

Peters, Jörg

Article — Published Version

Replikationskrise (?)

KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie

Provided in Cooperation with:

Springer Nature

Suggested Citation: Peters, Jörg (2021) : Replikationskrise (?), KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, ISSN 1861-891X, Springer Fachmedien Wiesbaden, Wiesbaden, Vol. 73, Iss. 2, pp. 311-314,
<https://doi.org/10.1007/s11577-021-00783-7>

This Version is available at:

<https://hdl.handle.net/10419/287040>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Replikationskrise (?)

Jörg Peters

Angenommen: 1. Juni 2021 / Online publiziert: 6. Juli 2021
© Der/die Autor(en) 2021

Ritchie, Stuart: Science Fictions. How Fraud, Bias, and Hype Undermine the Search for Truth. New York: Metropolitan Books 2020. 368 Seiten. ISBN: 978-1-250-22269-5. Preis: \$ 29,99.

Die gesellschaftliche Bedeutung von Wissenschaft steigt zurecht, weil Wissenschaft objektiv ist und die Fakten offenlegt. Politische Entscheidungen sollten deshalb stets auf wissenschaftlichen Erkenntnissen basieren.

Wer sich an diesem Diktum stört, aber auch, wer es für zutreffend hält, ist bei Stuart Ritchies Buch „Science Fictions“ goldrichtig. Natürlich sollten in vielen Bereichen wissenschaftliche Erkenntnisse politische Entscheidungen informieren, teilweise sogar leiten. Das stellt auch Ritchie nicht in Frage. Doch mahnt seine schonungslose Beschreibung des wissenschaftlichen Systems und seiner Fehleranfälligkeit zu mehr Vorsicht bei szientistischen Träumereien. Viele, der im Buch beschriebenen grundlegenden Prozesse gelten disziplinübergreifend, doch richtet es sich vor allem an quantitativ arbeitende Professionen. Das Buch fordert die beanspruchte Deutungshoheit gerade jener Disziplinen heraus, die vom Erkenntnisideal der falsifizierbaren Hypothesentestung geleitet sind.

Ritchie stellt gleich zu Beginn klar, dass er antritt, „um die Wissenschaft zu lobpreisen, nicht um sie zu beerdigen“. Konsolidierte wissenschaftliche Ergebnisse sind zentral für gesellschaftliches Vorankommen, daran lässt er keinen Zweifel. In Ritchies Kritik geht es eigentlich um eben dieses Wort „konsolidiert“. Denn

Jörg Peters (✉)

Forschungsgruppe „Klimawandel in Entwicklungsländern“, Leibniz Institut für Wirtschaftsforschung
Hohenzollernstraße 1-3, 45128 Essen, Deutschland
E-Mail: joerg.peters@rwi-essen.de

Lehrstuhl für Development Economics, Universität Passau
Innstraße 29, 94031 Passau, Deutschland

ein empirisches Ergebnis gilt im erwähnten Erkenntnisideal nur dann als konsolidiert und nicht zufällig ausgelöst, wenn es *replizierbar* ist. Es darf also nicht nur einmalig zu beobachten sein, sondern muss in mehreren Studien und unter unterschiedlichen Rahmenbedingungen nachgewiesen werden können. Werden politische Entscheidungen auf nicht replizierbaren Forschungsergebnissen basiert, laufen die politischen Maßnahmen in die Irre. Ritchie setzt bei der Beobachtung an, dass die Replizierbarkeit einflussreicher wissenschaftlicher Ergebnisse in vielen quantitativ arbeitenden Disziplinen erschreckend gering ist – die sogenannte Replikationskrise.

Das Buch ist eine unterhaltsame, wortgewandte, teilweise ins Sarkastische abdriftende, doch stets analytische Beschreibung des Wissenschaftsprozesses. Die Vergabe von Forschungsmitteln und wissenschaftlichen Posten, der Publikationsprozess mit samt Peer-Review, aber auch die Wissenschaftskommunikation werden schonungslos beschrieben als das, was sie sind: ein von Menschen gemachtes und deshalb fehleranfälliges System. Die Wissenschaft ist ein soziales Konstrukt – und so spielen die Menschen und ihre Eitelkeiten, ihr Überlebens- und Aufstiegszweck, ihre Hybris und ihre Beziehungen zueinander eine wichtige Rolle. Dies mag in Disziplinen mit Nähe zur Kritischen Theorie wie der Soziologie oder den Politikwissenschaften trivial sein – in vielen anderen Professionen jedoch hinterfragt es das dominierende Narrativ von „Objektivität“ und „Evidenzbasierung“.

Ritchie kritisiert, dass quantitative empirische Ergebnisse medial oft als unumstößliche Fakten interpretiert werden, betont aber gleichzeitig, dass die Verantwortung dafür auch bei der Wissenschaft liegt, weil Forschende oder institutionelle Pressemitteilungen der Forschungseinrichtungen nüchterne Studienergebnisse als spektakuläre Fakten präsentieren. Tatsächlich entspringen die Studienergebnisse aber dem oben geschilderten sozialen Konstrukt, das sie viel fehleranfälliger macht als meist dargestellt. Dafür liefert Ritchie viel anschauliches Material. Er rekapituliert einschlägige Wissenschaftsskandale, doch zeigt er auch, dass Fehler systematisch passieren. Auch in hypotheseentestenden empirischen Ansätzen geschieht dies, durch Betrug und Nachlässigkeit, wahrscheinlich aber noch mehr durch systemische Verzerrungen, bei denen nicht unbedingt einzelne Menschen böswillig agieren. Menschen, so auch Forschende, bevorzugen Spektakuläres gegenüber dem Unspektakulären. Über das sogenannte *p-hacking* und den *Publication Bias* führt diese Suche nach dem Spektakulären zu systematischen Verzerrungen. *Publication Bias* und *p-hacking* führen, kurz gesagt, dazu, dass die statistischen Methoden, auf die die quantitative Forschung so stolz ist, ad absurdum geführt werden. Es ist perfide, dass gerade diese Phänomene im Big-Data-Zeitalter immer wahrscheinlicher werden, weil die unbegrenzte Datenverfügbarkeit den Spielraum für *p-hacking* und den *Publication Bias* erhöht.

Wie funktioniert der *Publication Bias*? Die Gestaltung statistischer Hypothesentests sichert ein Ergebnis nur mit einer gewissen Wahrscheinlichkeit, sodass es stets eine Restwahrscheinlichkeit – meist per Konvention 5 % – gibt, dass man ein Ergebnis fälschlicherweise für statistisch gesichert erachtet. Ein solches Ergebnis wäre nicht replizierbar und damit wertlos. Dies bedeutet, dass selbst zu einer unsinnigen Theorie bei 100 Versuchen, beispielsweise in 100 Datensätzen, fünf davon diese Theorie dennoch statistisch signifikant bestätigen. Würden nun alle 100 Versuche publiziert, wäre es unproblematisch. Andere Wissenschaftlerinnen könnten dann

richtigerweise erkennen, dass die fünf erfolgreichen Studien dem Zufall geschuldet sind. Allerdings werden nicht alle Ergebnisse publiziert. Peer-Reviewer und Herausgeber der Fachzeitschriften befinden die fünf signifikanten Studien für interessanter und werden tendenziell eher diese publizieren und die nichtsignifikanten Studien ablehnen. Die so entstehende veröffentlichte Literatur zeigt dann ein falsches Bild der Wirklichkeit.

Verwandt damit ist das *p*-hacking: Unspektakuläre Ergebnisse können spektakulärer gemacht werden, indem das statistische Signifikanzniveau verbessert wird. Jede empirische Untersuchung umfasst dutzende, wenn nicht hunderte Mikroentscheidungen. Diese beginnen bei sehr fundamentalen Entscheidungen darüber, welche Forschungsfrage überhaupt gestellt wird, reichen aber weiter, hinein in auch für die Forschenden selbst kaum spürbare Entscheidungen über unterschiedliche Messungen oder die Rohdatenbereinigung. Entscheidet man sich bei diesen Mikroentscheidungen – vielleicht sogar unbewusst – für das jeweils spektakulärer aussehende Ergebnis, werde ich irgendwann wegen der 5%-Irrtumswahrscheinlichkeit einen signifikanten Zusammenhang finden. Ein solches Vorgehen ist legitim, wenn alle Versuche dokumentiert und publiziert werden. Das werden sie aber meist nicht, sodass, wie beim Publication Bias, wieder ein falsches Bild der Wirklichkeit entsteht.

Ritchies Buch ist besonders überzeugend, weil es die Wechselwirkung zwischen der Fehleranfälligkeit statistischer Methoden und dem sozialen Konstrukt, das die Wissenschaft darstellt, herausarbeitet. Denn in der Wissenschaft geht es mitnichten nur um Erkenntnisgewinn, sondern um Macht und Einfluss, um Drittmittel, um Posten und, gerade zu Beginn der Karriere, um entfristete Stellen und Professuren. In vielen Disziplinen werden diese zentralen Karriereentscheidungen zunehmend von Publikationen, in ihrer Anzahl und der rankingbasierten Journalqualität, abhängig gemacht. Die Anreize, spektakulärere Ergebnisse zu produzieren, um gut zu publizieren, sind also klar. Zugleich ist es naiv zu erwarten, Peer-Review könnte *p*-hacking und die besagten Mikroentscheidungen nachverfolgen und so die Qualität sichern. Im Gegenteil, durch den Publication Bias und das Faible der Gutachterinnen und Gutachter für spektakuläre Ergebnisse ist das Peer-Review sogar Teil des Problems.

Dies alles ist natürlich keineswegs neu. Dennoch ist Ritchies Buch in seiner Reichhaltigkeit und ganzheitlichen Betrachtung ein wichtiger Beitrag. Wie dramatisch die Lage ist oder ob Ritchie es allzu pessimistisch sieht, wird im Auge der Betrachtenden variieren. Manche seiner Argumente sind notgedrungen eher anekdotisch, doch viele seiner Punkte untermauert er auch eindrücklich mit systematischer Evidenz. Das Buch ist gespickt mit Referenzen zu einschlägigen Studien aus unterschiedlichen Disziplinen. Dass die „perversen Anreize“ im sozialen Konstrukt Wissenschaft den Mertonschen Regeln zuwiderlaufen, macht Ritchie sehr deutlich. Gerade auch für junge Wissenschaftlerinnen und Wissenschaftler in den weniger selbstkritischen Professionen ist das Buch deshalb eine gute Vorbereitung auf das, was sie erwartet.

Ritchie ist kein Wissenschaftsfundamentalskeptiker oder Relativist. Er fordert die quantitativen Methoden heraus, nicht um sie zu kompromittieren, sondern um sie zu reformieren. Deshalb trägt das Buch zum Abschluss eine Reihe von Lösungsansätzen vor. Dabei geht es zunächst um mehr Forschungstransparenz, also Prä-Spezifizierung

von Hypothesen, aber auch die Verfügbarmachung von Daten, die einer publizierten Studie zugrunde liegen. Ritchie macht aber ebenso deutlich, dass dies nur ein Schritt in Richtung einer Verbesserung ist. Diese Verbesserung tritt effektiv nur ein, wenn sich die Kultur nachhaltig ändert. Dazu gehört die Anwendung der Mertonschen Regel des „Organisierten Skeptizismus“, also Replikationen und Reanalysen. Die Anreize jedoch, Replikationen durchzuführen, fehlen gegenwärtig in den meisten Disziplinen. Replikationen werden größtenteils karrieristisch nicht belohnt, weil sie schwer in prominenten Fachzeitschriften zu veröffentlichen sind. Auch täte laut Ritchie mehr bayesianisches Denken der Glaubwürdigkeit von Wissenschaft gut, weil es die Standards für spektakuläre Ergebnisse dadurch erhöht, dass frühere Forschungsergebnisse stärker einbezogen werden müssen. Der abschließende und wahrscheinlich effektivste Vorschlag ist, Karriereentscheidungen über Entfristung, Professuren und Drittmittel nicht nur auf Publikationen zu basieren, sondern auch auf deren Replizierbarkeit. Eine fesselnde Idee, bis auf Weiteres aber wohl auch utopisch.

Zu solchen Veränderungen in der Wissenschaftskultur wird es jedoch nur dann kommen, wenn die Probleme benannt und auch der nichtwissenschaftlichen Öffentlichkeit bekannt werden. In diesem Sinne ist dieses Buch ein wichtiger Schritt in Richtung einer glaubwürdigeren Wissenschaft.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Jörg Peters Prof. Dr., Kompetenzbereichsleiter Klima & Entwicklung am RWI Leibniz Institut für Wirtschaftsforschung. Forschungsschwerpunkte: Umweltökonomik, Entwicklungsökonomik, Metascience. Ausgewählte Publikationen: Social norms and energy conservation beyond the US. *Journal of Environmental Economics and Management*, 103, 2020 (mit M.A. Andor, A. Gerster und C. M. Schmidt); Successful coal phase-out requires new models of development. *Nature Energy* 4 (11), 2019 (mit M. Kalkuhl, J. C. Steckel, L. Montrone, M. Jakob und O. Edenhofer); Generalization in the tropics – development policy, randomized controlled triads, and external validity. *World Bank Research Observer* 33 (1), 2018 (mit J. Langbein und G. Roberts).