarkeit von Forschungsergebnissen

Eine Stellungnahme der Deutschen Forschungsgemeinschaft

Stand: April 2017

Ansprechpartnerin:

Dr. Katja Hartig

Deutsche Forschungsgemeinschaft (DFG)

Katja.Hartig@dfg.de

Deutsche Forschungsgemeinschaft

Kennedyallee 40 · 53175 Bonn · Postanschrift: 53170 Bonn

Telefon: + 49 228 885-1 · Telefax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de



Replikation ist ein sehr wichtiges Verfahren zur Prüfung experimentalwissenschaftlich und quantitativ begründeter empirischer Wissensansprüche in der Medizin, in den Natur-, Lebens-, Ingenieur- sowie den Sozial- und Verhaltenswissenschaften und auch den Geisteswissenschaften. Seit im Jahr 2014 in der Zeitschrift *Lancet* eine Serie¹ zur Replizierbarkeit von Forschungsergebnissen (in der Biomedizin) erschien, entwickelte sich unter dem Schlagwort „Replikationskrise“ eine rege, auch öffentliche Debatte. Diese berührt im Kern die Frage der Qualität von Forschung und betrifft die gesamte Wissenschaft.

Replikation ist allerdings nicht das einzige Prüfverfahren. Es gibt andere, zu denen etwa die theoretisch-begriffliche Diskussion und Kritik, Modellbildung, mathematische Modellierung, Simulation und andere gehören. Insofern ist bei der Diskussion von Fragen der Replikation und Replizierbarkeit zu bedenken:

- Wissenschaftliche Ergebnisse können replizierbar sein, sie müssen es nicht. **Replizierbarkeit ist kein generelles Kriterium wissenschaftlicher Erkenntnis.** Die Erwartung, alle wissenschaftlichen Ergebnisse müssten replizierbar sein, ist schon deshalb nicht erfüllbar, weil zahlreiche Forschungsgebiete einmalige Ereignisse wie Klimawandel, Sternener Explosionen, Vulkanausbrüche oder vergangene Vorkommnisse untersuchen. Andere Forschungsgebiete konzentrieren sich auf die Beobachtung und Analyse kontingenter Phänomene (z. B. in den Erdsystemwissenschaften oder in der Astrophysik) oder untersuchen Phänomene, die aus anderen (etwa ethischen, finanziellen oder technischen) Gründen nicht wiederholt beobachtet werden können. Darüber hinaus gibt es Formen der Forschung, die einen solchen Grad der Komplexität der experimentellen Methodik erreicht haben, dass von daher ihre replizierende Wiederholung schwierig sein kann.
- **Die Feststellung der Replizierbarkeit oder Nicht-Replizierbarkeit eines wissenschaftlichen Ergebnisses ist ihrerseits ein wissenschaftliches Ergebnis.** Als solches ist es nicht endgültig, sondern steht, wie alle wissenschaftliche Erkenntnis, unter dem Vorbehalt methodischer Skepsis und weiterer Überprüfung.
- Nicht-Replizierbarkeit eines empirisch begründeten wissenschaftlichen Wissensanspruchs kann dessen Falsifizierung bedeuten, sie muss es nicht. **Nicht-Replizierbarkeit ist kein genereller Falsifikationsbeweis.** Unter bestimmten Umständen kann im Gegenteil Nicht-Replizierbarkeit auch als Derzeit-Noch-Nicht-Replizierbarkeit zu verstehen sein oder als Signal für die Existenz eines neuen, bisher

¹ Chalmers et al. *Lancet*, Vol. 383 (2014); Ioannidis et al. *Lancet*, Vol. 383 (2014); Salman et al. *Lancet*, Vol. 383 (2014); Chan et al. *Lancet*, Vol. 383 (2014); Glasziou et al. *Lancet*, Vol. 383 (2014)

noch unbekanntem wissenschaftlichen Zusammenhangs gedeutet werden und insofern ein Hinweis auf im Sinne des wissenschaftlichen Erkenntnisfortschritts positiv riskante Forschung sein.

- Nicht-Replizierbarkeit eines empirisch begründeten wissenschaftlichen Wissensanspruchs kann, muss aber nicht damit zusammenhängen, dass es sich um ein Resultat schlechter wissenschaftlicher Praxis oder von wissenschaftlichem Fehlverhalten handelt. **Nicht-Replizierbarkeit ist kein genereller Hinweis auf schlechte Wissenschaft.** Für die Prüfung der Frage, ob bei der Vorbereitung, Durchführung, Auswertung, Darstellung oder Publikation eines Experiments in diesem Sinne schlechte Wissenschaft vorliegt oder nicht, ist zu verweisen auf die „Empfehlungen zur Sicherung guter wissenschaftlicher Praxis“ der DFG und den in diesem Zusammenhang entwickelten Verfahrenleitfaden.²
- Der Sachverhalt, dass es tragfähige, obwohl nicht-replizierbare wissenschaftliche Erkenntnis gibt, darf in keinem Fall als Ausflucht oder Entschuldigung für Nicht-Replizierbarkeit dort missbraucht werden, wo die Replizierbarkeit eines wissenschaftlichen Wissensanspruchs methodisch erwartet werden muss.

Vor dem Horizont dieser Grundsätze sind die seit der *Lancet*-Serie 2014 erschienenen Fachbeiträge und die daran anschließende öffentliche Debatte differenziert zu beurteilen. Die Deutsche Forschungsgemeinschaft nimmt diese Diskussion zum Anlass für die nachfolgende Positionierung.

- Die Debatte in Wissenschaftspolitik und Öffentlichkeit identifiziert vielfach undifferenziert Replizierbarkeit mit guter, Nicht-Replizierbarkeit mit schlechter Wissenschaft; insofern folgt sie den oben angedeuteten Fehlschlüssen.
- Die durch die *Lancet*-Serie angestoßene Auseinandersetzung, Metastudien³ zur Replizierbarkeit und diesbezügliche Initiativen⁴ sind ihrerseits Zeugnis funktionierender Mechanismen wissenschaftlicher Selbstkontrolle.
- Die berichteten Fälle von Nicht-Replizierbarkeit in ihrer Gesamtheit sind auch dann besorgniserregend, wenn man davon ausgeht, dass sie auf eine Mehrzahl von Ursachen zurückzuführen sind und keineswegs alle schlechte wissenschaftliche Praxis oder wissenschaftliches Fehlverhalten bezeugen. Insofern indiziert die Diskussion über

² Empfehlungen zur Sicherung guter wissenschaftlicher Praxis und Verfahrenleitfaden zur guten wissenschaftlichen Praxis: www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/index.html

³ Open Science Collaboration (2015); Gilbert, King, Pettigrew & Wilson (2016)

⁴ NIH Rigor and Reproducibility Initiative; Center for Open Science; ARRIVE; Force 11

die sogenannte Replikationskrise ein **Qualitätsproblem von Forschung**, dessen Umfang freilich nicht genau bestimmbar ist. Forscherinnen und Forscher, Forschungseinrichtungen und Wissenschaftsorganisationen müssen dieses Qualitätsproblem sehr ernst nehmen. Es gefährdet die Leistungsfähigkeit der Wissenschaften ebenso wie das gesellschaftliche Vertrauen in sie.

- Insofern Elemente dieses Qualitätsproblems als schlechte wissenschaftliche Praxis oder **wissenschaftliches Fehlverhalten** individuell zurechenbar sind, greifen innerhalb der Wissenschaft die Kontrollmechanismen des Ombudsmans für die deutsche Wissenschaft, der lokalen Ombudsgremien und der in den wissenschaftlichen Einrichtungen etablierten Verfahren zur Aufklärung und Ahndung von Fällen wissenschaftlichen Fehlverhaltens sowie, bezüglich DFG-geförderter Forschung, insbesondere die des Ausschusses zur Untersuchung von Vorwürfen wissenschaftlichen Fehlverhaltens.
- Neben individuellem Fehlverhalten gibt es für das Qualitätsproblem von Forschung allerdings auch **strukturelle Gründe**. Das mittlerweile in der Wissenschaft erreichte Gewicht von quantitativ parametrisierenden Steuerungs-, Bewertungs- und Gratifikationssystemen wirkt sich auf die Forschung als gesteigener (und weiter steigender) Wettbewerbs- und Beschleunigungsdruck aus. Dieser manifestiert sich in Entscheidungen (und zugrunde liegenden Entscheidungskriterien) über Karriereschritte, finanzielle Förderung, Publikationsorte oder institutionelle Strukturentwicklungen. Die notwendige skrupulöse Sorgfalt bei der Vorbereitung, Durchführung, Auswertung, Darstellung und Publikation experimenteller und empirisch-quantitativer Forschung braucht Zeit, Gelegenheit, Mittel und Personal. Sie muss eher gegen diesen Wettbewerbs- und Beschleunigungsdruck durchgesetzt werden, als dass sie von ihm befördert würde. Gleiches gilt für die Würdigung und Publikation von sogenannten neutralen oder negativen oder redundanten Ergebnissen.

Angesichts von strukturellen Rahmenbedingungen, die allzu leicht missverstanden werden können als Einladung zu einer Forschungspraxis des *quick and dirty*, sieht sich auch die **DFG in der Verantwortung**. In der Wahrnehmung dieser Verantwortung verlangen verschiedene Aspekte ihres Handelns als Organisation der Forschungsförderung und der wissenschaftlichen Selbstverwaltung Aufmerksamkeit. Die DFG

- rückt bei der Projektförderung den spezifisch zu erwartenden Erkenntnisgewinn eines Forschungsvorhabens in den Mittelpunkt der Beurteilung;
- wird bei der Weiterentwicklung ihres Förderportfolios und der von ihr zu verantwortenden Begutachtungs-, Bewertungs- und Entscheidungsprozesse darauf achten, dass weniger die Quantität oder der Ort als vielmehr die Qualität von Publikationen als Kriterium im Zentrum der wissenschaftlichen Urteilsbildung steht;
- wird dabei auch berücksichtigen, dass Replikation als Verfahren der Prüfung von experimentellen und empirisch-quantitativen Forschungsergebnissen systematisch gestärkt werden muss;
- begleitet und unterstützt daher Prozesse der fachspezifischen Klärung jener Fragen, die sich in Bezug auf die Replizierbarkeit von Forschungsergebnissen ergeben; hierzu gehören auch die Entwicklung fachspezifischer Kriterien für die Förderung von Replikationsstudien⁵ sowie diese Förderung selbst;
- wird dabei weiterhin besonderes Augenmerk auf Fragen des Forschungsdatenmanagements sowie auf aktuelle Herausforderungen, die aus der Digitalisierung erwachsen, richten;
- fördert in diesem Zusammenhang die Entwicklung von Infrastrukturen und methodischen Werkzeugen sowie deren Nutzung;
- führt ihre vielfältigen Anstrengungen zur Förderung guter wissenschaftlicher Praxis, die im deutschen Wissenschaftssystem Maßstäbe setzen, uneingeschränkt fort;
- appelliert an die wissenschaftlichen Verlage, die wissenschaftlichen Einrichtungen und Ethikkommissionen sowie den Gesetzgeber und die Wissenschaftspolitik, alles zu tun, um den strukturellen Gründen für Replikationserschwerernisse entgegenzuwirken.

⁵ Empfehlungen des Fachkollegiums Psychologie: Psychologische Rundschau 67/3: Seiten 163ff. (2016)