

Zeitschrift:	Schweizerische Zeitschrift für Soziologie = Revue suisse de sociologie = Swiss journal of sociology
Herausgeber:	Schweizerische Gesellschaft für Soziologie
Band:	33 (2007)
Heft:	2
Artikel:	Peer-Review in der Wissenschaft : eine Analyse des Begutachtungsverfahrens aus der Sicht wissenschaftssoziologischer Theorien
Autor:	Bornmann, Lutz
DOI:	https://doi.org/10.5169/seals-814230

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 29.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Peer-Review in der Wissenschaft – eine Analyse des Begutachtungsverfahrens aus der Sicht wissenschaftssoziologischer Theorien

Lutz Bornmann*

1 Peer-Review in der Wissenschaft

In der modernen Wissenschaft ist das Peer-Review zum wichtigsten Bewertungsinstrument von wissenschaftlicher Arbeit avanciert (Ziman, 2000). Mit dem Verfahren werden nicht nur Manuskripte für die Publikation in Zeitschriften ausgewählt, es bestimmt auch darüber, wem Preise (wie z. B. der Nobelpreis), Stipendien und Arbeitsstellen zugewiesen werden (Hemlin und Rasmussen, 2006). Darüber hinaus wird es für die Evaluation von Forschungsgruppen und wissenschaftlichen Institutionen eingesetzt (Hemlin, 1996). Durch den Sparzwang der Universitäten zeichnet sich bei der Finanzierung von Forschungsvorhaben seit Jahren der Trend ab, dass Wissenschaftler immer weniger auf reguläre Forschungsmittel ihrer Universität zurückgreifen können und sich immer häufiger um externe Forschungsmittel bemühen müssen, die mit dem Peer-Review-Verfahren vergeben werden (Guston, 2003). In den typischen Anwendungsbereichen des Peer-Review, der Begutachtung von Manuskripten und Forschungsanträgen, haben Gutachter – als «Gralshüter» der Wissenschaft – im Begutachtungsprozess die Aufgabe, die «beste» wissenschaftliche Forschung unter der Bedingung knapper Ressourcen (d. h. begrenzten Forschungsmitteln bzw. einer begrenzten Anzahl von Druckseiten in Zeitschriften) zur Auswahl zu empfehlen (Hackett und Chubin, 2003). Darüber hinaus sollen die Gutachter Fehler in wissenschaftlichen Arbeiten oder wissenschaftliches Fehlverhalten im Zusammenhang mit den Arbeiten aufdecken (Smith, 2006).

Die Befürworter des Peer-Review halten das Verfahren für die Prüfung und Legitimation von wissenschaftlicher Arbeit für geeigneter als jedes andere bislang vorgeschlagene Verfahren (siehe z. B. die Vorschläge von Roy, 1985). Aktive Wissenschaftler aus demselben Forschungsgebiet werden als die am besten geeigneten Personen angesehen, die Arbeit ihrer Fachkollegen auf ihre wissenschaftliche Güte hin zu begutachten (Eisenhart, 2002). Die Kritiker des Peer-Review sehen die Schwächen des Verfahrens insbesondere darin, dass (1) verschiedene Gutachter kaum in der Bewertung ein und derselben wissenschaftlichen Arbeit übereinstimmen, dass (2) gutachterliche Empfehlungen systematische Urteilsfehler aufweisen (die Urteile

* Lutz Bornmann, Professur für Sozialpsychologie und Hochschulforschung, ETH Zürich, Zähringerstr. 24, CH-8092 Zürich, E-Mail: bornmann@gess.ethz.ch, Telefon: +41 44 632 48 25, Fax: +41 44 632 12 83.

würden nicht nur auf der wissenschaftlichen Güte einer Arbeit basieren, sondern auch auf nicht wissenschaftlichen Kriterien) sowie dass (3) der Zusammenhang zwischen den Urteilen im Peer-Review und der Qualität der begutachteten Arbeit gering sei (siehe zur Kritik am Peer-Review z. B. Eysenck und Eysenck, 1992; Ross, 1980). Der einzige Grund für die weitere Verwendung des Verfahrens sei, dass es keinen klaren Konsens über eine «bessere» Alternative gäbe (Young, 2003).

Die Peer-Review-Forschung, die die Kritik am Verfahren aufgegriffen und das Verfahren systematisch untersucht hat, bezieht sich zum grössten Teil auf das Peer-Review bei Zeitschriften (siehe im Überblick Campanario, 1998a; Campanario, 1998b; Overbeke und Wager, 2003; Weller, 2002), etwas seltener auf das Peer-Review bei Forschungs- und Stipendienanträgen (siehe im Überblick Bornmann und Daniel, 2003; Demicheli und Pietrantonj, 2004; Wessely, 1998). Forschung in den anderen Anwendungsbereichen des Peer-Review gibt es kaum (ausser z. B. von Bornmann, Mittag und Daniel, 2006; Wissenschaftliche Kommission Niedersachsen, 2006). Auch wenn bis heute eine Vielzahl von Studien zum Peer-Review durchgeführt wurde (Weller, 2002, recherchierte für den bis heute umfangreichsten Literaturüberblick zur Forschung in der Manuskriptbegutachtung 1439 Studien), sind davon nur wenige theoriegeleitet durchgeführt worden (Hirschauer, 2004). «There is a stark discrepancy between the number of empirical ‹peer review studies› and the theoretical understanding of the process» (Gläser und Laudel, 2006, 187). Die wenigen theoriegeleiteten empirischen Studien liegen überwiegend aus der älteren Peer-Review-Forschung vor, die durch die so genannte Nordamerikanische Schule von Merton (1973) geprägt war (siehe z. B. Cole, Cole und Simon, 1981; Cole und Rubin, 1978).

In der Peer-Review-Forschung war mit dem Verlust der dominierenden Stellung der Nordamerikanischen Schule in der Wissenschaftssoziologie zugunsten des Sozialkonstruktivismus in den späten 1970er- und frühen 1980er-Jahren der Übergang von einer (eher) theoriegeleiteten zu einer (eher) theorielosen empirischen Analyse des Begutachtungsverfahrens verbunden. «After Mertonian sociology of science had been supplanted (...) the literature shifts to atheoretical, empiricist approaches by scientists and editors who were mainly interested in the validity and reliability of peer review processes» (Gläser und Laudel, 2006, 187). Die sozialkonstruktivistische Wissenschaftssoziologie hat sich in den ethnmethodologischen Studien mit dem Thema Peer-Review nur ansatzweise beschäftigt (siehe dazu die wenigen Ansätze in Knorr-Cetina, 1981; Myers, 1990). Nach Gläser (2006) ist dieser «blinde Fleck» «der Tatsache geschuldet, dass der Peer-Review keine lokale, an einem Ort beobachtbare Interaktion ist, sondern eine räumlich verteilte Interaktion (...) Der Gesamtprozess entzieht sich ethnographischer Beobachtungen» (96). Da sich die sozialkonstruktivistische und die systemtheoretische Perspektive an ähnlichen methodologischen Prämissen orientiert, hat sich auch die Theorie sozialer Systeme nach Luhmann (1992; 1998) mit dem Peer-Review bislang kaum befasst.

Diese Entwicklung hat zu der Situation geführt, dass bei der empirischen Analyse des wissenschaftlichen Begutachtungsprozesses die Theorien aus der Wissenschaftssoziologie (insbesondere neuere Theorien) unberücksichtigt bleiben. Die Annahmen der Nordamerikanischen Schule werden vielfach als überholt angesehen und eine umfassendere theoretische Analyse des Peer-Review aus der Sicht des Sozialkonstruktivismus und der Theorie sozialer Systeme ist bislang ausgeblieben. Deshalb ist es das Ziel des vorliegenden Beitrags, einen theoretischen Rahmen für das Peer-Review vorzustellen, bei dem auch die neuere Theorieentwicklung berücksichtigt wird und der der zukünftigen empirischen Peer-Review-Forschung zugrunde gelegt werden kann. Dazu wird das Peer-Review aus der Sicht der drei wichtigsten Theorierichtungen in der Wissenschaftssoziologie (siehe dazu Bornmann, zur Publikation angenommen; Cole, 1992) analysiert: (1) der Nordamerikanischen Schule, die von Merton (1938) begründet wurde, (2) des Sozialkonstruktivismus (Knorr-Cetina, 1981; Latour und Woolgar, 1979) sowie (3) der Theorie sozialer Systeme nach Luhmann (1992). Bevor wir die theoretischen Analysen des Peer-Review im Folgenden vornehmen werden (siehe Kapitel 3), möchten wir die wichtigsten empirischen Ergebnisse aus der Peer-Review-Forschung vorstellen (siehe Kapitel 2). Aus den Ergebnissen der empirischen Studien können wir Rückschlüsse auf die Validität der theoretischen Annahmen – angewendet auf das Peer-Review – ziehen.

2 Empirische Studien zum Peer-Review

Formal kann das Peer-Review-Verfahren als ein sozialer Urteilsprozess von Individuen in einer Kleingruppe (beispielsweise ein oder mehrere Gutachter sowie ein oder mehrere Herausgeber einer disziplinären Eigengruppe bei der Manuskriptbegutachtung) über eine wissenschaftliche Arbeit konzeptualisiert werden (Krampen und Montada, 2002). Da dem Peer-Review-Verfahren nur dann eine hohe Güte zugeschrieben werden kann, wenn es zu zuverlässigen (reliablen), fairen und gültigen (validen) Urteilen kommt (Hackett und Chubin, 2003), hat ein Grossteil der Studien aus der Peer-Review-Forschung diese Gütekriterien der empirischen Analyse zugrunde gelegt (Daniel, 1993, 2004).

2.1 Reliabilität des Peer-Review

Urteile (Empfehlungen und Entscheidungen) über ein und dieselbe wissenschaftliche Arbeit werden im Peer-Review nur dann als reliabel bezeichnet, wenn die Übereinstimmung der einzelnen Urteile hoch ist. Der Grad der Übereinstimmung zwischen den Urteilen von Gutachtern wurde in der Peer-Review-Forschung zumeist bei der Manuskriptbegutachtung untersucht. Insgesamt gesehen zeigen die Befunde der Studien, dass der Zusammenhang zwischen den Gutachterurteilen überwiegend in den Bereich zwischen 0,2 und 0,4 fällt (siehe im Überblick Cicchetti, 1991). Koef-

fizienten (Kappa- oder Intraclass-Korrelations-Koeffizienten) mit Werten unter 0,5 deuten auf eine eher niedrige Urteilsübereinstimmung und somit wenig zuverlässige Urteile im Peer-Review hin.

Eckberg (1991) und Kostoff (1995) weisen vor dem Hintergrund dieser Zahlen darauf hin, dass unterschiedliche Urteile bei der Begutachtung nicht unbedingt ein Zeichen für Uneinigkeit über die Qualität einer wissenschaftlichen Arbeit sind, sondern sich auch durch unterschiedliche Positionen (Paradigmen), Urteilskriterien und Kompetenzbereiche der Gutachter erklären lassen. Nach Cole (2000) spiegeln niedrige Korrelationskoeffizienten bei Gutachterurteilen den kaum vorhandenen Konsens wider, der in allen wissenschaftlichen Disziplinen im Bereich der «research frontiers» herrsche. In diesem Bereich sei es für gewöhnlich unmöglich, eine zuverlässige Begutachtung von wissenschaftlichen Arbeiten vorzunehmen. Hargens und Herting (1990) kritisieren an den Studien zur Übereinstimmung von Gutachterurteilen, dass bei diesen Studien zwar implizit unterstellt werde, die Gutachterurteile würden auf einer latent vorhandenen, wissenschaftlichen Qualitätsdimension variieren, diese Annahme aber mit der Berechnung von Kappa- oder Intraclass-Korrelations-Koeffizienten kaum geprüft werden könne. Hargens und Herting (1990) berechnen daher so genannte «RC association models» (Goodman, 1984). Ihre Untersuchung zur Manuskriptbegutachtung bei fünf Zeitschriften zeigt, dass bei vier Zeitschriften eine einheitliche Qualitätsdimension bei den Urteilen der Gutachter angenommen werden kann. Im Gegensatz zu den Ergebnissen der anderen Studien zur Reliabilität des Peer-Review deuten die Ergebnisse von Hargens und Herting (1990) demnach darauf hin, dass zwischen den Gutachterurteilen ein substantieller statistischer Zusammenhang existiert.

Obwohl in der Regel eine hohe Übereinstimmung zwischen den Gutachtern im Peer-Review-Verfahren angestrebt wird, wird sie jedoch gleichzeitig von einigen Wissenschaftlern als nachteilig für den Begutachtungsprozess angesehen: «Too much agreement is in fact a sign that the review process is *not* [Hervorhebung vom Verfasser] working well, that reviewers are not properly selected for diversity, and that some are redundant» (Bailar, 1991, 138). Nach Marsh und Ball (1991) wird zwar durch die Auswahl von Gutachtern, die unterschiedliche Perspektiven vertreten, verschiedene Urteilskriterien verwenden etc. die Reliabilität verringert. Die Validität kann aber deutlich erhöht werden, da die Entscheidungsträger (wie z. B. Herausgeber oder Programmmanager) ihr Urteil über eine wissenschaftliche Arbeit auf einer breiteren Informationsbasis treffen können. Deshalb sollten nach Auffassung der Wissenschaftler, die eine geringe Reliabilität als förderlich für den Begutachtungsprozess ansehen, Gutachter gerade aufgrund ihrer unterschiedlichen Perspektiven, Urteilskriterien etc. ausgewählt werden (Stricker, 1991).

2.2 Fairness des Peer-Review

In Veröffentlichungen zum Peer-Review (Hojat, Gonnella und Caelleigh, 2003; Owen, 1982; Pruthi et al., 1997; Ross, 1980; Sharp, 1990; Wood und Wessely, 2003) werden bis zu 25 verschiedene Urteilsfehler (Biases) genannt, die die Fairness des Verfahrens gefährden würden. Ein Urteilsfehler wird wie folgt definiert: «Bias is any feature of an evaluator's cognitive or attitudinal mind-set that could interfere with an objective evaluation» (Shatz, 2004, 36). Auch wenn bereits die Ergebnisse zahlreicher Studien auf einen Mangel an Fairness des Peer-Review hinweisen, ist die Forschung zur Fairness mit zwei grundsätzlichen Problemen konfrontiert, die die Generalisierung der Ergebnisse erschweren. Zum einen sind inkonsistente Ergebnisse zum Einfluss der einzelnen Biases auf die Urteile im Verfahren publiziert worden. So konnte beispielsweise in zahlreichen Studien ein Geschlechterbias im Peer-Review von Forschungs- und Stipendienanträgen nachgewiesen werden (z. B. Brouns, 2000; Wennerås und Wold, 1997). Eine ähnlich hohe Anzahl an Studien konnte jedoch nur einen moderaten bzw. keinen Geschlechtereffekt feststellen (z. B. Sandström und Hällsten, 2006; Ward und Donnelly, 1998) oder hat über uneinheitliche Ergebnisse berichtet (z. B. Bornmann und Daniel, 2005). Da es in der Peer-Review-Forschung keine experimentellen Studien gibt, ist zum zweiten bei den Ergebnissen der Studien kaum zu entscheiden, ob eine bestimmte Gruppe von Wissenschaftlern (wie z. B. Antragsteller oder Autoren von wenig prestigeträchtigen Universitäten) ungünstige Urteile aufgrund von Urteilsfehlern seitens der Gutachter erfahren oder ob die ungünstigen Urteile aus der mangelnden Qualität der wissenschaftlichen Arbeit resultieren (Daniel, 1993, 2004).

2.3 Prädiktive Validität des Peer-Review

Mit der Begutachtung von Anträgen und Manuskripten ist für gewöhnlich das Ziel verbunden, die «beste» wissenschaftliche Arbeit unter den eingereichten Arbeiten auszuwählen (Smith, 2006). Die Selektionsfunktion wird als ein schwer zu untersuchender Gegenstand angesehen. Nach Jayasinghe, Marsh und Bond (2001) existiert keine einheitliche Definition darüber, was ein «publikationswürdiges» Manuskript oder einen «förderungswürdigen» Antrag ausmacht. Da jedoch die Anzahl an Zitierungen für eine Publikation die internationale Wirkung der berichteten Forschungsergebnisse widerspiegelt (Daniel, 2005; van Raan, 2004), ist es ein gängiger Ansatz in der Peer-Review-Forschung, den Erfolg des Verfahrens anhand der Zitierhäufigkeit für die begutachtete Forschung zu überprüfen.

Daniel (1993, 2004), Lock (1985), Wilson (1978) sowie Ophof, Furstner, van Geer und Coronel (2000) haben die prädiktive Validität der Begutachtung von Manuskripten anhand der Zitierhäufigkeit von *angenommenen und abgelehnten*, jedoch an anderer Stelle publizierten Aufsätzen überprüft. Die Ergebnisse der vier Studien zeigen übereinstimmend, dass die von der *Angewandten Chemie*, dem *British Medical Journal*, dem *Journal of Clinical Investigation* und *Cardiovascular*

Research angenommenen Arbeiten nach ihrer Veröffentlichung im Durchschnitt deutlich häufiger zitiert wurden wie jene, die zunächst abgelehnt wurden und später andernorts erschienen sind.

3 Das Peer-Review aus der Sicht wissenschaftssoziologischer Theorien

In den folgenden Kapiteln werden in einem ersten Schritt die theoretischen Annahmen der Nordamerikanischen Schule (Kapitel 3.1), des Sozialkonstruktivismus (Kapitel 3.2) sowie der Theorie sozialer Systeme (Kapitel 3.3) dargestellt, und es wird in einem zweiten Schritt eine Übertragung dieser Annahmen auf das Peer-Review-Verfahren vorgenommen.

3.1 Die Sicht der Nordamerikanischen Schule auf das Peer-Review

Mit dem Ethos der Wissenschaft hat Merton (1973) in der Tradition des Strukturfunktionalismus jene Normen formuliert, die als (positive) Eigenschaften des Wissenschaftssystems vorgeben, welche Handlungen von Wissenschaftlern als angemessen bzw. korrekt anzusehen sind: (1) Veröffentlichung von Wissen (Kommunismus), (2) Profitverzicht (Uneigennützigkeit), (3) Prüfung von Erkenntnissen (organisierter Skeptizismus) sowie (4) Anerkennung von Erkenntnissen, unabhängig von ihrem Ursprung (Universalismus). Das Ethos der Wissenschaft habe sich seit Beginn der modernen Wissenschaft in der Interaktion von (professionellen) Wissenschaftlern herausgebildet (Merton, 1973). Als präskriptive Normen (siehe Coleman, 1995) würden die Normen bestimmte Handlungen vorschreiben und würden so eine implizite Verfassung in der Wissenschaft schaffen. Als Imperative seien die Normen bei den Wissenschaftlern der heutigen Zeit durch die Sozialisation in wissenschaftlichen Institutionen internalisiert (Merton, 1985).

Nach Merton (1973) gilt das Peer-Review-Verfahren als institutioneller Garant dafür, dass sich die Wissenschaft an den Vorschriften und Grundsätzen «guten» wissenschaftlichen Handelns – dem Ethos der Wissenschaften – orientiert. Nur mit einer institutionalisierten Form der Begutachtung (d. h. mit einem anerkannten Qualitätssicherungssystem) könne sich die Wissenschaft als dasjenige Funktionsystem in der Gesellschaft behaupten, das *exklusiv* für die Erkenntnisgewinnung und Wissensakkumulation zuständig ist (Hansson, 2002). Zwei der insgesamt vier Normen im Ethos der Wissenschaften spielen bei der Analyse des Peer-Review-Verfahrens eine besonders wichtige Rolle: (1) die Norm des Universalismus und (2) die Norm des organisierten Skeptizismus (die Normen des Kommunismus und der Uneigennützigkeit betreffen eher die Gewährleistung des uneingeschränkten Zugangs zu Forschungsergebnissen):

Mit der Wirksamkeit der Norm des Universalismus im Peer-Review soll verhindert werden, dass ökonomische, politische, religiöse oder anderweitige wissenschafts-

ferne Interessen einen Einfluss auf den Prozess der Selektion von wissenschaftlicher Arbeit und somit den Prozess der Erkenntnisgewinnung insgesamt nehmen können (Ziman, 2000). Wenn Wissenschaftler die Arbeit eines Fachkollegen in einem Peer-Review-Verfahren zu begutachten hätten, würden sie bei ihrem Urteil dazu angehalten werden, ausschliesslich wissenschaftliche Gütekriterien und keine sachfremden Kriterien anzulegen (siehe dazu Cole, 1992). Die am besten geeigneten Personen für die Begutachtung nach wissenschaftlichen Gütekriterien seien die arrivierten Fachkollegen des begutachteten Wissenschaftlers (Antragstellers oder Autors, siehe dazu die Ergebnisse von Jayasinghe, 2003, zum Peer-Review). Bei der Bewertung der wissenschaftlichen Güte einer Arbeit würde für die Gutachter die Korrespondenz der Forschungsergebnisse mit den «Gesetzen der Natur» eine massgebliche Rolle spielen: Von den Gutachtern würden nur diejenigen Sachverhalte positiv beurteilt werden, die mit den «Gesetzen der Natur» übereinstimmen (Cole, 1992). Die Norm des Universalismus würde im Peer-Review letztlich dazu führen, dass Erkenntnisse unabhängig von lokalen und sozialen Bedingungen bei der Gewinnung (im Labor) oder Selektion (im Peer-Review) in der Wissenschaft akzeptiert und kommuniziert werden (Ziman, 2000).

Mit der Wirksamkeit der Norm des organisierten Skeptizismus könnte im Peer-Review sichergestellt werden, dass sich neue Erkenntnisse einer kontroversen Diskussion unter Fachkollegen zu stellen haben, bis ihnen bei einer positiven Begutachtung ein anerkanntes Mass an Glaubwürdigkeit und Sicherheit zugeschrieben wird. Durch die Norm des organisierten Skeptizismus würden nur solche Erkenntnisse in den Wissensbestand und die weitere Wissensverwendung eines Fachgebietes eingehen, die durch Fachkollegen eingehend geprüft und legitimiert seien. Damit hätte die Wirksamkeit der Norm des organisierten Skeptizismus letztlich zur Folge, dass ein hohes Mass an Übereinstimmung durch legitimierte und somit «konsensfähige» Erkenntnisse (Ziman, 1982) in einem Fachgebiet hergestellt werden. Die Anhänger der Nordamerikanischen Schule halten ein hohes Mass an Konsens im Prozess der Erkenntnisbildung für eine wichtige Voraussetzung, um die Akkumulation von wahrheitsfähigem Wissen in einem Fachgebiet dauerhaft zu gewährleisten (Mulkay, 1969).

«Normen haben die Gestalt von Vorschriften, Verboten und Grundsätzen, die bestimmen, was bevorzugt werden soll und was noch zulässig ist» (Merton, 1985, 88). Im Peer-Review würden die Vorschriften und Grundsätze für «gutes» wissenschaftliches Handeln mit Hilfe von Sanktionen durchgesetzt werden. Diese Sanktionen seien überwiegend inkrementeller Art (siehe dazu Coleman, 1995): Wissenschaftler, denen absichtsvolles Fehlverhalten nachgewiesen werde, würden (kollektiv) «geschnitten» werden. Zahlreiche Listen mit Verhaltensempfehlungen, die von wissenschaftlichen Institutionen zur Prävention gegen und zum Umgang mit Fehlverhalten publiziert wurden, können die Existenz dieser Sanktionen bestätigen (siehe z. B. European Association for Chemical and Molecular Sciences,

2006). Nach Merton (1985) hätten Angriffe auf die Integrität der Wissenschaft gezeigt, dass die Wissenschaft auf diese Sanktionen angewiesen sei, um den Prozess der Erkenntnisgewinnung nicht zu gefährden.

Ein Teil der Studien aus der Peer-Review-Forschung (siehe Kapitel 2) enthält Hinweise darauf, dass die Annahmen der Nordamerikanischen Schule das Handeln von Wissenschaftlern im Peer-Review angemessen beschreiben können: (1) Die Studien von Hargens und Herting (1990) zeigen, dass Gutachter bei der Manuskriptbegutachtung auf einer einheitlichen wissenschaftlichen Qualitätsdimension urteilen. Dieser Konsens über Qualität kann auf die Wirksamkeit beider Normen (Universalismus und organisierter Skeptizismus) im Peer-Review zurückgeführt werden. (2) Bislang wurde noch keine Untersuchung zur Fairness im Peer-Review durchgeführt, die mit einem experimentellen Studiendesign systematische Urteilsfehler im Peer-Review zweifelsfrei nachweisen konnte. Gutachter, Herausgeber und Programmmanager scheinen sich demnach in ihrem Handeln an der Norm des Universalismus zu orientieren. (3) Die Studien zur prädiktiven Validität verdeutlichen, dass Herausgeber und Gutachter den wissenschaftlichen Impact von Manuskripten korrekt einschätzen können: Abgelehnte und andernorts publizierte Manuskripte werden im Durchschnitt deutlich seltener zitiert als akzeptierte Manuskripte. Der Befund, dass sich die Begutachtung an der wissenschaftlichen Güte (bzw. dem möglichen Impact) einer Arbeit ausrichtet, kann wiederum als ein Hinweis auf die Wirksamkeit von beiden Normen im Peer-Review verstanden werden.

3.2 Die sozialkonstruktivistische Sicht auf das Peer-Review

Die Anhänger des Sozialkonstruktivismus unter den Wissenschaftssoziologen weisen die theoretischen Annahmen der Nordamerikanischen Schule – insbesondere die Existenz der Normen, die Merton (1973) im Ethos der Wissenschaften zusammengefasst hat – zurück und ersetzen sie durch eigene Annahmen (siehe dazu im Überblick Mulkay, 1979). Im Folgenden werden in einem ersten Schritt die wichtigsten Annahmen der Sozialkonstruktivisten beschrieben, und es wird in einem zweiten Schritt ihre Übertragung auf das Peer-Review vorgenommen. Wie wir sehen werden, weisen die theoretischen Annahmen erkennbare Bezüge zu jener Kritik auf, die seit Jahren am traditionellen Peer-Review-Verfahren (dessen Prinzipien durch die Nordamerikanische Schule geprägt sind) geäussert wird (siehe Kapitel 1).

Erste Annahme: Wissenschaftliches Handeln wird nicht durch die Normen, die im Ethos der Wissenschaften zusammengefasst sind, gesteuert

Die Anhänger des Sozialkonstruktivismus sind davon überzeugt, dass Wissenschaft kein normativ gesteuerter Prozess sei. Dabei stützen sie sich auf eine Reihe von mikrosoziologischen Studien (siehe beispielsweise Knorr-Cetina, 1981). Diese Studien würden kaum Hinweise dafür liefern, dass (1) Wissenschaftler sich darum bemühen würden, bei ihrem Handeln bestimmten Normen zu folgen, und dass

(2) das Verhalten von Wissenschaftlern durch das Ethos der Wissenschaften befriedigend erklärt werden könne (Knorr-Cetina, 1991). Wissenschaftliches Handeln wird daher nicht als normkonformes Verhalten angesehen. Aus der Annahme der Unabhängigkeit von wissenschaftlichem Handeln und normativen Vorgaben leiten die Anhänger des Sozialkonstruktivismus die These ab, dass normativ gesteuertes Handeln *keine* notwendige Voraussetzung für den Prozess der Erkenntnisbildung bzw. für den wissenschaftlichen Fortschritt darstellt.

In der Peer-Review-Forschung ist eine Reihe von Studien durchgeführt worden, die das Verhalten der Wissenschaftler im Peer-Review vor dem Hintergrund der Normen des Ethos der Wissenschaften untersucht haben. So haben alle Studien, die sich mit der Fairness im Peer-Review beschäftigt haben (siehe Kapitel 2.2), implizit oder explizit die Wirksamkeit der Norm des Universalismus im Begutachtungsprozess geprüft. Die Ergebnisse einiger dieser Studien deuten darauf hin, dass Urteile im Peer-Review-Verfahren durch nicht wissenschaftliche Kriterien, wie z. B. das Geschlecht, systematisch beeinflusst werden. In einer der bekanntesten Studien, deren Ergebnisse in der Zeitschrift *Nature* publiziert wurden, weisen Wennerås und Wold (1997) für das Peer-Review-Verfahren des Swedish Medical Research Council (MRC, Stockholm) nach, dass eine Frau etwa 2.5-mal wissenschaftlich produktiver sein muss (durch Publikationen in Zeitschriften mit einer hohen Qualität) als ein Mann, um von den Gutachtern das gleiche Urteil zu erhalten.

Während Studien, wie die von Wennerås und Wold (1997) normkonformes Verhalten von Wissenschaftlern im Peer-Review in Frage stellen, wird durch die Ergebnisse einer Studie, deren Ergebnisse auf dem fünften International Congress on Peer-Review and Biomedical Publication (Chicago, USA) präsentiert wurden, darüber hinaus deutlich, dass Wissenschaftler kaum dazu bereit sind, ihr Verhalten an institutionell vorgegebenen Grundsätzen und Vorschriften auszurichten. Ellison, Rosato und Outram (2005) haben untersucht, inwieweit Autoren, die im *British Medical Journal* (BMJ) publiziert haben, die ethischen Richtlinien der Zeitschrift berücksichtigt haben. Die Ergebnisse werden wie folgt zusammenfassend beschrieben: «It is clear that the BMJ's ‹guidelines on the use of ethnic, racial, and cultural descriptions in published research› have not been followed».

Zweite Annahme: Wissenschaftler besitzen kein Privileg für die Teilnahme am wissenschaftlichen Diskurs in ihrer Fachdisziplin

Gemäß den Annahmen der Nordamerikanischen Schule eignen sich Wissenschaftler durch eine langjährige Hochschulausbildung und Forschungstätigkeit spezifische Kenntnisse und Fähigkeiten in einem Fachgebiet an, die sie zu Autoritäten in diesem Gebiet heranreifen lassen (Collins und Evans, 2002). Aufgrund dieser Kenntnisse und Fähigkeiten könnten letztlich *nur* solche Autoritäten den wissenschaftlichen Diskurs in einem Fachgebiet auf einem angemessenen Niveau führen. Die Sozialkonstruktivisten widersprechen dieser Annahme und sprechen den Autoritäten

in einem Fachgebiet das Privileg ab, sich am wissenschaftlichen Diskurs (alleine) beteiligen zu können (siehe dazu Fuchs, 1996). «It might seem that anyone can be an expert» (Collins und Evans, 2002, 238). Auch für Außenstehende (fachfremde Personen bzw. Wissenschaftler), denen zwar ein technisches oder mathematisches Verständnis in einem Fachgebiet fehle, die aber mit Inhalten, Theorien und Fachbegriffen vertraut seien, sei es in ähnlicher Weise wie für Autoritäten möglich, am wissenschaftlichen Diskurs teilzunehmen.

In einem Experiment, das Harry M. Collins – einer der einflussreichsten Wissenschaftler unter den Sozialkonstruktivisten – durchgeführt hat und über dessen Ergebnisse in *Nature* berichtet wurde (Giles, 2006), konnte gezeigt werden, dass die Antworten von einem Experten auf Fragen aus seinem Fachgebiet (Gravitationswellen) kaum von den Antworten eines fachfremden Wissenschaftlers (einem Wissenschaftssoziologen) unterschieden werden konnten: «The experiments show that the linguistic performance of those well socialized in the language of a specialist group is indistinguishable from those with full blown practical socialization but distinguishable from those who are not well socialized» (Collins, Evans et al., 2006, 656). Eine andere Studie, die das Privileg von Autoritäten für die Tätigkeit als Gutachter im Peer-Review Prozess in Frage stellt, wurde auf dem fünften International Congress on Peer-Review and Biomedical Publication (Chicago, USA) präsentiert. Bryan, Fletcher und Kale (2005) haben in ihrer Studie untersucht, inwieweit ein (erfahrener) Mitarbeiter des «editorial administrative team» beim BMJ zu ähnlichen finalen Entscheidungen über eingereichte Manuskripte kommen würde, wie die Herausgeber der Zeitschrift aufgrund der Gutachtervoten in einem bestimmten Zeitraum gekommen sind. Die Ergebnisse zeigen, dass der Mitarbeiter bei 90% der ursprünglich abgelehnten Manuskripte dieselbe Entscheidung getroffen hätte; nur 10% der abgelehnten Arbeiten hätte er zur Publikation angenommen.

Die Ergebnisse von Collins, Evans, Ribeiro und Hall (2006) sowie Bryan, Fletcher und Kale (2005) widersprechen der Annahme der Nordamerikanischen Schule, dass aktive Wissenschaftler aus demselben Forschungsgebiet die am besten geeigneten Personen seien, um entsprechend der Norm des organisierten Skeptizismus die Arbeit ihrer Fachkollegen auf ihre wissenschaftliche Güte hin zu begutachten. Außenstehende (fachfremde Personen bzw. Wissenschaftler), die mit Inhalten, Theorien und Fachbegriffen vertraut seien, könnten zu ähnlichen Urteilen kommen.

Dritte Annahme: Wissenschaftliche Erkenntnisse geben keine «natürliche Realität» wider, sondern werden sozial und lokal konstruiert

Eine der wichtigsten Annahmen der Sozialkonstruktivisten lautet, dass wissenschaftliche Erkenntnisse sozial und lokal konstruiert werden (Knorr-Cetina, 1991). Erkenntnisse würden keine «natürliche Realität» wiedergeben, sondern entstünden im Forschungsprozess unter den gegebenen Bedingungen vor Ort (Knorr-Cetina, 1981; Kuhn, 1962). Zu diesen Bedingungen würden sowohl das kulturelle und

gesellschaftliche Umfeld als auch persönliche Interessen, soziale Bindungen und Urteilsfehler der Wissenschaftler zählen. Um die Prozesse der Konstruktion von Erkenntnissen vor Ort zu beschreiben, werden deshalb von den Sozialkonstruktivisten Begriffe wie Situiertheit, Kontextualität, interpretative Flexibilität, opportunistische Rationalität und lokale Idiosynkrasien herangezogen (Krohn, 2000).

Für die Anhänger des Sozialkonstruktivismus spielen nicht nur im Prozess der Erkenntnisgewinnung, sondern auch bei der Durchführung des Peer-Review die «Regeln der Natur» oder die «Wahrheit» kaum eine Rolle. Günstige und ungünstige Urteile von Gutachtern, Herausgebern und Programmanagern seien (ausschliesslich) das Resultat sozialer Prozesse und lokaler Bedingungen. Die Beteiligten an einem Peer-Review-Verfahren würden vor dem Hintergrund ihrer persönlichen Biographien, der individuellen Interpretation des Fachwissens, der sozialen Netzwerke sowie der lokalen Arbeitsbedingungen begutachten (Gläser und Laudel, 2006).

Für Cole (1992) wird die Annahme der Sozialkonstruktivisten, dass nicht die wissenschaftliche Güte einer Arbeit, sondern lokale und soziale Bedingungen die Begutachtung bestimmen, durch die in einigen Peer-Review-Studien berichtete niedrige Gutachterübereinstimmung belegt (siehe Kapitel 2.1): «If consensus can be determined by comparing a contribution with nature or by the application of a set of rational rules to evaluate the validity of a contribution, why do we find little consensus» (Cole, 1992, 83)? Auch die vielen Urteilsfehler, die im Zusammenhang mit dem Peer-Review diskutiert werden (siehe Kapitel 2.2), lassen Zweifel an einer objektiven Begutachtung ohne lokalen und sozialen Bezug aufkommen.

Die Ergebnisse der Studien zur prädiktiven Validität (siehe Kapitel 2.3) zeigen einerseits, dass akzeptierte Manuskripte einen grösseren Impact auf die spätere Forschung haben als abgelehnte und andernorts publizierte Manuskripte. Das Peer-Review-Verfahren scheint demnach seiner Selektionsfunktion von «erfolgreicher» und «erfolgloser» Forschung nachzukommen. Andererseits konnten die Studien zur prädiktiven Validität nachweisen, dass das Peer-Review-Verfahren bei denselben Manuskripten zu unterschiedlichen Urteilen kommt: Viele Manuskripte, die bei einer Zeitschrift (mit einem Peer-Review-Verfahren) abgelehnt werden, werden bei einer anderen Zeitschrift (mit einem Peer-Review-Verfahren) akzeptiert. Dieser Befund deutet darauf hin, dass der Manuskriptbegutachtung *keine* allgemein gültigen Qualitätskriterien zugrunde gelegt werden, die eine wissenschaftliche Arbeit erfüllen (Akzeptanz) oder nicht erfüllen (Ablehnung) kann; die Begutachtung (bzw. das Ergebnis der Begutachtung) scheint stattdessen von den lokalen und sozialen Bedingungen abhängig zu sein, unter denen das Peer-Review bei den einzelnen Zeitschriften stattfindet.

Vierte Annahme: Eine wissenschaftliche Arbeit ist ein soziales Konstrukt der begutachteten und begutachtenden Wissenschaftler

Für die Anhänger der Nordamerikanischen Schule ist ein Manuscript das Resultat der Forschungsarbeit eines Wissenschaftlers, der sich mit dem Einreichen des Manuscriptes bei einer Zeitschrift die Eigentumsrechte mit dem alleinigen Recht auf Anerkennung zusichern möchte (Merton, 1985). Die Anhänger des Sozialkonstruktivismus widersprechen dieser Annahme. Nicht nur der Autor sei für den Inhalt eines Manuscriptes verantwortlich (und könne dafür Eigentumsrechte geltend machen); der Inhalt eines Manuscriptes sei ein gemeinsames Produkt von Autor, Gutachter und Herausgeber (diese Sicht wurde auch auf Forschungsanträge übertragen, siehe Laudel, 2006). Autoren würden beim Schreiben den Begutachtungsprozess einer Zeitschrift antizipieren (Knorr-Cetina, 2002); Gutachter und Herausgeber würden im Begutachtungsprozess eigene Interessen im Hinblick auf den Inhalt des Manuscriptes zu verwirklichen suchen (Gläser, 2006; Gläser und Laudel, 2006). Die theoretische Annahme, dass ein Manuscript von den begutachteten *und* den begutachtenden Wissenschaftlern in einem Peer-Review-Verfahren sozial konstruiert wird, ist von den Anhängern des Sozialkonstruktivismus aus Beobachtungen des Forschungs- und Publikationsprozesses (Knorr-Cetina, 1981) sowie aus Analysen von Manuscripten und Gutachten (Myers, 1990) entwickelt worden.

3.3 Das Peer-Review aus der Sicht der Theorie sozialer Systeme

Die Vertreter der Nordamerikanischen Schule und diejenigen des Sozialkonstruktivismus stehen sich seit jeher mit einer tendenziell ablehnenden Haltung gegenüber (siehe dazu z. B. Knorr-Cetina, 1991; Sokal und Bricmont, 1999), obwohl einige Arbeiten publiziert wurden, die eine vermittelnde Position vorgeschlagen haben (z. B. Murphy, 1994). Krohn (2000) schätzt das Potenzial der Theorie sozialer Systeme von Luhmann (1992) «zur Überwindung der Krise der social studies of science» (320), die durch die theoretische Fragmentierung entstanden sei, als besonders hoch ein. Mit der Berücksichtigung einer Beobachterperspektive hat Luhmann (1992) die Differenz, die die Konstruktivisten zwischen Erkenntnis und Gegenstand vornehmen, beibehalten. Für ihn ist Erkenntnis die Konstruktion eines Unterschieds, wobei das in der Unterscheidung Bezeichnete in der Realität keine direkte Entsprechung hat. Im Unterschied zum sozialkonstruktivistischen Ansatz (und in Anlehnung an die Nordamerikanische Schule) ist für Luhmann (1992) die Wissenschaft durch eine spezifisch ausgebildete Kommunikationsform gekennzeichnet, die sich an «Wahrheit» orientiert. Durch diese Orientierung kann sich die Wissenschaft in der Gesellschaft funktional ausdifferenzieren und als autonomes System erhalten.

Luhmann (1992) hat für seine Analyse der Wissenschaft zwei Komponenten aus der Theorie sozialer Systeme herangezogen, die bei der Erzeugung gesicherter Erkenntnisse eine grundlegende Bedeutung haben sollen: 1) die Beobachtung zweiter

Ordnung sowie 2) die Evolution von Erkenntnissen. Im Unterschied zu den theoretischen Annahmen der Nordamerikanischen Schule und der Sozialkonstruktivisten handelt es sich bei diesen Komponenten *nicht* um Mechanismen, die (bestimmte) Phänomene in der Wissenschaft durch das Handeln von Personen im sozialen Kontext zu erklären versuchen (siehe dazu Hedström, 2005). Stattdessen können diese Komponenten den Ablauf von wissenschaftlicher Kommunikation grundsätzlich beschreiben und in diesem Ablauf die zentrale Bedeutung von Begutachtung und Selektion herausstellen.

Nach der Theorie sozialer Systeme hat sich die Wissenschaft in der Gesellschaft als ein autopoietisches System mit der Funktion ausdifferenziert, Erkenntnisse mit einer hohen Verbindlichkeit bzw. Sicherheit zu erzeugen (Luhmann, 1992). Autopoietische Systeme zeichnen sich durch eine rekursive, operationale Geschlossenheit ihrer Elemente gegenüber der Umwelt aus (Maturana, 1985). Das operative Element der Wissenschaft sei jene Kommunikation von Erkenntnissen, die sich im symbolisch generalisierten Medium «Wahrheit» reproduziere (Stichweh, 1994). Kommunikation sei nur dann als spezifisch wissenschaftlich zu bezeichnen, wenn sie sich an der Differenz (dem binären Code) von «wahr» und «nicht wahr» orientiere. Mit der Zuschreibung von «wahr» werde im System symbolisiert, dass eine Erkenntnis als eine verallgemeinerbare Wirklichkeitsaussage gelten könne.

Die Beobachtung zweiter Ordnung

Luhmann (1992) gründet die Theorie sozialer Systeme auf einen formalen Begriff des Beobachtens, der auf Spencer Brown (1969) zurückgeht. Spencer Brown (1969) hat die sogenannten «Gesetze der Form» konzipiert, bei denen es ihm mit Hilfe von zwei Begriffen (Bezeichnung und Unterscheidung) gelingen soll, alle möglichen Operationen eines Beobachters zu beschreiben: Grundlage einer jeden Bezeichnung sei eine Unterscheidung zwischen innen und aussen, die ein Beobachter vornehme. In diesem Vorgang stehe dem Beobachter kein vollständiges und konsistentes Beschreibungssystem zur Verfügung, d. h. «no observation can observe what it cannot observe» (Fuchs, 1996, 321). Der Beobachter treffe Unterscheidungen mit einem «blinden Fleck», der erst durch eine Beobachtung höherer Ordnung erkennbar werden könne. Da jede Beobachtung mit einem eigenen «blinden Fleck» behaftet sei, seien mit jeder weiteren Beobachtung höherer Ordnung keine «besseren» Realitätsbeschreibungen zu erwarten als mit der Beobachtung einer niedrigeren Ordnung.

Luhmann (1992) hat den formalen Begriff des Beobachtens auf den Prozess der Erkenntnisbildung in der Wissenschaft übertragen. Weil «the natural world cannot speak for itself» (Gieryn, 1982, 288), sei man im Forschungsprozess darauf angewiesen, durch eine Beobachtung erster Ordnung Erkenntnisse mit dem Anspruch zu generieren, ein bestimmtes Naturphänomen (bzw. Sozialphänomen) zu erklären. Da diese Beobachtung ausschliesslich auf das Phänomen und nicht auf die Beobachtung selber bezogen sein könne, könne sie nur als «naiv realistisch» und

nicht als «konstruktivistisch» bezeichnet werden (Fuchs, 1996): Der Beobachter erster Ordnung könnte auf die Annahme von Wirklichkeit nicht verzichten; er treffe Unterscheidungen in einer Objektwelt (Schmidt, 1998). Wenn die Beobachtung als Erkenntnis formuliert werde, könne sie durch eine Beobachtung zweiter Ordnung als «wahr» (gesichert) oder «nicht wahr» (ungesichert) kodiert werden.

«On the second level, communication must decide, amongst other issues, which observations are relevant and deserve being noticed by others» (Fuchs, 1996, 317). Erkenntnisse (aus der Beobachtung erster Ordnung) würden in der Regel in Manuskripten beschrieben werden, die bei wissenschaftlichen Zeitschriften zur Begutachtung eingereicht werden. Fachkollegen würden die Erkenntnisse durch eine Beobachtung zweiter Ordnung (das kritische Lesen) als «wahr» oder «nicht wahr» (bzw. gesichert oder nicht gesichert) kodieren. Da die Zuschreibung von «wahr» und «nicht wahr» nur durch eine Beobachtung zweiter Ordnung vorgenommen werden könne (und *nicht* durch eine Beobachtung erster Ordnung in einer Objektwelt), kann demnach erst durch die kritische Begutachtung gewährleistet werden, dass jene *gesicherten* Erkenntnisse entstehen, die die Gesellschaft als funktionsspezifische Leistung von der Wissenschaft erwarte. «Erst auf der Ebene der Beobachtung zweiter Ordnung kann [...] der Differenzcode wahr/unwahr voll zum Zuge kommen; erst auf dieser Ebene kann [...] Wissenschaft als System ausdifferenziert werden» (Luhmann, 1992, 170).

Die Evolution von Erkenntnissen

Luhmann (1992) hat für die Analyse der Wissenschaft eine Übertragung der Evolutionstheorie auf den Prozess der Generierung von (gesicherten) Erkenntnissen vorgenommen. Dafür hat er die Hauptelemente der Evolutionstheorie Variation, Selektion und Stabilisierung herangezogen. Erkenntnisse würden in der Wissenschaft in einer unüberschaubaren Menge mit einem hohen Mass an *Variation* produziert werden. Wir hätten es mit einem globalen Angebot jener Erkenntnisse zu tun, die zur Erklärung von Natur- und Sozialphänomenen beitragen (können). Die hohe Variationsbreite der Erkenntnisse habe dazu geführt, dass Erkenntnisse in der Wissenschaft über das Mass der Rezeptionsfähigkeit hinaus produziert würden. Im Prozess der *Selektion* werde deshalb ein Teil der Erkenntnisse kommunikativ aufgegriffen und als «wahr» markiert (der andere Teil werde nicht aufgegriffen bzw. als «nicht wahr» markiert). In der Wissenschaft wird der (kritische) Selektionsprozess (unter anderem) durch das Peer-Review-Verfahren vollzogen. Ein anderer, dem Peer-Review nachgeordneter Selektionsprozess ist die Verwendung (bzw. das Ignorieren) von Ergebnissen und Gedankengut in der weiteren Forschungsarbeit in einem Fachgebiet, die ihren Ausdruck in der Zitierung durch Fachkollegen findet. Durch die Ablehnung von Manuskripten, Forschungsanträgen etc. im Peer-Review wird nur ein Teil der Erkenntnisse der Möglichkeit ausgesetzt, in den Wissensbestand bzw. die weitere Wissensverwendung in einem Fachgebiet aufgenommen zu werden. Der

andere Teil kann die «Hürde» des Peer-Review nicht überwinden und wird damit dem Schicksal des Vergessens ausgesetzt. Mit der Selektion von wissenschaftlicher Arbeit kann das Peer-Review gleichzeitig die Funktion der Stabilisierung (von Wissen) in der Wissenschaft erfüllen. Im evolutionären Prozess bezeichnet – nach Luhmann (1992) – die *Stabilisierung* die Tendenz im System, Veränderungen im Wissensbestand weitgehend zu vermeiden und beim bisherigen Wissen zu bleiben bzw. im Zustand der bisherigen Struktur zu verharren. Letztlich könne sich entgegen der stabilisierenden Tendenzen nur jene «neue» Erkenntnis im System als gesicherte Erkenntnis durchsetzen, die einerseits hinreichend auffalle und sich andererseits intersubjektiv als «wahr» herausstelle.

4 Diskussion

In den vorangegangenen Kapiteln ist das Peer-Review aus der Perspektive der drei wichtigsten theoretischen Richtungen in der Wissenschaftssoziologie analysiert worden: (1) der Nordamerikanischen Schule, (2) des Sozialkonstruktivismus und (3) der Theorie sozialer Systeme. Angesichts der inkonsistenten Befunde in der Peer-Review-Forschung (siehe Kapitel 2) liessen sich bei der Analyse des Peer-Review aus der Sicht der Nordamerikanischen Schule und des Sozialkonstruktivismus *jeweils* empirische Ergebnisse (aus Studien mit einer strengen Methodik) heranziehen, die die «diametralen» theoretischen Annahmen stützen konnten. Die inkonsistenten Befunde in der Peer-Review-Forschung spiegeln demnach die theoretische Fragmentierung in der Wissenschaftssoziologie wider (und umgekehrt). Die empirische Validierung der Annahmen von *beiden* theoretischen Richtungen (wie sie in den vorangegangenen Kapiteln aufgezeigt wurde), kann letztlich jene wissenschaftssoziologischen Arbeiten stützen, die eine vermittelnde Position zwischen beiden Theorierichtungen vorgeschlagen haben (siehe z. B. Murphy, 1994). Da es sich bei den Komponenten aus der Theorie sozialer Systeme (Beobachtung zweiter Ordnung und Evolution von Erkenntnissen), die in diesem Beitrag für die Analyse des Peer-Review verwendet wurden, um «abstrakte» Kommunikationsprozesse (ohne direkten Bezug zur Handlungsebene von Wissenschaftlern) handelt, liessen sich empirische Ergebnisse aus der Peer-Review-Forschung für die Validierung dieser Prozesse nicht heranziehen.

Mit der Analyse des Peer-Review aus der Perspektive der drei Theorierichtungen hat sich ein theoretischer Rahmen für die empirische Untersuchung des Peer-Review ergeben, der abschliessend zusammenfassend beschrieben werden soll: Für die Beschreibung grundlegender Prozesse im Peer-Review können die Beobachtung zweiter Ordnung und die Evolution von Erkenntnissen herangezogen werden. Im Ablauf von wissenschaftlicher Kommunikation können diese Komponenten angeben, auf welche Weise Begutachtung im und Selektion durch das Peer-Review

dazu beitragen, dass die Wissenschaft ihre spezifische Funktion – die Erzeugung gesicherter Erkenntnisse – für die Gesellschaft fortlaufend erfüllen kann: Gesicherte Erkenntnisse werden (nur) *dann* erzeugt, wenn 1) Beobachtungen zweiter Ordnung von (ungesicherten) Erkenntnissen die Zuschreibung von «wahr» und «nicht wahr» vornehmen (Begutachtung im Peer-Review) sowie 2) unter der Bedingung begrenzter Ressourcen ein Teil der Erkenntnisse aus einem variationsreichen Angebot ausgewählt werden (Selektion durch das Peer-Review).

Für die Analyse des Peer-Review sind neben der Beobachtung zweiter Ordnung und der Evolution von Erkenntnissen (auf der Ebene der Kommunikation in der Wissenschaft) (1) die Normen des Ethos der Wissenschaften und (2) die Annahmen der Anhänger des Sozialkonstruktivismus (auf der Ebene des Handelns von Wissenschaftlern) verwendet worden. (1) Die Normen des Universalismus und des organisierten Skeptizismus gelten als Imperative für die Disziplinierung des Verhaltens von Wissenschaftlern. Diese Disziplinierung kann (bzw. soll) jenem System von Begutachtung und Selektion Stabilität verleihen, das durch die Beobachtung zweiter Ordnung und die Evolution von Erkenntnissen fortlaufend generiert wird. Normen können (bzw. sollen) zu einer «Steuerung» des Verhaltens von Wissenschaftlern nach den Grundsätzen für «gutes» wissenschaftliches Handeln führen, indem sie Willkürlichkeiten entgegenwirken (Wirksamkeit des Universalismus) und die Intersubjektivität der Erkenntnisbildung garantieren (Wirksamkeit des organisierten Skeptizismus). (2) Die theoretischen Annahmen der Anhänger des Sozialkonstruktivismus weisen auf der einen Seite auf die Notwendigkeit hin, das Verhalten von Wissenschaftlern im Peer-Review zu disziplinieren. Normkonformes Verhalten ist – so zeigen die Ergebnisse der mikrosoziologischen Studien und ein Teil der Studien aus der Peer-Review-Forschung – keine Selbstverständlichkeit; mit normdivergentem Handeln muss gerechnet werden. Auf der anderen Seite hat die mikrosoziologische Forschung der Sozialkonstruktivisten deutlich gemacht, dass die lokalen und sozialen Bedingungen, unter denen die Wissenschaftler arbeiten, massgeblich den Begutachtungsprozess beeinflussen können. «*Facts* needed long construction and acceptance procedures» (Evetts, Mieg und Felt, 2006, 116).

Mit den Hinweisen auf das normdivergente Handeln von Wissenschaftlern sowie auf die lokalen und sozialen Bedingungen bei der Begutachtung können wir die theoretischen Annahmen der Sozialkonstruktivisten auf der einen Seite als Ergänzung zu den Annahmen der Nordamerikanischen Schule ansehen, die zu einer «dichten» (d. h. umfassenden) Beschreibung des Handelns von Wissenschaftlern im Peer-Review führen (siehe dazu Collins und Evans, 2002; Gilbert und Mulkay, 1984). Auf der anderen Seite zeigen die Annahmen der Sozialkonstruktivisten auf, dass bestimmte Bedingungen, die im traditionellen Peer-Review als notwendig für einen «erfolgreichen» Begutachtungsprozess angesehen werden, der Flexibilität und Variabilität unterliegen: (1) Begutachtungen müssen nicht notwendigerweise von Fachkollegen (Experten) vorgenommen werden. (2) Nicht nur der begutach-

tete Wissenschaftler (Autor oder Antragsteller), sondern auch die begutachtenden Wissenschaftler (Gutachter, Herausgeber oder Programmmanager) in einem Peer-Review-Verfahren tragen dazu bei, dass eine wissenschaftliche Arbeit entsteht. Anerkennung für den Inhalt einer wissenschaftlichen Arbeit gebührt also nicht nur dem Autor oder Antragsteller, sondern auch dem Gutachter und dem Herausgeber bzw. Programmmanager.

Insgesamt gesehen kommen wir mit den Annahmen der Sozialkonstruktivisten, der Nordamerikanischen Schule und der Theorie sozialer Systeme zu einer Beschreibung von Prozessen der Kommunikation und Mechanismen des Handelns, die für die Durchführung des Peer-Review in einer Wissenschaft, die sich für die Erzeugung von gesicherten Erkenntnissen in der Gesellschaft funktional ausdifferenziert hat, als grundlegend angesehen werden können. Die Annahmen der drei wichtigsten Theorierichtungen in der Wissenschaftssoziologie stellen eine wertvolle Grundlage für die vertiefte, theoretisch fundierte Auseinandersetzung mit den (zukünftigen) Ergebnissen der Peer-Review-Forschung dar.

5 Literaturverzeichnis

- Bailar, J. C. (1991), Reliability, fairness, objectivity, and other inappropriate goals in peer review, *Behavioral and Brain Sciences*, 14, 1, 137–138.
- Bornmann, L. (zur Publikation angenommen), Grundlegende Kommunikationsprozesse und stabilisierte Mechanismen für die Erzeugung gesicherter Erkenntnisse – eine Übersicht über den Theoriestand in der Wissenschaftssoziologie, *Sociologica Internationalis*.
- Bornmann, L. und H.-D. Daniel (2003), Begutachtung durch Fachkollegen in der Wissenschaft. Stand der Forschung zur Reliabilität, Fairness und Validität des Peer-Review-Verfahrens, in: S. Schwarz und U. Teichler (Hrsg.), *Universität auf dem Prüfstand. Konzepte und Befunde der Hochschulforschung*, Frankfurt am Main: Campus, 211–230.
- Bornmann, L. und H.-D. Daniel (2005), Selection of research fellowship recipients by committee peer review. Analysis of reliability, fairness and predictive validity of Board of Trustees' decisions, *Scientometrics*, 63, 2, 297–320.
- Bornmann, L.; S. Mittag und H.-D. Daniel (2006), Quality assurance in higher education – meta-evaluation of multi-stage evaluation procedures in Germany, *Higher Education*, 52, 4, 687–709.
- Brouns, M. (2000), The gendered nature of assessment procedures in scientific research funding: the Dutch case, *Higher Education in Europe*, 25, 193–199.
- Bryan, G.; J. Fletcher und R. Kale (2005, September 17, 2005), *Do you need to be an editor to accept or reject research papers sent to the BMJ?*, Paper präsentiert auf: 5th International Congress on Peer Review and Biomedical Publication, Chicago, IL, USA.
- Campanario, J. M. (1998a), Peer review for journals as it stands today – part 1, *Science Communication*, 19, 3, 181–211.
- Campanario, J. M. (1998b), Peer review for journals as it stands today – part 2., *Science Communication*, 19, 4, 277–306.
- Cicchetti, D. V. (1991), The reliability of peer review for manuscript and grant submissions: a cross-disciplinary investigation, *Behavioral and Brain Sciences*, 14, 1, 119–135.

- Cole, J. R. (2000), The role of journals in the growth of scientific knowledge, in: B. Cronin und H. B. Atkins (Hrsg.), *The web of knowledge. A festschrift in honor of Eugene Garfield*, Medford: Information Today, 109–142.
- Cole, S. (1992), *Making science. Between nature and society*, Cambridge: Harvard University Press.
- Cole, S.; J.R. Cole und G.A. Simon (1981), Chance and consensus in peer-review, *Science*, 214, 4523, 881–886.
- Cole, S. und L. Rubin (1978), *Peer Review in the National Science Foundation. Phase one of a study*, Washington DC: National Academic Press.
- Coleman, J. S. (1995), *Grundlagen der Sozialtheorie. Handlungen und Handlungssysteme* (Bd. 1), München: R. Oldenbourg Verlag.
- Collins, H.; R. Evans, R. Ribeiro und M. Hall (2006), Experiments with interactional expertise, *Studies in History and Philosophy of Science Part A*, 37, 4, 656–674.
- Collins, H. M. und R. Evans (2002), The third wave of science studies: studies of expertise and experience, *Social Studies of Science*, 32, 2, 235–296.
- Daniel, H.-D. (1993, 2004), *Guardians of science. Fairness and reliability of peer review*, Weinheim: Wiley-VCH.
- Daniel, H.-D. (2005), Publications as a measure of scientific advancement and of scientists' productivity, *Learned Publishing*, 18, 143–148.
- Demicheli, V. und C. Pietrantonj (2004), Peer review for improving the quality of grant applications (Cochrane Methodology Review), in: *The Cochrane Library, Issue 1*, Chichester: John Wiley & Sons, Ltd.
- Eckberg, D. L. (1991), When nonreliability of reviews indicates solid science, *Behavioral and Brain Sciences*, 14, 1, 145–146.
- Eisenhart, M. (2002), The paradox of peer review: admitting too much or allowing too little?, *Research in Science Education*, 32, 2, 241–255.
- Ellison, G. T. H.; M. Rosato und S. Outram (2005, September 17, 2005), *The impact of editorial guidelines on the classification of race/ethnicity in the BMJ*, Paper präsentiert auf: 5th International Congress on Peer Review and Biomedical Publication, Chicago, IL, USA.
- European Association for Chemical and Molecular Sciences (2006), *Ethical guidelines for publication in journals and reviews*, Brussels: European Association for Chemical and Molecular Sciences (EuCheMS).
- Evetts, J.; H. A. Mieg und U. Felt (2006), Professionalization, scientific expertise, and elitism: a socio-logical perspective, in: K. A. Ericsson, N. Charness, P. J. Feltovich und R. R. Hoffman (Hrsg.), *The Cambridge Handbook of Expertise and Expert Performance*, New York: Cambridge University Press, 105–123.
- Eysenck, H. J. und S.B.G. Eysenck (1992), Peer review: advice to referees and contributors. *Personality and Individual Differences*, 13, 4, 393–399.
- Fuchs, S. (1996), The new wars of truth: conflicts over science studies as differential modes of observation, *Social Science Information Sur Les Sciences Sociales*, 35, 2, 307–326.
- Gieryn, T. F. (1982), Relativist/ constructivist programs in the sociology of science – redundancy and retreat, *Social Studies of Science*, 12, 2, 279–297.
- Gilbert, G. N. und M. Mulkay (1984), *Opening Pandora's box: a sociological analysis of scientists' discourse*, Cambridge: Cambridge University Press.
- Giles, J. (2006), Sociologist fools physics judges, *Nature*, 442, 7098, 8–8.
- Gläser, J. (2006), *Wissenschaftliche Produktionsgemeinschaften. Die soziale Ordnung der Forschung*, Frankfurt am Main: Campus.

- Gläser, J. und G. Laudel (2006), Advantages and dangers of «remote» peer evaluation, *Research Evaluation*, 14, 3, 186–198.
- Goodman, L. A. (1984), *The analysis of cross-classified data having ordered categories*, Cambridge: Harvard University Press.
- Guston, D. H. (2003), The expanding role of peer review processes in the United States, in: P. Shapira und S. Kuhlmann (Hrsg.), *Learning from science and technology policy evaluation. Experiences from the United States and Europe*, Cheltenham: Edward Elgar, 81–97.
- Hackett, E. J. und D.E. Chubin (2003), *Peer review for the 21st century: applications to education research*, Paper präsentiert auf: Peer review of education research grant applications. Implications, considerations, and future directions, Washington, DC, USA.
- Hansson, F. (2002), How to evaluate and select new scientific knowledge? Taking the social dimension seriously in the evaluation of research quality, *VEST*, 15, 2–3, 27–52.
- Hargens, L. L. und J.R. Herting (1990), A new approach to referees assessments of manuscripts, *Social Science Research*, 19, 1, 1–16.
- Hedström, P. (2005), *Dissecting the social: on the principles of analytical sociology*, Cambridge: Cambridge University Press.
- Heimlin, S. (1996), Research on research evaluations, *Social Epistemology*, 10, 2, 209–250.
- Heimlin, S. und S.B. Rasmussen (2006), The shift in academic quality control, *Science Technology & Human Values*, 31, 2, 173–198.
- Hirschauer, S. (2004), Peer Review Verfahren auf dem Prüfstand. Zum Soziologiedefizit der Wissenschaftsevaluation, *Zeitschrift für Soziologie*, 33, 1, 62–83.
- Hojab, M.; J.S. Gonnella und A.S. Caelleigh (2003), Impartial judgment by the «gatekeepers» of science: fallibility and accountability in the peer review process, *Advances in Health Sciences Education*, 8, 1, 75–96.
- Jayasinghe, U. W. (2003), *Peer review in the assessment and funding of research by the Australian Research Council*, University of Western Sydney, Greater Western Sydney, Australia.
- Jayasinghe, U. W.; H.W. Marsh und N. Bond (2001), Peer review in the funding of research in higher education: the Australian experience, *Educational Evaluation and Policy Analysis*, 23, 4, 343–346.
- Knorr-Cetina, K. (1981), *The manufacture of knowledge: an essay on the constructivist and contextual nature of science*, Oxford: Pergamon Press.
- Knorr-Cetina, K. (1991), Merton sociology of science: the first and the last sociology of science, *Contemporary Sociology*, 20, 4, 522–526.
- Knorr-Cetina, K. (2002), *Die Fabrikation von Erkenntnis. Zur Anthropologie der Naturwissenschaft* (2. Auflage), Frankfurt am Main: Suhrkamp.
- Kostoff, R. N. (1995), Federal, research impact assessment – axioms, approaches, applications, *Scientometrics*, 34, 2, 163–206.
- Krämer, G. und L. Montada (2002), *Wissenschaftsforschung in der Psychologie*, Göttingen: Hogrefe.
- Krohn, W. (2000), Wissenschaftssoziologie: zwischen Modernisierungstheorie und Sozialkonstruktivismus auf schwankendem epistemischem Boden, *Soziologische Revue*, 23, 314–325.
- Kuhn, T. S. (1962), *The structure of scientific revolutions*, Chicago: University of Chicago Press.
- Latour, B. und S. Woolgar (1979), *Laboratory life: the social construction of scientific facts*, London: Sage.
- Laudel, G. (2006), Conclave in the Tower of Babel: how peers review interdisciplinary research proposals, *Research Evaluation*, 15, 1, 57–68.
- Lock, S. (1985), *A difficult balance: editorial peer review in medicine*, Philadelphia: ISI Press.
- Luhmann, N. (1992), *Die Wissenschaft der Gesellschaft*, Frankfurt am Main: Suhrkamp.

- Luhmann, N. (1998). *Die Gesellschaft der Gesellschaft*, Frankfurt am Main: Suhrkamp.
- Marsh, H. W. und S. Ball (1991), Reflections on the peer review process, *Behavioral and Brain Sciences*, 14, 1, 157–157.
- Maturana, H. R. (1985), *Erkennen: die Organisation und Verkörperung von Wirklichkeit: ausgewählte Arbeiten zur biologischen Epistemologie* (2. Auflage), Braunschweig: Vieweg.
- Merton, R. K. (1938), *Science, technology and society in 17th century England*, New York: Harper & Row.
- Merton, R. K. (1985), *Entwicklung und Wandel von Forschungsinteressen. Aufsätze zur Wissenschaftssoziologie*, Frankfurt am Main: Suhrkamp.
- Merton, R. K. (Hrsg.) (1973), *The sociology of science: theoretical and empirical investigations*, Chicago: University of Chicago Press.
- Mulkay, M. (1969), Some aspects of cultural growth in natural sciences, *Social Research*, 36, 1, 22–52.
- Mulkay, M. (1979), *Science and the sociology of knowledge*, London: George Allen and Unwin.
- Murphy, R. (1994), The sociological construction of science without nature, *Sociology – the Journal of the British Sociological Association*, 28, 4, 957–974.
- Myers, G. (1990), *Writing biology: texts in the social construction of scientific knowledge*, Madison: The University of Wisconsin Press.
- Ophof, T.; F. Furstner, M. van Geer und R. Coronel (2000), Regrets or no regrets? No regrets! The fate of rejected manuscripts, *Cardiovascular Research*, 45, 1, 255–258.
- Overbeke, J. und E. Wager (2003), The state of the evidence: what we know and what we don't know about journal peer review, in: F. Godlee und T. Jefferson (Hrsg.), *Peer review in health sciences* (2. Auflage., 45–61), London: BMJ Books.
- Owen, R. (1982), Reader bias, *Journal of the American Medical Association*, 247, 18, 2533–2534.
- Pruthi, S.; A. Jain, A. Wahid, K. Mehra und S.A. Nabi (1997), Scientific community and peer review system – a case study of a central government funding scheme in India, *Journal of Scientific & Industrial Research*, 56, 7, 398–407.
- Ross, P. F. (1980), *The sciences' self-management: manuscript refereeing, peer review, and goals in science*, Massachusetts: The Ross Company, Todd Pond.
- Roy, R. (1985), Funding science – the real defects of peer-review and an alternative to it, *Science Technology & Human Values*, 52, 73–81.
- Sandström, U. und M. Hällsten (2006, 7–9 September), *Productivism in grant peer review: Wennerås and Wold revisited*, Paper präsentiert auf: Ninth International Conference on Science and Technology Indicators, Leuven, Belgium.
- Schmidt, S. J. (1998), *Die Zähmung des Blicks. Konstruktivismus – Empirie – Wissenschaft*, Frankfurt am Main: Suhrkamp.
- Sharp, D. W. (1990), What can and should be done to reduce publication bias – the perspective of an editor, *Journal of the American Medical Association*, 263, 10, 1390–1391.
- Shatz, D. (2004), *Peer review: a critical inquiry*, Lanham: Rowman & Littlefield.
- Smith, R. (2006), Peer review: a flawed process at the heart of science and journals, *Journal of the Royal Society of Medicine*, 99, 4, 178–182.
- Sokal, A. und J. Bricmont (1999), *Eleganter Unsinn: wie die Denker der Postmoderne die Wissenschaften missbrauchen*, München: Beck.
- Spencer-Brown, G. (1969), *Laws of Form*, New York: Dutton.
- Stichweh, R. (1994), *Wissenschaft, Universität, Professionen: soziologische Analysen*, Frankfurt am Main: Suhrkamp.

- Stricker, L. J. (1991), Disagreement among journal reviewers: no cause for undue alarm, *Behavioral and Brain Sciences*, 14, 1, 163–164.
- van Raan, A. F. J. (2004), Measuring science. Capita selecta of current main issues, in: H. F. Moed, W. Glänzel und U. Schmoch (Hrsg.), *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems* (19–50), Dordrecht: Kluwer Academic Publishers.
- Ward, J. E. und Donnelly, N. (1998), Is there gender bias in research fellowships awarded by the NHMRC? *Medical Journal of Australia*, 169, 11–12, 623–624.
- Weller, A. C. (2002), *Editorial peer review: its strengths and weaknesses*, Medford: Information Today, Inc.
- Wennerås, C. und A. Wold (1997), Nepotism and sexism in peer-review, *Nature*, 387, 6631, 341–343.
- Wessely, S. (1998), Peer review of grant applications: what do we know?, *Lancet*, 352, 9124, 301–305.
- Wilson, J. D. (1978), Peer review and publication, *Journal of Clinical Investigation*, 61, 4, 1697–1701.
- Wissenschaftliche Kommission Niedersachsen (2006), *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen. Bewertung des Evaluationsverfahrens*, Hannover: Wissenschaftliche Kommission Niedersachsen.
- Wood, F. Q. und S. Wessely (2003), Peer review of grant applications: a systematic review, in: F. Godlee und T. Jefferson (Hrsg.), *Peer review in health sciences* (2. Auflage., 14–44), London: BMJ Books.
- Young, S. N. (2003), Peer review of manuscripts: theory and practice, *Journal of Psychiatry & Neuroscience*, 28, 5, 327–330.
- Ziman, J. (1982), *Wie zuverlässig ist wissenschaftliche Erkenntnis?*, Braunschweig: Friedrich Vieweg & Sohn.
- Ziman, J. (2000), *Real science. What it is, and what it means*, Cambridge: Cambridge University Press.