

### III.1.6

## Peer Review

*Martin Reinhart*

**Abstract |** Peer Review ist die zentrale Institution zur Sicherung wissenschaftlicher Qualität. Die Begutachtung von Ergebnissen wissenschaftlicher Arbeit durch Fachkolleg\*innen wird prominent bei der Publikation wissenschaftlicher Zeitschriftenartikel und der Förderung von Forschungsprojekten praktiziert. Im Begutachtungsprozess werden wissenschaftliche Arbeitsergebnisse ein erstes Mal außerhalb ihres Entstehungskontextes kritisch gelesen, diskutiert und allenfalls überarbeitet. Die daraus verfahrensförmig hervorgehenden Entscheidungen über Publikation oder Förderung verteilen knappe Ressourcen (Publikationsfläche, Forschungsmittel) und steuern damit, welche Forschung und welche Forschenden Chancen auf Aufmerksamkeit und Anerkennung erhalten. Als Selbststeuerungsmechanismus verteilt das Peer Review aber nicht nur Reputation und sichert dadurch die Qualität nach innen. Es sichert gleichzeitig die Autonomie der Wissenschaft nach außen, bspw. gegenüber der Politik, da es durch wissenschaftliche Expertise und Verfahrensfähigkeit als legitim erachtete Entscheidungen hervorbringt.

**Stichworte |** Peer Review, Bewerten, Selbststeuerung

## Einleitung

Was ist schlimmer, die Publikation von falschem Wissen oder gar keine Publikation? (Kennefick 2005). Erstaunlicherweise sind sich Forschende wenig einig in dieser Frage. Woher soll denn neues und v. a. originelles Wissen kommen, wenn das Risiko nicht eingegangen wird, Gewagtes aber möglicherweise Falsches zu publizieren, fragen sich einige. Andere halten die Vorstellung für schwer erträglich, dass die publizierte wissenschaftliche Literatur voller Fehler und Falschheiten sein könnte. Albert Einstein war 1936 noch überzeugt, dass es sowohl unnötig war als auch einen Vertrauensbruch darstellte, als eine renommierte wissenschaftliche Zeitschrift sein eingereichtes Manuskript an einen anonymen Kollegen zur Begutachtung schickte. Obwohl er das Manuskript später entlang der gut-achterlichen Vorschläge stark überarbeitet bei einer anderen Zeitschrift publizierte, blieb er empört. Er reichte nie wieder ein Manuskript bei der Zeitschrift ein, die sein Manuskript einer kollegialen Begutachtung, einem Peer Review, unterzogen hatte. Einsteins Reaktion und Haltung wirken aus heutiger Sicht nicht mehr angemessen, was zeigt, dass sich die Vorstellungen gewandelt haben, wie sich die Qualität der wissenschaftlichen Wissensproduktion sichern lässt.

Beim Peer Review geht es darum, die Produktion von wissenschaftlichem Wissen auf seine Qualität zu prüfen. Dabei soll die Möglichkeit zum Austausch wissenschaftlicher Ideen aber nicht unnötig eingeschränkt werden. Peer Review ist der primäre Mechanismus, über den das geschieht, und zwar indem v. a. Manuskripte zur Publikation und Projektvorschläge zur finanziellen Förderung von Fachkolleg\*innen begutachtet werden. Als Ergebnis davon kommen Entscheidungen darüber zustande,

was von Zeitschriften publiziert und von Forschungsförderungsorganisationen finanziert wird. Diese Entscheidungen sind folgenreich, weil die Publikation in einer renommierten Zeitschrift oder die Finanzierung eines Forschungsprojektes in Konkurrenz mit Fachkolleg\*innen erfolgt, die dieselben knappen Güter (Publikationsfläche, Aufmerksamkeit, finanzielle Mittel) auch erhalten möchten (Merton 1957, Latour/Woolgar 1986: 188ff.; → III.1.4 Fachkommunikation). Damit ist das Peer Review nicht nur ein Mechanismus zur Qualitätssicherung, sondern auch zur Verteilung knapper Güter und damit zur Steuerung von Wissenschaft (Reinhart 2012). Als Folge ist das Peer Review umstritten und immer wieder der Kritik ausgesetzt, so dass laufend Veränderungsvorschläge diskutiert werden, aber ohne, dass eine Mehrheit der Forschenden die Abschaffung des Peer Review fordern würde (Hirschauer 2004, Reinhart/Schendzielorz 2024).

Der folgende Beitrag geht deshalb von der Einsicht aus, dass das Peer Review den zentralen Qualitäts- sicherungs- und Steuerungsmechanismus der Wissenschaft darstellt, dabei aber einem historischen Wandel unterliegt, der gegenwärtig vor allem anhand der Digitalisierung der wissenschaftlichen Kommunikation kontrovers diskutiert wird. Glücklicherweise hat das Peer Review als Forschungsgegenstand in den letzten Jahren vermehrt Aufmerksamkeit erfahren, so dass an vielen Stellen auf relevante Forschungsliteratur zurückgegriffen werden kann. Wie noch darzustellen sein wird, ist der Forschungsstand jedoch sehr uneinheitlich was verschiedene Aspekte des Peer Review betrifft und insbesondere theoretische Arbeiten sind äußerst selten.

## Historische Entstehung und Verbreitung des Peer Review

Obwohl Peer Review heute als fast selbstverständlicher Bestandteil der Wissenschaft gesehen wird, beginnt sein weitverbreiteter Einsatz bei Zeitschriften und Forschungsförderungsorganisationen erst in der zweiten Hälfte des 20. Jahrhunderts. Davor kommt es eher vereinzelt und informell zum Einsatz, so dass sich Einstein in der oben beschriebenen Anekdote noch 1936 überrascht zeigen kann, wenn es bei einer Zeitschrift benutzt wird. Historisch reichen die Anfänge des Peer Review jedoch bis ins 17. Jahrhundert zu den in Europa entstehenden wissenschaftlichen Gesellschaften, wie bspw. der Royal Society in England, zurück. Diese gelten als Ort der sog. wissenschaftlichen Revolution und entstehen aus elitären Klubs zur Diskussion neuer naturphilosophischer Erkenntnisse (Shapin 1994). Universitäten sind zu diesem Zeitpunkt eher Orte der Wissensvermittlung, aber nicht der Wissensproduktion. Ziel hinter der Gründung wissenschaftlicher Gesellschaften ist die Schaffung eines sozialen Raumes, in dem frei und unvoreingenommen über Forschungsergebnisse diskutiert werden kann, die entweder aus den Experimenten der eigenen Mitglieder stammen oder durch Zuschriften an die Gesellschaft berichtet werden (“republic of letters”, Daston 1991).

Damit diese wissenschaftlichen Gesellschaften einen solchen Raum entstehen lassen können, müssen sie zwei Herausforderungen begegnen, die einerseits ihre Unabhängigkeit und andererseits die Qualität des produzierten Wissens betreffen. Eine freie und unabhängige Diskussion ist nur möglich, wenn andere gesellschaftliche Kräfte und Interessen (insb. kirchliche) minimiert werden können und wenn die Wahrheitsorientierung der Beteiligten maximiert werden kann (Shapin 1994). Beides, Unabhängigkeit nach außen und Wahrheitsorientierung nach innen, wird im ersten Schritt durch den Adelsstand der Mitglieder sichergestellt. Adlingen (im Englischen: peers) wird zugeschrieben,

dass sie aufgrund ihrer gesicherten sozialen Stellung und ihres Reichtums relativ unvoreingenommen Wahrheitsfragen nachgehen können. Diese herausgehobene soziale Stellung macht es zudem unwahrscheinlich, dass sie das neue Wissen dazu nutzen, die bestehende gesellschaftliche Ordnung (König, Kirche) zu gefährden. Damit in den Gesellschaften nicht nur frei diskutiert, sondern die Ergebnisse auch gedruckt und verbreitet werden können, müssen die entstehenden Zeitschriften und Buchreihen von der staatlichen Zensur ausgenommen werden (Biagioli 2002). Die Gesellschaften verpflichten sich auf einen kritischen internen Diskurs, den sie u. a. durch elitäre Mitgliedschaftsregeln absichern, der das neue Wissen auf Wahrheit und auf Publizierbarkeit prüft. Die Treffen der Mitglieder, in denen Experimente und Zuschriften diskutiert werden, können deshalb als eine erste Form der kollegialen wissenschaftlichen Begutachtung, des Peer Review, gelten.

Eine umfassende historische Aufarbeitung der Entwicklung des Peer Review liegt bisher nicht vor, aber der für die Gegenwart entscheidende Moment scheint aus zwei Gründen in der Mitte des 20. Jahrhunderts zu liegen. Einerseits hat das Peer Review bis dahin eine gewisse Verbreitung in wissenschaftlichen Zeitschriften gefunden, so dass es den meisten Forschenden zumindest bekannt ist, auch wenn sie damit noch keine Erfahrung als Verfahrensbeteiligte gehabt haben sollten. Andererseits werden nach dem zweiten Weltkrieg in vielen forschungsstarken Nationen Forschungsförderungsorganisationen geschaffen, die in Ergänzung zu den Hochschulen Forschungsprojekte finanziell ermöglichen, sog. Drittmittförderung. Anfänglich werden die Entscheidungen darüber, welche Projekte gefördert werden sollen, noch von Programmverantwortlichen innerhalb der Organisationen getroffen. Sowohl die Bedeutung als auch die staatlich aufgewendeten Mittel der Drittmittforschung nehmen kontinuierlich zu, bis diese auch mehr Aufmerksamkeit von der Wissenschaftspolitik erfahren.

Insbesondere ab den 1970er Jahren werden in den Parlamenten und in der Öffentlichkeit teilweise kontroverse Diskussionen darüber laut, welche Art von Forschung finanziert werden soll. So sind es u. a. Kontroversen um Forschung zur menschlichen Sexualität, die zu Steuerungsversuchen durch die Politik auf die Forschungsförderung führen (Baldwin 2018). In der Folge setzen Forschungsförderungsorganisationen zunehmend das von Zeitschriften bekannte Peer Review ein, um der Öffentlichkeit zu signalisieren, dass die Wahl von Forschungsthemen keinesfalls politisch motiviert ist, sondern nach innerwissenschaftlichen Qualitätskriterien stattfindet. Damit werden Forschungsförderungsorganisationen zu Grenzorganisationen (Guston 2000), die ebenso wie die wissenschaftlichen Gesellschaften des 17. Jahrhunderts Autonomie der Wissenschaft nach außen und wissenschaftliche Qualität nach innen absichern. Erst zu diesem Zeitpunkt wird das Peer Review zum selbstverständlichen Verfahren (Baldwin 2018), das überall da eingesetzt werden kann, wo sich wissenschaftliche Qualitätsfragen stellen, und in der Folge breitet sich dieses verstärkt bei Zeitschriften und in der Forschungsförderung aus.

## Definition und Funktionen von Peer Review

Die starke Verbreitung des Peer Review in der Gegenwart hat dazu geführt, dass eine Definition dessen, was Peer Review ausmacht, schwierig und gar umstritten geworden ist (Hirschauer 2004, Schendzielorz/Reinhart 2020). Umstritten deshalb, weil Peer Review auch zu einer Art Gütesiegel geworden ist, das Zeitschriften, Förderorganisationen oder Forschende nutzen, um die Qualität ihrer

Arbeit auszuweisen. Schwierig ist die Definition deshalb, weil Peer Review in so vielen verschiedenen Varianten zum Einsatz kommt und eben nicht nur bei Zeitschriften und in der Forschungsförderung, sondern bspw. sogar bei der Hochschulakkreditierung (Schneijderberg et al. 2022). Dagegen wird oft versucht die Bezeichnung Peer Review auf Begutachtungsverfahren bei wissenschaftlichen Zeitschriften zu begrenzen, bei denen eingereichte Manuskripte an zwei bis drei Fachkolleg\*innen zur Kommentierung und Bewertung verschickt werden und von den Herausgeber\*innen darauf basierend dann über Publikation, Überarbeitung oder Ablehnung entschieden wird. Mit Blick auf die historische Herleitung und die beiden Herausforderungen der Autonomiesicherung nach außen und der Qualitätssicherung nach innen wird in diesem Beitrag jedoch eine umfassendere Definition zugrunde gelegt. Als Peer Review soll hier jedes Verfahren in der Wissenschaft gelten, mit dem knappe Ressourcen verteilt werden, indem Fachurteile herangezogen werden, um die epistemische Qualität von wissenschaftlichen Objekten zu bewerten (Reinhart/Schendzielorz 2024). Als Objekt kommen bspw. Artikel, Bücher, Forschungsprojekte, Karrieren, Förderprogramme oder Organisationen in Frage, während als Ressourcen Publikationsfläche, finanzielle Förderung, Stellen, Aufmerksamkeit oder Prestige verteilt werden. Damit wird ein großer Teil der institutionellen Evaluationsverfahren in der Wissenschaft auch jenseits des Publikationswesens und der Forschungsförderung, bspw. bei Konferenzbeiträgen, der Besetzung von Professuren oder Programmevaluationen, abgedeckt.

Weil im Peer Review fachliche Qualitätsurteile und Entscheidungen über Mittelverteilung verbunden sind, ist die Vorstellung unter den Beteiligten weit verbreitet, dass es sich um eine Art Messverfahren mit abschließendem Qualitätsurteil handelt. Forschende nehmen dies oft so wahr, insbesondere wenn ihr Manuskript oder ihr Forschungsantrag abgelehnt wird. In der Folge konzentrieren sich sowohl die Fachliteratur als auch die praktischen Versuche zur Verfahrensverbesserung darauf, diese Einzelentscheidungen so zu optimieren, dass sie möglichst effizient und vorurteilsfrei zustande kommen. Dementsprechend sind Fairness, Minimierung von Bias und Einhelligkeit der Gutachter\*innenurteile wichtige normative Bezugspunkte in diesem Diskurs. Die daran anschließenden Forderungen zielen dann primär auf mehr Transparenz, (formalisierbare) Qualitätskriterien und, im Extremfall, quasi-algorithmische Auswahlverfahren (wie bspw. Lotterien; siehe z. B. Reinhart/Schendzielorz 2020). Aus Sicht der Wissenschaftsforschung stellt sich das als Praxisperspektive dar, die einseitig auf die individuellen Verfahrensergebnisse fokussiert und dabei vorschnell evaluativ urteilt (Hirschauer 2004). Aus Sicht der Forschung über Peer Review als Forschungsgegenstand gilt es evaluative Fragen auszuklamern und das Phänomen Peer Review hinsichtlich drei grundsätzlicher Fragen zu beforschen: Was wissen wir darüber, was innerhalb von Peer Review Verfahren überhaupt geschieht (Prozess)? Welche Ergebnisse lassen sich aus den Verfahren festhalten (Outcome)? In welche größeren Kontexte ist Peer Review eingebettet (Kontext)?

Zur Frage nach dem Prozess, also wie Peer Review Verfahren praktisch ablaufen, gibt es vergleichsweise wenig empirische Forschung. Aus öffentlich zugänglichen Beschreibungen von Verfahren durch Zeitschriften und Forschungsförderungsorganisationen lässt sich schliessen, dass es eine große Bandbreite von Verfahren gibt (Schendzielorz/Reinhart 2020). Im Minimalfall folgt auf die Einreichung eines Manuskripts oder Antrags eine schriftliche Begutachtung durch selten mehr als zwei oder drei fachliche Expert\*innen und eine darauf basierende Entscheidung durch Verfahrensverantwortliche, wie bspw. Zeitschriftenherausgeber\*innen. Oft haben die Verfahren aber zusätzliche Elemente: Die Entscheidung wird nicht durch Einzelne, sondern durch ein Gremium getroffen, das gleichzeitig eine

größere Zahl von Einzelentscheidungen in Abhängigkeit voneinander trifft (Lamont 2009, Olbrecht 2014). Die Gutachtenden beurteilen mehrere Einreichungen gleichzeitig und treffen sich in der Gruppe (Panel), um im Gespräch Einigkeit über ihre Urteile herzustellen. Die Einreichenden werden nach einer Vorauswahl persönlich eingeladen, um sich den kritischen Fragen eines Begutachtungspanels zu stellen. Die Einreichungen werden in Bestandteile aufgeteilt, bspw. Projekтиde und Lebenslauf der Einreichenden, die dann von unterschiedlichen Gutachtenden nach spezifischen Kriterien beurteilt werden. In der Kombination solcher Verfahrenselemente finden sich von sehr einfachen bis sehr komplexen Peer Review Verfahren zahlreiche Möglichkeiten, um verschiedenste Begutachtungs- und Entscheidungsbedarfe abzudecken. Sehr aufwändige Verfahren, manchmal mit mehr als zehn Verfahrensschritten, finden sich v. a. in der Forschungsförderung von großen Verbundvorhaben mit einer größeren Zahl von gemeinsam antragstellenden Forschenden und Forschungsinstitutionen (Schendzielorz & Reinhart 2020).

Studien, die das Bewertungsgeschehen direkt beforschen, gibt es relativ wenige, da einerseits der Zugang zu den als vertraulich angelegten Peer Review Verfahren aufwändig ist und oft verwehrt wird. Andererseits ist das Geschehen auch bei vorhandenem Zugang oft schwer empirisch greifbar, da zentrale Aspekte wie das Lesen von Einreichungen und das Schreiben von Gutachten nur sehr vermittelt beobachtet werden können. Studien, denen das trotzdem gelingt, liefern deshalb besonders wichtige Einsichten. So konnte durch teilnehmende Beobachtung bspw. gezeigt werden, dass in interdisziplinären Begutachtungspanels neben den Bewertungskriterien vor allem Standards für Fairness ausgehandelt werden, um unterschiedliche Fachexpertisen abwägen zu können (Lamont 2009). Gezeigt werden konnte auch, dass die Herausgebergruppen von Zeitschriften in eine hochkomplexe Kommunikationsdynamik eingebunden sind (Hirschauer 2005). Untereinander sowie mit den Autor\*innen und Gutacher\*innen entstehen über die einzelne Entscheidung hinaus strategische Argumentations- und Beeinflussungsmuster, die deswegen aber nicht zu einseitigen, interessegeleiteten oder gar korrupten Entscheidungen führen. Vielmehr zeigt sich das Peer Review als ein Ort, an dem Einreichungen mit großer Aufmerksamkeit gelesen, diskutiert, verbessert und im Rahmen einer inneren Öffentlichkeit bewertet werden (Hirschauer 2019). Das Peer Review leistet damit, trotz eingeschränkter Transparenz, in erheblichem Masse den wissenschaftlichen Anspruch auf organisierten Skeptizismus (Merton 1973) im Wissensproduktionsprozess.

Zur Frage nach den Outcomes, also welche Ergebnisse das Peer Review produziert, finden sich hunderte von (quantitativen) Studien, die Bewertungen und Entscheidungen zu einzelnen Einreichungen betrachten (Weller 2001, Guthrie et al. 2018). Diese fragen meist danach, ob sich die Gutachtenden in ihrem Urteil einig sind (Reliabilität), ob die Entscheidungen mit anderen Qualitätssäulen übereinstimmen (Validität) oder ob sich Voreingenommenheiten in den Entscheidungen nachweisen lassen (Bias). Die Aussagekraft der meisten dieser Studien ist entweder dadurch beschränkt, dass die Fallzahlen eher gering sind oder dass sie jeweils nur Daten aus einem einzelnen Verfahren haben. Vergleichende Studien oder solche mit einem sehr robusten Forschungsdesign gibt es wenige. Trotzdem lässt sich festhalten, dass die Reliabilität bei unabhängig voneinander Gutachtenden eher gering ist. Die Validität scheint insofern hoch, als Peer Review Entscheidungen gut mit dem früheren Publikationserfolg der Einreichenden korreliert. Bezüglich Bias ist das Bild wenig einheitlich für Benachteiligungen aufgrund von Alter und Geschlecht. Gut nachgewiesen scheint hingegen eine Benachteiligung aufgrund der Nationalität bei internationalen (englischsprachigen) Zeitschriften.

Eher selten diskutiert wird das robusteste Ergebnis dieser Forschung, nämlich dass der Erfolg einer Einreichung etwa zur Hälfte durch zufällige Faktoren bei der Auswahl der Gutachtenden bestimmt ist (Cole et al. 1981). Dass die Entscheidungen im Einzelfall erheblich durch Zufälligkeiten beeinflusst wird, verträgt sich schlecht mit dem Selbstverständnis der Beteiligten, ist aber aus einer aggregierten oder systemischen Perspektive aufschlussreich (Reinhart 2012).

Über eine Einzelentscheidung über Publikation oder Finanzierung hinaus ist auffällig, dass Forschende nicht nur bewertet werden, sondern auch selbst bewerten. Die Gutachtenden müssen schließlich Fachkolleg\*innen mit geeigneter Expertise und damit aus derselben Fachgemeinschaft sein. Dieser regelmäßige Rollenwechsel führt zu einer professionellen Selbstkontrolle, die übermäßig scharfe oder ungerechtfertigte Kritik reduziert und eine Kultur der gutachterlichen Kollegialität etabliert (Neidhardt 2016). Dementsprechend findet sich der sprichwörtliche Reviewer 2, der übermäßig scharf oder gar korrupt begutachtet, in empirischen Untersuchungen von Gutachten allenfalls als Ausnahme. Die meisten Gutachten zeugen vom ernsthaften Versuch, ein sachgerechtes Urteil abzugeben. Für die Vermutung, dass mit einer an Kollegialität orientierten Begutachtung Ideen außerhalb des Mainstreams benachteiligt sein könnten, gibt es hingegen anfängliche Evidenz, auch wenn ein systematischer Konservatismus schwer nachzuweisen zu sein scheint. Aufschlussreich sind zudem Studien, die die Annahme- oder Ablehnungshäufigkeit von Peer Review Verfahren im Vergleich und über die Zeit untersuchen. Insbesondere in der Forschungsförderung zeigt sich eine starke Zunahme der Ablehnungshäufigkeit über die letzten Dekaden, so dass Förderprogramme keine Seltenheit mehr sind, bei denen mehr als 90% der Einreichungen abgelehnt werden. In der Folge stellen sich zunehmend auch Fragen nach den Kosten all dieser Einreichungs-, Begutachtungs- und Entscheidungsarbeit, zu denen es aber nur wenig belastbare Zahlen gibt (Guthrie et al. 2017).

Zur Frage nach den Kontexten, also in welche größeren systemischen Prozesse das Peer Review eingebettet ist, kann mehrheitlich theoretische Literatur herangezogen werden. Soziologisch lässt sich das Peer Review allgemein als eine Form professioneller Selbstkontrolle oder Selbststeuerung einordnen (Weingart 2001). Im Vergleich zu anderen Professionen, bspw. der praktischen Medizin, fällt auf, dass Kontrolle sehr dezentral erfolgt, indem die Wissenschaft von unzähligen selbstorganisierten Peer Review Verfahren durchsetzt ist. Auch auffällig ist, dass diese Form der Selbstkontrolle nicht an inhaltlichen Standards orientiert ist, die sich formalisieren ließen, wie das bspw. bei Standards für medizinische Behandlungen der Fall ist. Schließlich deutet dies auf den besonderen wissenschaftlichen Anspruch, die Qualität von Forschung nicht abschließend zu prüfen, sondern einem kritischen Diskurs zu unterziehen, der primär an neuem oder gar originellem Wissen interessiert ist. In diesem Sinn kann das Peer Review als eine Form professioneller Selbstkontrolle gesehen werden, die auf minimal formalisierte Einzelfallprüfungen setzt, deren Aufwand sich vor allem dadurch rechtfertigt, dass ein organisierter Skeptizismus aufrechterhalten wird, der als produktiver Teil und nicht als abschließende Prüfung der Erkenntnisproduktion gedacht ist (→ III.1.3 Profession und akademische Selbstorganisation). Über die Funktion im Erkenntnisprozess hinaus werden mittels Peer Review knappe Ressourcen wie Publikationsfläche, Forschungsmittel oder Aufmerksamkeit und Reputation verteilt. Darüber entscheidet sich, wie akademische Karrieren verlaufen, welche Forschungsthemen zentral sind und welche Texte gelesen werden sollen (Merton 1957, 1968).

Während das Moment der Selbstkontrolle oder Selbststeuerung auf das Innern der Wissenschaft gerichtet ist, kommt dem Peer Review auch eine entscheidende Rolle im Verhältnis der Wissenschaft

nach außen, bspw. zu Politik oder Öffentlichkeit, zu. Der oben erwähnte Fall der US-amerikanischen Wissenschaftspolitik der 1970er Jahre macht deutlich, dass sich die Wissenschaft nach außen darüber legitimiert, dass sie auf diesen Fall der Selbstkontrolle zeigen kann. Grenzorganisationen, wie bspw. nationale Forschungsförderungsorganisationen, rechtfertigen die Ausgabe von Steuermitteln für thematisch offene Grundlagenforschung dadurch, dass nur mittels Peer Review sichergestellt werden kann, dass qualitativ hochwertige Forschung gefördert wird (Reinhart 2012, Baldwin 2018). Als Schnittstelle zur Gesellschaft trägt das Peer Review auch zur Sicherung der Autonomie der Wissenschaft bei, so wie sie in Deutschland im Grundgesetz (Art. 5: „Kunst und Wissenschaft, Forschung und Lehre sind frei.“) verankert ist. Damit das Peer Review diese Funktion erfüllen kann, muss es sich den Veränderungen im Verhältnis von Wissenschaft und Gesellschaft anpassen können (Weingart 2001). So hat Musselin (2013) noch argumentiert, dass die Wissenschaftsprofession durch den gestiegenen Einsatz von Peer Review (z. B. in nationalen Exzellenzinitiativen) an Einfluss gewinnt. Aktuell wird hingegen häufiger diskutiert, ob das Peer Review in dieser Funktion nicht überlastet werden könnte und eine Art Krise drohen würde. Durch zunehmende Evaluationsbedarfe und steigende Anteile der Forschungsfinanzierung, die durch Drittmittel erfolgt, sinken die Bewilligungsquoten und steigt die Kompetitivität im Wissenschaftssystem. Die im Peer Review getroffenen Entscheidungen sind dadurch immer folgenreicher, wirken aber aufgrund der hohen Ablehnungsquoten entweder als ob sie zufallsbasiert zustande kommen oder als ob es immer dieselben sind, die davon profitieren würden (Loveday 2018).

## Fazit und Ausblick

Das vermehrte Bewerten und Evaluieren in der Wissenschaft ist eingebettet in einen umfassenderen gesellschaftlichen Prozess, den Michael Power schon 1997 als „audit society“ (Power 1997) beschrieben hat. In der Wissenschaft sind deshalb nicht nur mehr und vielfältigere Peer Review-Verfahren zum Einsatz gekommen, sondern auch eine große Zahl von quantitativen Messverfahren, wie bspw. Hochschulrankings (Brankovic et al. 2018). Die gegenwärtige Wahrnehmung, dass das Peer Review in einer Krise sei, speist sich deshalb aus zwei Quellen. Einerseits hat die Zunahme an Peer Review-Verfahren zu einer Verschärfung der Konkurrenz bei der Rekrutierung von Gutachtenden geführt. Da Zeitschriften und Forschungsförderungsorganisationen hier aber vornehmlich, um eine kleine Wissenschaftselite zu konkurrieren scheint, ließe sich dieser Verschärfung durch eine Ausweitung und damit Diversifizierung des Gutachtendenpools, aus dem rekrutiert wird, begegnen. Andererseits ist es der größere Kontext von Forschungsinformation und Forschungsevaluation, der zunehmend, als systemisches Problem betrachtet wird, das Fehlanreize und Qualitätsprobleme schafft. Forschende und Forschungsorganisationen, die prominent anhand ihres quantitativen Outputs bewertet werden, bspw. durch die Anzahl von Publikationen oder Zitationen, reagieren darauf notwendigerweise mit einer Steigerung genau dieser Aspekte, was teilweise zulasten der Qualität geht. Als Gegenmaßnahme wurde zuletzt in Europa eine Koalition zur Reform der Forschungsevaluation ins Leben gerufen (European Commission 2021, coara.org).

Vor diesem Hintergrund sind gegenwärtige Versuche einzuordnen, das Peer Review zu reformieren, indem bspw. mehr Transparenz geschaffen wird (Open Peer Review) oder indem früher im Erkennt-

nisproduktionsprozess begutachtet wird (präregistrierte Publikationen) oder indem neue technologische Lösungen (bspw. künstliche Intelligenz) eingesetzt werden. Es ist zu erwarten, dass einige dieser Maßnahmen dauerhaft im Peer Review zum Einsatz kommen werden und damit die Vielfalt der Verfahren noch weiter erhöht wird. Es deutet alles darauf hin, dass die Bedeutung des Peer Review damit eher noch steigen wird. Was dabei bisher aber kaum Beachtung findet ist die Frage, woher der Aufwuchs dieser, bisher auf Freiwilligkeit basierenden, wissenschaftlichen Arbeit des Organisierens und Begutachtens kommen soll? Da diese Arbeit vor allem im Rahmen von hochschulischen Beschäftigungsverhältnissen stattfindet, böte sich hier ein explizit hochschulforscherischer Zugang an, der in der bisherigen Peer Review-Forschung eher selten vorzufinden ist.

## Literaturempfehlungen

Shapin, Steven (1994) rekonstruiert im Buch „A Social History of Truth: Civility and Science in Seventeenth-Century England“ den sozio-historischen Prozess, der die Herausbildung eines eigenständigen gesellschaftlichen Bereichs für Wissenschaft vorbereitet, der sich dadurch auszeichnet, dass Kritik an vorgebrachten Wissensbehauptungen nicht als gesellschaftliche Rufschädigung betrachtet werden muss, sondern als kollektiver Lernprozess gedeutet werden kann.  
Biagioli, Mario (2002) schliesst an Shapin an und zeigt in „From Book Censorship to Academic Peer Review“, dass das Peer Review schon bei der Entstehung im 17. Jahrhundert sowohl innerwissenschaftliche (Qualitätssicherung) als auch gesellschaftlich-politische (Zensur) Funktionen hatte.  
Stefan Hirschauer (2019) zeigt im Artikel „Urteilen unter Beobachtung“ wie der oft geäußerte Vorwurf, das Peer Review sei intransparent, einer soziologischen Analyse nicht standhält, da die Begutachtungspraxis in einem komplexen Geflecht von Kommunikations- und Beobachtungsverhältnissen stattfindet.

## Literaturverzeichnis

- Baldwin, Melinda (2018): Scientific Autonomy, Public Accountability, and the Rise of „Peer Review“ in the Cold War United States, in: Isis, Jg. 109, Nr. 3, 538–558, doi: 10.1086/700070.
- Biagioli, Mario (2002): From Book Censorship to Academic Peer Review. In: Emergences: Journal for the Study of Media & Composite Cultures, Jg. 12, Nr. 1, 11–45, doi: 10.1080/104572202200003435.
- Brankovic, Jelena/Ringel, Leopold/Werron, Tobias (2018): How rankings produce competition: The case of global university rankings. In: Zeitschrift für Soziologie, Jg. 47, Nr. 4, 270–288.
- Cole, S./Cole, J. R./Simon, G. A. (1981): Chance and consensus in peer review. In: Science (New York, NY.), Jg. 214, Nr. 4523, 881–886.
- Daston, Lorraine (1991): The Ideal and Reality of the Republic of Letters in the Enlightenment. In: Science in Context, Jg. 4, Nr. 2, 367–386, doi: 10.1017/S0269889700001010.
- European Commission, Directorate-General for Research and Innovation (2021): Towards a reform of the research assessment system: scoping report, LU: Publications Office of the European Union.
- Guston, David H. (2000): Between Politics and Science: Assuring the Integrity and Productivity of Research, 2. Aufl. Cambridge/New York: Cambridge University Press.
- Guthrie, Susan/Ghiga, Ioana/Wooding, Steven (2018): What do we know about grant peer review in the health sciences? In: F1000Research, Jg. 6, 1335.
- Hirschauer, Stefan (2004): Peer Review Verfahren auf dem Prüfstand: Zum Soziologiedefizit der Wissenschaftsevaluation. In: Zeitschrift für Soziologie, Jg. 33, Nr. 1, 62–83.

- Hirschauer, Stefan (2005): Publizierte Fachurteile. Lektüre und Bewertungspraxis im Peer Review. In: Soziale Systeme, Jg. 11, Nr. 1, 52–82.
- Hirschauer, Stefan (2019): Urteilen unter Beobachtung. In: Nicolae, Stefan/Endreß, Martin/Berli, Oliver u. a. (Hg.), (Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit, Wiesbaden: Springer Fachmedien Wiesbaden (Soziologie des Wertens und Bewertens), 275–298, doi: 10.1007/978-3-658-21763-1\_12.
- Kennefick, Daniel (2005): Einstein Versus the Physical Review, in: Physics Today, Jg. 58, Nr. 9, 43–48, doi: 10.1063/1.2117822.
- Lamont, Michèle (2009): How Professors Think: Inside the Curious World of Academic Judgment. Harvard: Harvard University Press.
- Latour, Bruno/Woolgar, Steve (1986): Laboratory Life: The Construction of Scientific Facts, 2nd Edition with a New Postscript. London: Sage Publications.
- Loveday, Vik (2018): „Luck, chance, and happenstance? Perceptions of success and failure amongst fixed-term academic staff in UK higher education“. In: The British Journal of Sociology, Jg. 69, Nr. 3, 758–775, doi: 10.1111/1468-4446.12307.
- Merton, Robert K. (1957): Priorities in scientific discovery: A chapter in the sociology of science. In: American Sociological Review, US: American Sociological Assn, Jg. 22, 635–659, doi: 10.2307/2089193.
- Merton, Robert K. (1968): The Matthew Effect in Science: The reward and communication systems of science are considered. In: Science, Jg. 159, Nr. 3810, 56–63, doi: 10.1126/science.159.3810.56.
- Merton, Robert K. (1973): The Normative Structure of Science. In: Storer, Norman W. (Hg.), The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University of Chicago Press, 267–279.
- Musselin, Christine (2013): How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate. In: Research Policy, Jg. 42, Nr. 5, 1165–1173, doi: 10.1016/j.respol.2013.02.002.
- Olbrecht, Meike (2014): Entscheidungsfindungsprozesse von Gutachtergruppen, Berlin: Humboldt-Universität zu Berlin, Philosophische Fakultät III, doi: 10.18452/16946.
- Power, Michael (1997): The Audit Society: Rituals of Verification. Oxford: Oxford University Press.
- Reinhart, Martin (2012): Soziologie und Epistemologie des Peer Review. Baden-Baden: Nomos.
- Reinhart, Martin/Schendzielorz, Cornelia (2024): Peer-review procedures as practice, decision, and governance – the road to theories of peer review. In: Science and Public Policy, Jg. 51, Nr. 3, 543–552, doi: 10.1093/scipol/scad089.
- Reinhart, Martin/Schendzielorz, Cornelia (2020): The lottery in Babylon – On the role of chance in scientific success. In: Journal of Responsible Innovation, Jg. 7, Nr. sup2, 25–29, doi: 10.1080/23299460.2020.1806429.
- Schendzielorz, Cornelia/Reinhart, Martin (2020): Die Regierung der Wissenschaft im Peer Review / Governing Science Through Peer Review. In: dms – der moderne staat – Zeitschrift für Public Policy, Recht und Management, Jg. 13, Nr. 1, 101–123.
- Schneijderberg, Christian/Götze, Nicolai/Steinhardt, Isabel u. a. (2022): (Be)Werten im organisationsabhängigen Interaktionssystem: das Beispiel Peer Review in Akkreditierungsverfahren von Hochschulen und Studiengängen. In: Zeitschrift für Evaluation, Jg. 2022, Nr. 02, 217–244, doi: 10.31244/zfe.2022.02.03.
- Shapin, Steven (1994): A Social History of Truth: Civility and Science in Seventeenth-Century England. Chicago: University of Chicago Press.
- Weingart, Peter (2001): Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft. Weilerswist: Velbrück Wissenschaft.
- Weller, Ann C. (2001): Editorial peer review: its strengths and weaknesses. Medford, N. J.: Information Today.

