

„Informed Peer Review“ – Ausgleich der Fehler und Verzerrungen?

Gerhard Fröhlich

1. Wozu Evaluation? Kritik als ein Definiens von Wissenschaft

Die Evaluation wissenschaftlicher Leistungen ist zurzeit Gegenstand vieler Debatten. Doch stützen sich diese eher auf private bzw. ideologische Meinungen, als auf die Erkenntnisse der Wissenschaftstheorie, Wissenschaftsethik und Wissenschaftsforschung.

Wissenschaftstheoretisch gesehen ist Evaluation (im Sinne von Kritik, Bewertung von Theorien, Methoden, Befunden nach ‚wissenschafts-internen‘ Kriterien) von vornherein konstitutiv für alle wissenschaftliche Methoden. Was viele WissenschaftlerInnen in letzter Zeit empört, ist die Festlegung der Parameter von außen, etwa geleitet von Prinzipien der Marktforschung oder des Controllings. Durch einen Verzicht auf normative, d.h. wissenschaftstheoretische und wissenschaftsethische Festlegungen guter wissenschaftlicher Praxis, auf die Nutzung der Befunde der Wissenschaftsforschung etwa zu Innovationshemmnissen dient Evaluation weniger dem Erkenntnisfortschritt, als eher der Legitimation von Verteilungskämpfen und Sparmaßnahmen.

Mein wissenschaftstheoretischer Ausgangspunkt ist der öffentliche Charakter wissenschaftlicher Methoden. Kritik und kognitive Konkurrenz (Theorien- und Methodenpluralismus) sind Definitionskriterien wissenschaftlicher Rationalität (Karl Popper). Rücksichtslose (Popper) Kritik und offene Kommunikation sind Definiencia von Wissenschaft: „Die Wissenschaft, und insbesondere der wissenschaftliche Fortschritt, ist nicht das Ergebnis isolierter Leistungen, sondern der freien Konkurrenz des Denkens. Denn die Wissenschaft braucht immer mehr Konkurrenz zwischen Hypothesen und immer rigorosere Prüfungen. Und die konkurrierenden Hypothesen müssen durch Personen vertreten werden: sie brauchen Anwälte, Geschworene und sogar ein Publikum. Diese persönliche Vertretung muss institutionell organisiert werden, wenn sie verlässlich funktionieren

soll. Und diese Institutionen müssen unterhalten und gesetzlich geschützt werden.“ (Popper 1969, 120f.)

2. Nicht-intendierte Effekte: Verteidigungsstrategien Betroffener

Wie jedes soziale Handeln haben normative Forderungen, Organisationsformen oder evaluative Praktiken nicht-intendierte Effekte. Vor allem entwickeln die Betroffenen Ausweich- und Verteidigungsstrategien.

Kritik zu üben, kann gefährlich sein – schon bei der nächsten Begutachtung eines Projekts, einer Publikation, einer Bewerbung könnte die Rache kommen (im Schutz der Anonymität). So dominiert in der wissenschaftlichen Öffentlichkeit laue Höflichkeit: Komplimente zu verteilen ist risikoarm. Brisante Informationen erhält man über informelle Kanäle, Internetlisten und Massenmedien.

Die öffentlich formulierte Kritik bildet nur eine kleine Spitze des Eisberges kritischer Forschungskommunikation. Diese findet vor allem in informellen Dyaden und Kleingruppen (z.B. in Form des Gegenlesens von Manuskripten im Freundeskreis) und im Rahmen der Arkanpraxis der Zeitschriftenverlage statt. Herausgeber, Referees, ManuskriptenreicherInnen interagieren nach dem Modell konspirativer Organisationen: mit Ausnahme des Herausgeber weiß kaum einer der Beteiligten voneinander,¹ und nur selten bekommen die Begutachteten die Gutachten überhaupt oder gar vollständig zu Gesicht. Als Folgen dieser Informationsvorenthaltung werden Fehler und Irrwege ressourcenvergeudend wiederholt.

Das Problem jeder nicht absolut geheimgehaltenen Evaluation: Die Evaluierten wissen von den Evaluationskriterien (zumindest gerüchteweise) und passen sich diesen an – scheinbar. In Australien wurde der Output (ohne Rücksicht auf Impact-Werte) zum Kriterium Nr.1 erklärt. Die Zahl australischer Publikationen stieg tatsächlich erheblich. Deren Journal-Impact-Faktor-Werte sanken allerdings, da kein Evaluationskriterium.

Die erhöhte Produktivität, als Artefakt der Evaluation, dürfte weder Forschungsqualität noch Erkenntnisfortschritt gefördert haben – eher im Gegenteil: Das Einhalten der Evaluationskriterien kostet Zeit und Energie. In

¹ So sind drei Viertel der Referees nicht einmal über das endgültige „Schicksal“ der von ihnen referierten Manuskripte informiert (Glogoff 1988).

den letzten Jahren nimmt neben verdeckten Widerstandsformen auch öffentlicher Protest Betroffener gegen Verzerrungen, Fehler und Dysfunktionen qualitativer wie quantitativer Evaluationen zu – vor allem gegen uninformatierten Einsatz durch Laien („Discounter-, „Quick-and-Dirty-Evaluation“). Dies äußert sich auch in der Unterstützung von OPEN-ACCESS-Initiativen, der regen Nutzung eingerichteter Server, im eigenmächtigen (halb- bis illegalen) kostenlosen Stellen eigener Journalartikel auf Webseiten.

3. Das Peer-Review-System gibt es nicht. Blinde Flecken & nicht-intendierte Effekte

Aus Platzgründen beschränke ich mich im Folgenden auf stichwortartige Kritik der Manuskript-Peer-Review-Verfahren im Journalwesen. Ich bestehe auf dem Plural: Es gibt weder ein „Peer-Review-System“ noch ein „Peer-Review-Verfahren“ im Singular, sondern nur höchst unterschiedliche Praktiken, die oft bloß den Namen gemeinsam haben (z.B. wenn als „Peers“ nicht FachkollegInnen, sondern Politiker, Pfarrer, Theologen, Funktionäre fungieren). Ein einheitliches Prüfsystem gibt es nicht: Zu unterschiedlich sind die Betriebsitten in einzelnen Disziplinen, Organisationen, Redaktionen. Die vielen Peer-Review-Varianten haben unterschiedliche Funktionen: Bei niedrigen Abweisungsraten (20-30 Prozent in der Physik) sollen Gutachter Aufsätze zum Druck vorschlagen, es dominieren serielle Verfahren: Ein Gutachter wird bestimmt, wenn dieser das Manuskript empfiehlt, wird es gedruckt; lehnt dieser ab, wird ein weiterer Gutachter beauftragt. Bei hohen Abweisungsraten (80-90 Prozent in der Psychologie) sollen Gutachter Argumente für die Ablehnung von Manuskripten liefern, es werden parallele Verfahren (mit zwei bis vier Gutachtern zugleich) bevorzugt. Bei diskrepanten Beurteilungen wird das Manuskript oft abgelehnt.

Peer Review fungiert als Prestige-Schmuck: Etliche Autoren bezeichnen ihre Veröffentlichungen, manche Herausgeber ihre Journale fälschlicherweise als anonym bzw. extern begutachtet. Selbst auf Journal-Verzeichnisse ist kaum Verlass: Nur die Hälfte von 784 klinischen Journalen wurde

von zwei verschiedenen Directories übereinstimmend als referiert angeführt (Eldredge 1997).²

Nur ein oft recht kleiner Teil der Beiträge wird in „referierten“ Journalen tatsächlich (und auch in der Endfassung) extern begutachtet. Beim US-„Passivraucherskandal“ nutzten dies Justiziere des beklagten Konzerns: Sie formulierten unter strategischen Gesichtspunkten „Letters“ an referierte Journale, die dort – gegen Honorierung mit den Namen akademischer Experten geschmückt – wie üblich nichtreferiert veröffentlicht wurden. Bei der Verhandlung verwiesen die Konzernanwälte erfolgreich auf diese in „referierten Journalen“ veröffentlichten wissenschaftlichen Beiträge (vgl. Kieran 1998).

Peer-Review ersetzt kritische Wissenschaftsöffentlichkeit. Da mit „Peer-Review-Prüfstempel“ versehen, werden Aufsätze auch ohne kritische Lektüre zitiert.³ Nach einer aufwendigen Auswertung und Hochschätzung befand sich nur 20 Prozent der zitierten Literatur jemals in der Hand der Zitierenden. 80 Prozent der Zitationen sollen demnach auf ‚Plünderung‘ fremder Literaturlisten beruhen (Simkin/Roychowdhury 2003). Referierte Journale ermöglichen „Schnellschusspublikationen“ und heizen so die Publikationsflut an.

Der Widerstand gegen das Neue zeigt sich in zahlreich dokumentierten Fehltritten von Gutachtern bzw. Herausgebern renommierter Journale (z.B. in Form abgelehnter, später mit dem Nobelpreis ausgezeichneten Arbeiten, vgl. Campanario 1996, Fischer 2002). Nach solchen ‚Eigentoren‘ setzten Journale die Begutachtung oft über viele Jahre aus. WissenschaftlerInnen halten das Gutachterwesen generell für konservativ (Armstrong 1997):

² Diese Unklarheit potenziert sich in der einschlägigen Forschung: Für Weller 2001 bleibt in ihrer extensiven Bestandsaufnahme wiederholt unklar, ob sich eine Studie auf referierte Journale bezieht oder nicht.

³ Auch Herausgeber und Gutachter scheinen eher lesefaul zu sein: Zwei Psychologen (Peters/Ceci 1982) hatten zwölf angesehenen Journalen ihrer Zunft je einen Aufsatz entnommen und diese bereits veröffentlichten Manuskripte (nach kosmetischen Änderungen) bei jeweils denselben Zeitschriften erneut eingereicht. Nur 3 von 38 Herausgeber bzw. Referees erkannten die in ihren Journalen vor 1 1/2 bis 3 Jahren bereits publizierten Aufsätze wieder. Universitätsprofessoren (n=672), befragt über ihre Erfahrungen mit Gutachtern, beklagten die achtlose, unaufmerksame Lektüre der Manuskripte (Bradley 1981).

Eingefahrene Themen, Modelle, Theorien hätten bei der Begutachtung gute Chancen, da den Referees geläufig. Bei innovativen oder gar interdisziplinären Ansätzen sei hingegen die Gefahr der Ablehnung enorm – aufgrund der Unkenntnis der Gutachter.

Wissenschaftliches Fehlverhalten: Nur ganz wenige Fälle von Plagiat, Betrug und Täuschung konnten durch Referees bzw. Editoren aufgedeckt werden. Selbst banalste Fehler (etwa: auch von Laien erkennbare Fälschungen in Bildern) entgingen Herausgebern wie Gutachtern – jedoch nicht aufmerksamen LeserInnen. Die derzeitigen Prüfprozeduren sind selbst bei Top-Journalen offensichtlich nicht imstande, die elementarste Qualitätsforderung, die man an wissenschaftliche Publikationen stellen kann – nicht gefälscht zu sein – zu gewährleisten (vgl. Fröhlich 2003b,c). „Nature“-Chefredakteur Campbell gab in einem Interview zu: Nicht die Journale, sondern nur ihre kritischen aktiven LeserInnen können gefälschte Artikel entlarven. Zudem mehren sich die Fälle schweren Missbrauchs der Gutachterposition für eigene Interessen und Zwecke („Wissenschaftsspionage“ nennt dies Stefanie Stegemann-Boehl 1994).

Mängel der Qualitätskontrolle: Bei experimentellen Untersuchungen mittels Fehlereinsatz in Manuskripten übersahen die Referees auch schwere Fehler. Qualitative Bewertungen von Artikeln in ausreferierten Journalen zeigen schwerste Mängel.⁴ Es scheint keineswegs ausgemacht, dass die Bearbeitungsforderungen von Herausgebern und Gutachtern die eingereichten Manuskripte verbessern. In der Ökonomie mehren sich kritische Gegenstimmen: Durch die wiederholten Überarbeitungen würden die Manuskripte nicht verbessert, sondern verschlechtert: Mit jeder Revision würden die Papers mehr und mehr formaler, ihr Informationsgehalt sinke (Frey 2004, Gans/Shepherd 1994).

In der Konkurrenz der Journale fördern künstlich überhöhte Abweisungs-raten das Prestige (der Mythos: je höher, desto wissenschaftlich hochwertiger; manche Journale deklarieren daher sogar Kürzungsforderungen als

⁴ Die Qualität referierter Aufsätze wird in inhaltsanalytischen Studien kritisiert. Z.B. könnten Artikel angesehener deutschsprachiger Soziologie-Journale durch andere SozialwissenschaftlerInnen kaum methodisch überprüft werden: Unverzichtbare Angaben dazu fehlten in fast allen Beiträgen der Stichprobe (Meinefeld 1985). Herausgebern wie Gutachtern scheint dies entgangen zu sein.

Abweisung). Die niedrigen Abweisungsraten etwa in der Physik und die hohen etwa in der Psychologie lassen sich durch jährliche hohe Seitenzahlen physikalischer Blätter (z.B. 10 000 Seiten) im Gegensatz zu geringen in den Sozialwissenschaften (z.B. 250 Seiten) erklären. Bei Zusammenfassung sozialwissenschaftlicher Journale zu ‚virtuellen Gesamtjournalen‘ wären ihre Abweisungsraten niedriger als in den Naturwissenschaften.

Denn wie zahlreiche Verbleibsstudien zeigen, werden fast alle abgelehnten Manuskripte nach (mitunter oftmaligen) Wiedereinreichungen in anderen Journalen veröffentlicht – z.T. mit gleich hohen, ja sogar höheren Impact Faktoren. Die derzeitigen Praktiken belohnen die Frustrationstoleranten und Hartnäckigen, sie bestrafen die Misserfolgvermeider (Atkinson). Sie dienen eher Herstellung und Erhalt von Journal- bzw. Verlagshierarchien als der Qualitätsfilterung. In der Ökonomie werden nach Azar (2004) Artikel zwischen drei- und sechsmal eingereicht. Die wiederholte Begutachtung mehr oder minder gleicher Texte durch eine ganze Heerschar an Referees (nach dem ‚Gesetz‘: Je ‚schlechter‘ ein Manuskript, umso mehr LeserInnen erhält es) sollte eine Gesamtrechnung der Peer-Review-Kosten berücksichtigen.

4. Kontingenzen: Von der Buchzensur zum Peer Review, von der sowjetischen Planwirtschaft zur Impact-„Messung“

„Peer“ bedeutet „Ebenbürtiger“, „hoher Adelige“, „Mitglied des britischen Oberhauses“. Aus ursprünglich informellen Treffen nobler Herren entstand zwischen 1660 und 1663 die Londoner „Royal Society“. Der König gestand eine Reihe von Privilegien zu. Für dieses königliche Wohlwollen war allerdings – wie in Frankreich – ein Preis zu bezahlen. „Die Organisation der Forschung in den Akademien und wissenschaftlichen Gesellschaften war ein Teil des Arrangements mit den etablierten kulturellen Gewalten Englands im Zeitalter der Restauration und des Ancien Régime in Frankreich“ (Felt et al. 1995, 38f.).

Mario Biagoli zeichnet die kontingente Entwicklung der Begutachtungspraktiken als Delegation der kirchlichen bzw. staatlichen Zensur: „Both of the first two state-sponsored or state-chartered academies, the Royal Society of London (1662) and the Académie Royale des Sciences of Paris

(1699), were granted the privilege to publish their own works. This was an extraordinary exception from the licensing and censorship systems that since the 16th century had been established by political and religious authorities throughout Europe in response to the perceived political and religious threats posed by the printing press. All texts, scientific or not, had to be reviewed and licensed in order to be printed and sold legally. The first scientific academies were not exempted from these requirements, but were allowed to administer them on their own.“

(Biagioli 2002, 14).

Bei der Gründung wissenschaftlicher Journale ging es weniger um effiziente Kommunikation (rege Korrespondenz fand ja statt), sondern eher um die Festsetzung von Priorität. Ehre kommt in den Naturwissenschaften jenem zu, der für den Erstentdecker gehalten wird. Dies führte zu erbitterten Streitigkeiten. Wichtige Funktion wissenschaftlicher Gesellschaften und Journale war die Bestätigung des Anspruchs auf Erstentdeckung, d.h. die Festsetzung des „eigentlichen“ Entdeckers bzw. Erfinders. Standesgemäße (adelige) Herkunft der ‚Zeugen‘ war wesentlich wichtiger als fachspezifische Kenntnis: Kaufleute seien unfreie Menschen, sie hätten keine ausreichende Kontrolle über ihre Handlungen und Meinungen, ebenso wenig Techniker – so hieß es (Felt et al. 1995, 38f.).

Die quantitative „Wissenschaftswissenschaft“ entwickelte sich im Rahmen der leninistischen Planwirtschaft. Maßzahlen dienten zur Festschreibung und Überprüfung der Fünfjahrespläne. Noch heute wird das Flaggschiff der quantitativen Evaluation, das Journal „Scientometrics“, in Budapest herausgegeben. Die Logik des rein Quantitativen, des Messens von Publikations- und Zitationshäufigkeiten, hat in ihren Effekten fatale (inverse) Ähnlichkeit mit der sowjetischen Planwirtschaftslogik: Die Messung des Plansolls von Weihnachtsbaumständern nach Tonnen Gewicht führte zur Produktion möglichst klobiger Exemplare (um so leicht und schnell das Soll zu erfüllen und die Planüberschreitungsprämie zu kassieren). Die Belohnung möglichst vieler Journalbeiträge ohne Berücksichtigung von Seiten- und CoautorInnenzahl, aber mit möglichst hohem Journal Impact (und ohne Berücksichtigung des Article Impacts) führt zur „Salamipublikationstaktik“: zur Zerteilung des Forschungsertrags in möglichst viele, möglichst hauch-

dünne Scheibchen mit wechselseitiger Beteiligung möglichst vieler CoAutorInnen aus möglichst vielen Institutionen aus möglichst vielen Ländern, gefördert von möglichst vielen Fonds.⁵ Dies wird in Evaluationen meist als Erfolg gefeiert. Doch zeigen nach der Zahl der CoautorInnen gewichtete Auswertungen einen Rückgang der Produktivität und Resonanz je Einzelautor.

5. Peer Review & Szientometrie: Gleichen sich Fehler & Verzerrungen wirklich aus?

Verteidiger des Systems korrelieren qualitative Bewertungen mit Zitationshäufigkeiten. Doch diese sind für dieselben Verzerrungen anfällig: Bereits etablierte Theorien, Methoden, WissenschaftlerInnen, Institutionen werden durch „Impact“-Werte belohnt, innovative, noch statusniedrige Außenseiter hingegen bestraft. Der „Matthäus-Effekt“⁶ wird in der Interaktion quantitativer und qualitativer Evaluation kaum eingedämmt, sondern weiter aufgeschaukelt. Etliche positive Korrelationen zwischen qualitativen Ratings und Journal-Impact-Faktoren sind Artefakte unzulänglicher Forschungsdesigns: Dem Gutachtergremium wurden die Impact-Werte der Journale bekannt gegeben, bzw. auf die Kontrolle verzerrender Einflüsse dieser Kenntnis verzichtet.

Zahlreiche Faktoren, die weniger mit wissenschaftlicher Qualität, als mit Insider-Bevorteilung zu tun haben, beeinflussen die Zitationshäufigkeit eines Artikels: Je länger der Artikel und je weiter vorne im Journal platziert, umso mehr Zitate werden erwirtschaftet. Zugestandene Länge wie Stellung des Artikels im Heft hängen von der sozialen Nähe zum Herausgeber ab (vgl. Campanario 1996).

Zahlreiche kritische Untersuchungen (v.a. zur verbreiteten Unsitte der „Ehrenautorenschaft“) belegen: Output- wie Input-Indikatoren, d.h. Produktivitätskennziffern wie Resonanzwerte einzelner AutorInnen, Gruppen und In-

⁵ Alle Beteiligten können stolz gestiegene Leistungsbilanzen vermelden: Die AutorInnen, Journalherausgeber, Institutionen, Forschungsförderer und Staaten. Die "least publishable unit" beträgt in der Psychologie - unter Abzug von Titelseite, Abstract, Danksagungen und Literaturliste – inzwischen eine Seite Text (inkl. Tabellen) für drei Coautoren, d.h.: eine Drittel Seite je Coautor.

⁶ Von Merton (1968, 1988) benannt nach dem Gleichnis vom anvertrauten Geld im Evangelium nach Matthäus: „Denn jene die haben, denen wird gegeben werden; jene die nicht haben, wird sogar das noch genommen werden.“

stitutionen sind eher Indikatoren für Macht, für soziales und ökonomisches Kapital, als für ‚reine‘ wissenschaftliche Leistung. Qualitätsbestimmungen wissenschaftlicher Leistungen im „universalistischen“ (Merton) Sinne – d.h. gerade: unter Absehung der Position der Sender – müssten diese kumulative Bevorteilungen durch intelligentes methodisches Vorgehen ausschalten.

6. Notwendige Reformen des wissenschaftlichen Publikationswesens – zu spät?

Herausgeber üben in Arkanpraxis und oft feudaler Kronprinzipfolge nicht selten über Jahrzehnte unbeschränkte Macht und Willkür aus. Gutachter dienen als Verteidigungslinie: So erfolgt nur ein Bruchteil der Manuskriptablehnungen gutachtergestützt, auch wenn in den Ablehnungsschreiben anders behauptet. Editoren und ihre Boards sind massiv US-dominiert, Frauen in entscheidenden Positionen extrem unterrepräsentiert (vgl. Braun/Dióspatonyi 2005; Addis/Villa 2003, Green 1998).

Es sollte daher nicht nur das Peer-Reviewing verbessert, sondern das wissenschaftliche Journalwesen insgesamt reformiert, Journale einem Akkreditierungsprozess unterzogen werden. Herausgabe und Begutachtung sollte unter konsequenter Nutzung informationswissenschaftlicher Techniken erfolgen. Plagiate wären z.B. mittels der „related documents“-Funktion (=Nachweis von Dokumenten mit ähnlicher Zitationsstruktur) leicht eruierbar. Fortbildung von Herausgebern und Gutachtern (oder Beistellung einschlägig qualifizierter AssistentInnen), der Einsatz professioneller Datenbanken und Plagiatsüberprüfungssoftware wären für „informed peer review“ unverzichtbar, aber kostenintensiv.

Doch ist es für eine Reform vielleicht schon zu spät? Werden sich konventionelle Journale halten können? Das Argument der „Knappheit“ des Platzes für Veröffentlichungen ist im digitalen Zeitalter obsolet. Die großen Wissenschaftsverlage mit logarithmischen Gewinnsteigerungen geraten durch die „Open Access“-Initiativen bereits in Bedrängnis: „Open Access“-Publikationen erhalten höhere Article Impacts als kostenpflichtige. So stellen immer häufiger AutorInnen mehr oder minder illegal ihre Papiere kostenlos ins Netz. Journale offerieren vermehrt kostenfreie digitale

Schlupflöcher. Der Niedergang der Papierjournale (ein Kind der Postkutschenära – welchen Sinne sollte es in digitalen Zeiten haben, Papiere, die nur wenig miteinander zu tun haben, zu einem Bündel zu schnüren?) ist aufgrund der ständig zunehmenden Bedeutung von Volltextdatenbanken und WWW-Servern unvermeidlich. Die Bedeutung einzelner Zeitschriften sinkt: Die Verlage stellen Datenbanken aller edierten Artikel und Bücher mit Suchfunktionen ins Netz. Zitationshinweise versuchen die Aufmerksamkeit auf andere Publikationen im eigenen Pool umzulenken.

Abschließend ein Gedankenexperiment: Publizieren ohne Peer-Review-„Zensur“ könnte die Qualität wissenschaftlicher Publikationen erhöhen. Öffentliche Kritik oder gar Aufdeckung als Plagiator oder Fälscher durch „Peer Monitoring“ ist blamabler als geheime Beanstandung hinter den Kulissen. Konventionelle Evaluation befrachtet die wissenschaftliche Kommunikation mit den Personalbewertungs- und Mittelverteilungsproblemen der Ministerien und Institutionen: Wir sollten vielleicht Mut aufbringen, hier neue Formen zu entwickeln, zwecks Stärkung geschwächter epistemologischer Funktionen wissenschaftlicher Kommunikation (Kritik, kognitive Konkurrenz, Theorien- und Methodenpluralismus) – und damit Förderung der „Wettbewerbsfähigkeit“ und Innovationskraft der europäischen Wissenschaft.

ZITAT:

Gerhard Fröhlich, 2006, „Informed Peer Review“ - Ausgleich der Fehler und Verzerrungen? In HRK (Hochschulrektorenkonferenz), Hg., Von der Qualitätssicherung der Lehre zur Qualitätsentwicklung als Prinzip der Hochschulsteuerung. Bonn, 193-204

Literaturverzeichnis

Addis, E./Villa, P. (2003): The Editorial Boards of Italian Economics Journals: Women, Gender, and Social Networking. *Feminist Economics* 9 (1), 75-91

Armstrong, J. S. (1997): Peer Review for Journals: Evidence on Quality Control, Fairness, and Innovation. *Science and Engineering Ethics* 3 (1), 63-84

Azar, O. F. (2004): Rejections and the Importance of First Response Times. *International Journal of Social Economics*, February 2004

Biagioli, M. (2002): From Book Censorship to Academic Peer Review. *Emergences* 12 (1), 11-45

Bradley, J. V. (1981): Pernicious publication practices. *Bulletin of the Psychonomic Society* 18 (1), 31-34

Braun, T./Dióspatonyi, I. (2005): US Scientists dominate as Journal Gatekeepers. *The Scientist* 19 (5), 10

Campanario, J. M. (1996): The Competition for Journal Space among Referees, Editors, and Other Authors and its Influence on Journals' Impact Factors. *Journal of the American Society for Information Science* 47 (3), 184-192

Eldredge, J. D. (1997): Identifying peer-reviewed journals in clinical medicine. *Bulletin of the Medical Library Association* 85 (4); 418-422

Felt, U. et al. (1995): *Wissenschaftsforschung*. Frankfurt a. M./New York

Fischer, K. (2002): „Ist Wissenschaftsevaluation unvermeidlich innovationshemmend?“, in: Pipp, E. (Hg.): *Drehscheibe E-Mitteuropa*. Wien, 109-128

Frey, B. S. (2004): Publizieren als Prostitution? *Perspektiven der Wirtschaftspolitik* 5 (3), 333-336

Fröhlich, G. (1999): Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? *GMD-Report* 61, 27-38

Fröhlich, G. (2002): Anonyme Kritik. Peer Review auf dem Prüfstand der empirisch-theoretischen Wissenschaftsforschung, in: Pipp, E. (Hg.): *Drehscheibe E-Mitteuropa*. Wien, 129-146

Fröhlich, G. (2003a): Wie rein ist die Wissenschaft? Fälschung und Betrug im rauen Wissenschaftsalltag, in: Etlzdorfer, Hannes et al. (Hrsg.) *echt_falsch*. Will die Welt betrogen sein? Wien, 72-93

Fröhlich, G. (2003b): „Visuelles in der wissenschaftlichen Kommunikation – z.B. Betrug und Fälschung“. *S – European Journal for Semiotic Studies* 15 (2-4), 627-655

Gans, J. S./Shepherd, G. B. (1994): How are the Mighty Fallen: Rejected Classic Articles by Leading Economists. *Journal of Economic Perspectives* 8 (1), 165 -180

Glogoff, S. (1988): Reviewing the Gatekeepers: A Survey of Referees of Library Journals. *American Society for Information Science Journal* 39 (6), 400-407

- Green, K. (1998): The Gender Composition of Editorial Boards in Economics. Royal Economic Society's Committee on Women in Economics, <<http://www.res.org.uk/society/pdfs/editoria.pdf>>
- Kiernan, V. (1998): Tobacco Industry Paid Letter-Writing Scientists. The Chronicle of Higher Education, 14. 8. 1998, A16
- Meinefeld, W. (1985): Die Rezeption empirischer Forschungsergebnisse – eine Frage von Treu und Glaube? Zeitschrift für Soziologie 14 (4), 297-314
- Merton, R. K. (1968): The Matthew Effect in Science. Science 159 (3810), 56-83
- Merton, R. K. (1988): The Matthew Effect in Science, II. Cumulative Advantage and the Symbolism of Scientific Property. ISIS 79, 606-623
- Peters, D. P./Ceci, S. J. (1982): Peer review practices of psychological journals: The fate of published articles, submitted again. Behavioral and Brain Science 5, 187-195
- Popper, K. R. (1969): Das Elend des Historizismus. Tübingen (c 1965)
- Simkin, M. V./Roychowdhury, V. P. (2003): Read before you cite! Complex Systems 14 (3), 269-274
- Stegemann-Boehl, S. (1994): Fehlverhalten von Forschern. Stuttgart
- Weller, A.C. (2001): Editorial Peer Review: Its Strengths and Weaknesses. Medford/NJ •