

# Irrläufer einer missverstandenen Szientometrie

ULRICH NAUMANN

*Die Szientometrie untersucht quantitative Dimensionen der wissenschaftlichen Entwicklung. Ihre Methoden und Ergebnisse werden auch zur Evaluation des wissenschaftlichen Outputs herangezogen. Ohne ein grundsätzliches Verständnis über die begrenzte Aussagefähigkeit massenstatistischer Datenerhebungen besteht die Gefahr der Fehlinterpretation, da sich ‚Wissenschaftlichkeit‘ nicht allein mit diesem Instrumentarium messen lässt. Im vorliegenden Beitrag wird darauf Bezug genommen, indem die Versuche geschildert werden, anhand der Publikationsziffern zwischen Universitäten und innerhalb von Universitäten Rangfolgen mit finanziellen Auswirkungen zu bilden. Ein zweites Thema ist die Kritik an der vermuteten Qualität von peer reviewed-Zeitschriften, die bei der quantitativen Messung eine Vorzugsstellung genießen. Ein drittes Thema ist die kritische Betrachtung des für Zeitschriften ermittelten impact factor als untaugliches Instrument der qualitativen Positionierung von Forschungsleistungen.*

## Einleitendes

Zu den wissenschaftlichen Interessengebieten von Walther Umstätter gehört die Szientometrie. Er vertritt sie seit vielen Jahren in Forschung und Lehre. Schon aus der Ulmer Zeit 1975-1982 stammen mehrere Beiträge, die er meist zusammen mit Margarete Rehm verfasst hat. Von Walther Umstätter stammt der Beitrag „Szientometrie“ in der 2004 erschienenen 5. Auflage des Standardwerkes „Grundlagen der praktischen Information und Dokumentation“ [24], und erst kürzlich hat er in dem von ihm wesentlich konzipierten und mitgetragenen „Berliner Bibliothekswissenschaftlichen Kolloquium“ einen Vortrag zu *Bradford's Law of Scattering* [25] gehalten und über die „Garfield'sche Konstante“ publiziert [26]. Umstätter kommt bei diesem Interessengebiet zugute, dass er neben der bibliothekarischen Professionalität über eine profunde naturwissenschaftliche Ausbildung verfügt, die er in diesem stark mathematisch-statistisch geprägten Wissenschaftsgebiet nutzbringend anzuwenden weiß.

Seinen eigenen Bekundungen zufolge ist er in seinen szientometrischen Überlegungen von Derek J. de Solla Price beeinflusst worden, mit dem er sich auch kritisch auseinandergesetzt hat [16].

Insbesondere dessen einprägsame Trennung wissenschaftlicher Forschungsarbeit in die Phasen des *Little Science* und des *Big Science* [16], deren Phasenwechsel Umstätter in die Zeit des *Big Bang*, also des Beginns der Weltraumfahrt mit dem *Man in the Moon*-Projekt der USA Ende der fünfziger, Anfang der sechziger Jahre des 20. Jahrhunderts legt<sup>1</sup>, greift er immer wieder gern auf, um Bezugspunkte zur Entwicklung der Szientometrie und ihrem Forschungsobjekt, den Publikationen, zu finden. Es sei nur am Rande erwähnt, dass Umstätter hier mit einer nicht unumstrittenen eigenen Definition von ‚Veröffentlichung‘ arbeitet, die unter anderem nur publik gemachtes Wissen von publizierter Information unterscheidet [23]. Dies hat aber wesentliche Auswirkungen auf die szientometrische Betrachtung und Abgrenzung der vielfältigen Formen, die uns heute in einer weit über die herkömmliche Form der auf Papier vorgelegten Veröffentlichungen begegnen, etwa im Bereich der digitalen und multimedialen Produkte. Hier erscheinen ganz neue Forschungsansätze unter Einbeziehung von Publikationen, die bisher von den herkömmlichen bibliometrischen ‚Gesetzmäßigkeiten‘ nicht erfasst werden (können) oder deren Existenz zumindest Veranlassung genug ist, diese Gesetzmäßigkeiten einer Überprüfung auf ihre Validität unter Einbeziehung der neuen Publikationsformen zu unterziehen.

Was erstaunt, ist, dass trotz gesichertem und laufendem Erkenntnisgewinn auf diesem wichtigen Gebiet der Messung von Forschungsleistungen, vor allem auch hinsichtlich ihrer Wirkung auf das Voranbringen weiterer Forschungen und dem damit verbundenen Fortschritt, die Ergebnisse szientometrischer Forschung im politisch und ökonomisch regulierten ‚Wissenschafts-Wettbewerb‘ weitgehend unbeachtet bleiben und Auswüchse zu beobachten sind, die jedem Fachmann unverständlich bleiben müssen [4; 6; 7; 8; 11; 19; 21; 23]. Einige dieser Auswüchse bei der Messung von ‚Forschungsleistungen‘ sollen mein Thema sein.

### **Little Science, Big Money**

Im Wettbewerb zwischen den deutschen Hochschulen um auskömmliche Etats und Forschungsförderung spielt unter vielen Parametern, die eine Position im ‚Ranking‘ beeinflussen, auch die Menge an Publikationen eine Rolle. Insbesondere Steuerungsmodelle zu einer leistungsgerechten Mittelverteilung zwischen den Hochschulen und innerhalb der Hochschulen greifen gern auf

---

<sup>1</sup> Wissenschaftshistoriker könnten auch andere Zeitpunkte nennen, etwa zwanzig Jahre zuvor das „Manhattan Projekt“ in Los Alamos, New Mexico (USA) mit Ziel, eine Atombombe zu bauen.

diesen Parameter zurück, weil deutsche Wissenschaft sich immer noch über die Menge an Forschungsoutput, seltener über seine Qualität definiert.

#### Interuniversitärer Vergleich von Publikationsleistungen

Man kann das exemplarisch an der Berliner Hochschulpolitik aufzeigen. Hier versucht der Senat von Berlin, die Finanzerwartungen der zahlreichen Berliner Hochschulen, darunter vier Universitäten, durch Hochschulverträge zu befriedigen. Mit ihnen soll durch die Vereinbarung bestimmter Summen über mehrere, meist vier, Jahre hinweg Planungssicherheit geschaffen werden, zugleich sollen damit aber auch die Obergrenzen der Hochschulfinanzierung festgelegt werden. Ob das Politikinstrument ‚Hochschulverträge‘ für die Hochschulentwicklung in Berlin vorteilhaft war, wird von anderen zu beurteilen sein. Um nun ein Wettbewerbselement in die Globalfinanzierung der Berliner Hochschulen hineinzubringen, werden immer größere Prozentsätze der Gesamtsumme nach Leistungskriterien vergeben. In der laufenden Vertragsphase (2006 bis 2009) wird dieser Prozentsatz von ursprünglich 6 Prozent (2002) auf 30 Prozent im Jahr 2008 gesteigert.

Die Leistungskriterien orientieren sich an Leistungen in Forschung und Lehre sowie an der Internationalität der jeweiligen Universität. Nicht zuletzt auf Betreiben der Freien Universität Berlin ist für die Jahre 2008 und 2009 der Parameter ‚Publikationen‘ in diesen Leistungsvergleich aufgenommen worden. Ohne mich in weiteren Einzelheiten verlieren zu wollen, sei festgestellt, dass immerhin etwa 10 Millionen Euro allein aufgrund dieses Parameters zur Verteilung kommen werden [10].

Bei dem Versuch, eine Vergleichsbasis für die Publikationen der drei großen Universitäten Freie Universität Berlin, Technische Universität Berlin und Humboldt-Universität zu Berlin zu finden, nach der diese Leistungsmittel ausgeschüttet werden können, zeigt sich allerdings, dass es eine solche Vergleichsbasis trotz relativ vergleichbarer Aufgabenstellung nicht gibt. Auf der einen Seite steht die seit 25 Jahren gedruckt und seit 12 Jahren auch kumulativ als Datenbank veröffentlichte Universitätsbibliographie der Freien Universität Berlin mit ca. 7 000 bis 8 000 Publikationsnachweisen pro Jahr [27]. Dazwischen gibt es eine unvollständige Publikationssammlung der Technischen Universität Berlin, die auf dem Prinzip der jährlichen Selbstaufschreibung durch die Lehrstühle beruht (LINF-System) und wesentlich schärfere Auswahlkriterien für zu meldende Publikationen vorsieht (Bücher werden nur aus ‚wissenschaftlichen‘ Verlagen, Zeitschriftenaufsätze werden nur bei *peer reviewed*-Zeitschriften akzeptiert). Auf der anderen Seite, bei der Humboldt-Universität zu Berlin, gibt es gar keine Universitätsbibliographie. Publikationen werden lediglich im Rahmen der mit einer fünfjährigen

Periodizität vorgesehenen Forschungsevaluation eines Faches durch Nennung von drei bis fünf *key publications* der letzten fünf Jahre und eine als Anhang mögliche Auflistung der übrigen, nach verschiedenen Aspekten und Erscheinungsformen getrennten übrigen Publikationen genannt, wobei auch hier wiederum die Aufsätze nach dem vermeintlichen Qualitätsmerkmal *peer reviewed* (s. unten) in zwei Gruppen getrennt werden.

Es ist also festzustellen, dass es aufgrund der bisherigen inneruniversitären Vorgaben keine Vergleichsbasis gibt: der relativ vollständigen Universitätsbibliographie der Freien Universität Berlin steht eine limitierte Erfassung der Publikationsleistungen bei der Technischen Universität Berlin und ein sehr unvollständiger, weil auch von den Rhythmen der Forschungsevaluation abhängiger Teilausschnitt aus den Publikationsleistungen der Humboldt-Universität zu Berlin gegenüber<sup>2</sup>. Diesem Problem der fehlenden Vergleichsbasis wird man auch nicht mit objektiven szientometrischen Methoden beikommen. Es wird spannend sein zu beobachten, welche Kompromisse für die Bewertung von publizierten Forschungsleistungen eingegangen und auch von der Senatsverwaltung akzeptiert werden, um nicht wegen mangelnder Ausgestaltung des Publikationsparameters die Summe von 10 Millionen EURO zu verlieren. Erste Vorschläge gehen in die Richtung, nur einen Teilausschnitt der Publikationsleistungen heranzuziehen, etwa in Fächern wie Mathematik oder Physik, die an allen drei Hochschulen vertreten sind. Die Zahl solcher Vergleichsfächer ist aber wegen des sehr unterschiedlichen Fächerspektrums der Hochschulen sehr gering und scheint eine sehr schmale Basis zu sein, um darauf aufbauend jährlich 10 Millionen EURO zu verteilen.

#### Inneruniversitärer Vergleich von Publikationsleistungen

Die Freie Universität Berlin setzt im Rahmen ihres hochschulinternen Steuerungsmodells auf das Konzept, neben der Ausweisung eines begrenzten Grundetats den Fachbereichen Leistungsmittel aufgrund der erzielten Ergebnisse in Forschung und Lehre zuzuteilen. Auch hier spielen wegen des gut eingeführten Instruments ‚Universitätsbibliographie‘ die Publikationsleistungen eine Rolle. Sie werden bei der Erfassung für die Bibliographie in die verschiedenen Publikationsformen Buch, Sammelwerksbeitrag, Zeitschriftenaufsatz, Rezension, Patente, Rundfunk- und Fernsehbeitrag unabhängig von der medialen Form getrennt, bei den Büchern sogar weiter in vier Umfangskategorien. Aber auch die Tätigkeiten als Herausgeber werden in einer eigenen

---

<sup>2</sup> Der Erste Vizepräsident der HU, Hans Jürgen Prömel, schätzt die Zahl der jährlich erscheinenden Aufsätze auf etwa 2 200, eine im Vergleich etwa zur etwas größeren FU wohl zu geringe Zahl [17].

Erfassungsmaske berücksichtigt. Die unterschiedlichen Publikationsformen werden für die einzelne wissenschaftliche Einrichtung statistisch getrennt in Listenform an das Präsidium gemeldet, dort nach einem festgelegten Punkteschema (10 Punkte für eine Monographie, 8 Punkte für ein Patent, 4 Punkte für einen Aufsatz, 1 Punkt für die übrigen Publikationen) ausgewertet und dann in die Leistungsmittelberechnung einbezogen. Dissertationen, die in der Universitätsbibliographie ebenfalls nachgewiesen werden, bleiben unberücksichtigt, weil Promotionen bei den Lehrleistungen bewertet werden.

Hier zeigt sich nun in den letzten Jahren ein Phänomen, das ein für mich peinliches Schlaglicht auf den Charakter solcher Wettbewerbe um die knappen Mittel wirft. Ohne Steuerung von außen, etwa die Bitte, auch solche Beiträge zu melden, werden vermehrt Leistungen, die dem ‚Wissenstransfer‘ zwischen Universität und Öffentlichkeit zugeordnet werden können, als Publikationen mit der Forderung um Aufnahme in die Universitätsbibliographie gemeldet: die Interviews, die mit Angehörigen der Universität zu tagesaktuellen Themen geführt werden. Solche Interviews zeugen zwar von dem Renommee, das einzelne Wissenschaftler dieser Universität genießen, können aber schwerlich als eigenständige wissenschaftliche Leistungen angesehen werden, zumal ihnen regelmäßig auch das Kennzeichen der geistigen Urheberschaft, also die Autoreneigenschaft, fehlt. Für das Publikationsjahr 2005 liegen bisher über 500 solcher Meldungen vor. Es ist daher festgelegt worden, solche Interviews nicht (mehr) in der Universitätsbibliographie zu verzeichnen, da ihnen der Publikationscharakter durch einen Angehörigen der Hochschule fehlt, obwohl, und das ist eine erfreuliche Feststellung, viele dieser Dokumente über die Archivserver der Zeitungen und Rundfunkanstalten zumindest vorübergehend (Problem der gesicherten Langzeitarchivierung!) online nachgewiesen sind.

Bezeichnend ist auch, dass solche Interviews in der Regel nicht in den Publikationslisten auftauchen, die von den Wissenschaftlern selbst im Internet präsentiert werden. Es wird also deutlich, dass das Bestreben, solche Produkte des ‚Wissenstransfers‘ als Publikationen auszuweisen, ausschließlich mit der Hoffnung verbunden wird, im Rahmen des nur statistisch basierten inneruniversitären Leistungsvergleichs besser als andere Einrichtungen dazustehen und sich ein größeres Stück vom Verteilungskuchen abschneiden zu können. Dass dieses Ziel in den vergangenen Jahren auch bei den früheren Meldungen einer wesentlich geringeren Zahl von Interviews nicht erreicht wurde, lag daran, dass solche Interviews bei der Weitergabe der Publikationsstatistik an das Präsidium ‚ausgeblendet‘ wurden, was aber mehr auf einer fehlenden Programmanpassung unserer Auswertungsprogramme als auf einer inhaltlichen Wertung beruhte, da ein Bibliograph kein Zensor sein darf. Bei einer solchen Gesamtschau der Publikationsleistungen, wie sie die Universitätsbibliographie der Freien Universität Berlin mit Recht für sich in

Anspruch nehmen kann, wäre bei der inneruniversitären Mittelverteilung ein Einsatz szientometrischer Verfahren wie etwa der Zitationsanalyse analog dem ISI-Konzept zwar denkbar (Ralf Brugbauer berichtet über die Verwendung des Impact Factors nur beim medizinischen Fachbereich der Philipps-Universität Marburg aufgrund der Meldeergebnisse für die Universitätsbibliographie [3, S. 520]) würde aber bei dem in Betracht zu ziehenden Fächerspektrum, den Publikationsformen und den Publikationsorten instrumentell versagen [13].

### **Sleeping peers?**

Wenn das britische Fernsehen aus dem *House of Lords*, der parlamentarischen Versammlung der *Peers of England*, berichtet, zeigt es bei Plenardebatten des öfteren Bilder von hochangesehenen, aber schlafenden älteren Herren [14]. Das sollte kein Vorbild bei der Kennzeichnung einer Zeitschrift als *peer reviewed* sein. Da allgemein diese Kennzeichnung als besonderes Qualitätsmerkmal gilt (auch die Technische Universität Berlin und die Humboldt-Universität zu Berlin stellen bei den oben genannten Publikationslisten qualitativ unterscheidend darauf ab), wirkt sich die verstärkte Berücksichtigung solcher Zeitschriften bei der Platzierung von Aufsätzen durch die Autorenschaft auch auf die Messergebnisse szientometrischer Methoden aus.

Es gibt viele Definitionsversuche darüber, was ein *peer reviewed journal* charakterisiert. Eine relativ umfassende und viele Aspekte beschreibende Definition stammt vom Evergreen State College in Olympia, Wash.:

“A peer reviewed journal can be recognized by the following characteristics:

1. It is concerned with academic study.
2. The articles contain bibliographies of cited and related works.
3. The articles are written and reviewed by experts in the field.
4. The author’s credentials and institutional affiliation are included.
5. Often an abstract of the article is included at the beginning of the article.
6. The articles report on original research, or reviews and evaluates material that has already been published, or expands and refines a theory.
7. The journal is published by a professional association, society, research association, or academic institution.” [5]

Zweifellos ist das Etikett *peer reviewed* auch ein Marketing-Aspekt und wird daher auch von Zeitschriften angestrebt, die einen solchen Begutachtungsprozess nicht kennen oder nur dem Anschein nach durchführen. Selbst wenn es

einen solchen anerkannten Überprüfungsprozess gibt, haben jüngste Ereignisse gezeigt, dass da einiges schief gehen kann. Dabei ist ausgerechnet eine der anerkanntesten Zeitschriften auf dem naturwissenschaftlichen Gebiet, *Science*, in den Verdacht geraten, dass ihre *peers* geschlafen haben.

Es geht um den Fall des südkoreanischen Klon-Forschers Hwang Woo Suk. Der Bielefelder Soziologe Peter Weingart, der durch seine präzisen Analysen des Wissenschaftsbetriebs bekannt ist, bezeichnet ihn als „Klon-König“, dessen Regentschaft [...] nur wenig länger als die eines Karnevalsprinzen gedauert (hat)“ [29]. Weingart geht in dem Artikel der Frage nach, warum Betrug gerade in der medizinischen Forschung häufig geworden zu sein scheint, und identifiziert neben dem vielen Geld, das Verlags-, Werbe- und Pharmakonzerne für die Marktbesserstellung ihrer Produkte aufwenden, vor allem das Versagen des zentralen Mechanismus für die Herstellung verlässlichen Wissens, des *peer review*, also das Prüfen von Forschungsergebnissen vor ihrer Veröffentlichung durch kompetente Fachkollegen.

Nach Angaben aus dem *Web of Science* vom 23.3.2006 wurde der in Band 303 der Zeitschrift *Science* veröffentlichte Aufsatz von Hwang und anderen insgesamt 192 mal zitiert. Es wäre interessant, in den den Aufsatz zitierenden Quellen zu überprüfen, ob sich dort eine kritische Auseinandersetzung mit den Ergebnissen von Hwang findet. Eine anhand der Abstracts angestellte nicht repräsentative Stichprobe, die sich auf die Zeit vor der Aufdeckung des Fälschungsskandals bezog, zeigte oberflächlich ein anderes Ergebnis. Eine genaue Prüfung sollte aber Spezialisten vorbehalten bleiben.

Für die Dissemination von Wissen, für die die Szientometrie ebenfalls einige Gesetzmäßigkeiten aufzudecken versucht, ist ebenfalls beachtlich, dass der in Band 303 von *Science* veröffentlichte Aufsatz von Hwang und anderen in einer GOOGLE-Recherche vom 10.1.2006 in über 200 000 Dokumenten enthalten ist, also dem Tausendfachen im Vergleich zu den im *Web of Science* aufgeführten Aufsätzen. Dies zeigt die Mächtigkeit des Internets für die Verbreitung von Wissen, die unabhängig von der teilweise stark verzögerten Erscheinungsdauer in Fachzeitschriften rasend schnell vonstatten gehen kann. Auch hier wäre am Beispiel dieses Aufsatzes überprüfenswert, wie viele ältere Beiträge im Vertrauen auf die durch *peers* geprüfte Qualität des Aufsatzes die dort veröffentlichten ‚Forschungsergebnisse‘ für bare Münze genommen haben. Andererseits waren es südkoreanische Forscher, die das ‚urdemokratische‘ Internet nutzten, um kritische Fragen zu den Forschungsergebnissen von Hwang zu stellen. Es war wohl die einzige Möglichkeit, den mit nationalen Ehren und Millionen von Forschungsgeldern überhäuften Klon-Forscher ohne die Kontrollmechanismen ‚etablierter‘ Wissenschaft zu stoppen.

Der Editor-in-Chief von *Science*, Donald Kennedy, zog am 12.1.2006 die ‚Notbremse‘ und erklärte in *Science Express* zwei der in *Science* veröffent-

lichten Artikel von Hwang für „retracted“. Etwas kurios klingt seine Entschuldigung gegenüber den weltweit getäuschten Wissenschaftlern:

“*Science* regrets the time that the peer reviewers and others spent evaluating these papers as well as the time and resources that the scientific community may have spent trying to replicate these results.” [18; 20]

Hier werden auch die *peer reviewers* von *Science* in Schutz genommen, deren ureigenste Aufgabe es gewesen wäre, mit ihrem Expertenwissen diesen Skandal zu verhindern.

Der Skandal um die Veröffentlichungen von Hwang zeigt aber nur die kleine Spitze eines Eisbergs von Fehlurteilen, die sich in den *peer reviewed*-Zeitschriften finden. Er macht in seinen Folgewirkungen auch bedenkliche Erscheinungsformen von wissenschaftlicher Forschung deutlich, wenn einer der Mit-Autoren eines anderen sich als Fälschung herausgestellt habenden Artikels von Hwang, Gerald Schatten, nun erklärt, er habe nur seinen Namen unter diesen Artikel gesetzt, ohne sich mit dem Inhalt auseinanderzusetzen. Bei Schatten handelt es sich immerhin um einer der renommiertesten amerikanischen Reproduktionsmediziner mit hoher nationaler und internationaler Reputation [28]. Es hat den Anschein, als wollte hier jemand ohne Billet auf einen fahrenden Zug aufspringen, um am Ende mit den anderen erfolgreich ans Ziel zu kommen, aber bei dem ersten Nothalt abgesprungen ist.

In den Vereinigten Staaten von Amerika ist man sich der Unzuverlässigkeit des *peer review* von Zeitschriftenartikeln gerade im medizinischen Bereich bewusst. Regelmäßig veranstaltet die *American Medical Association* Kongresse speziell mit dieser Thematik, zuletzt den fünften Kongress in Pittsburgh im September 2005. Aus der Kongressanzeige des *American Medical Association*: “The Fifth International Congress on Peer Review and Biomedical Publication was held September 16-18, 2005, in Chicago, Illinois. As with the previous Congresses, our aim was to improve the quality and credibility of biomedical peer review and publication and to help advance the efficiency, effectiveness, and equitability of the dissemination of biomedical information throughout the world. Four hundred seventy participants from 38 countries attended the Congress.” [1] Drei Tage lang diskutierten also fast 500 Teilnehmer aus 38 Ländern über Probleme des *peer review* vor allem in medizinischen Zeitschriften.

In eigens dazu angestellten Studien über die Verlässlichkeit des *peer review*-Modells für die Begutachtung wissenschaftlicher Zeitschriften zeigen sich hohe Fehlerraten, die einerseits dadurch bedingt sind, dass durch verbesserte Forschungsmöglichkeiten andere Aussagen zustande kommen (das ist das Schicksal jeder wissenschaftlichen Forschung, die durch bessere instrumentelle Ansätze widerlegt werden kann), aber auch Fehlerraten für eine nicht

unbeachtliche Zahl von Forschungsergebnissen, die eindeutig auf bewusste Falschinterpretation der Forschungsergebnisse zurückzuführen sind. In Deutschland sind noch die Skandale um die Krebsforscher Friedhelm Herrmann und Marion Bruch (1997) (wo durch eine wissenschaftliche Kommission letztlich 29 Publikationen als gefälscht ermittelt wurden und bei 65 Arbeiten konkreter Verdacht auf Fälschung besteht) und den Physiker Jan Hendrik Schön (2004) (der zum „retract“ von 8 Artikeln in Science und fünf Artikeln in Nature führte!) in ungueter Erinnerung.

Im deutschsprachigen Raum kann Gerhard Fröhlich von der Johannes-Kepler-Universität Linz zu den profundesten Kritikern des *peer review*-Systems, aber auch der Szientometrie gerechnet werden. In seinem hier angeführten Aufsatz legt er an stupenden Beispielen dar, wie unzuverlässig das *peer review*-System als Qualitätskontrolle für wissenschaftliche Forschungsergebnisse funktioniert [7]. In seinen Reformvorschlägen fordert er, wenn man nicht ganz auf dieses fragwürdige Begutachtungsverfahren verzichten will, unter anderem neben der Offenlegung des gesamten Begutachtungsverfahrens ein Dreifachblindverfahren, bei dem weder die Herausgeber noch die Gutachter die Autoren der zur Begutachtung eingereichten Beiträge kennen, und angemessen für den Aufwand entschädigte *peers*. Umso unverständlicher ist dann – pars pro toto – die Forderung des Leiters eines großen Wirtschaftsforschungsinstituts an seine mehr als hundert wissenschaftlichen Mitarbeiterinnen und Mitarbeiter, dass sie wenigstens einen Zeitschriftenaufsatz pro Jahr in einer *peer reviewed*-Zeitschrift platziert haben müssen, um die Qualität ihrer Arbeit nachzuweisen [30].<sup>3</sup>

Würde man in der Hochschulpolitik und der Wissenschafts- und Forschungspolitik solche fundiert kritischen Beiträge zum *peer review*-System ernster nehmen, würde die Vorstellung, mit dem vermeintlichen Qualitätsmerkmal *peer reviewed* eine Vergleichbarkeit von Forschungsleistungen herstellen zu können, rasch schwinden.

---

<sup>3</sup> In den Jahresberichten des Instituts werden diese Aufsätze in einer eigenen Rubrik nachgewiesen. Der Leiter selbst ist *peer* für 33 Fachzeitschriften. Hierzu noch eine Aussage von ihm, die die Objektivität von *peer reviews* in Frage stellt: „Bevor man mich in Deutschland kannte, wusste die internationale Wissenschaft etwas mit meiner Arbeit anzufangen“, ist XXX überzeugt und genießt den Spaß am Erfolg. Das ist wohl auch der Grund, wieso er gerne auf seinen Posten als erster deutscher Programmdirektor des Centre for Economic Policy Research (CEPR) in London von 1991 bis 2001 zurückblickt. „Dort konnte ich beeinflussen, wer wo welche Artikel publiziert und wer auf welcher Tagung vorträgt.“

„Die Tatsache, dass die Verfahren sehr differenziert betrachtet werden müssen, nur eine eingeschränkte Aussagekraft haben und mitunter (abhängig vom Fach oder der Datenlage) aus prinzipiellen Gründen völlig versagen können, wird von den Auftraggebern bibliometrischer Evaluationen allerdings kaum wahrgenommen – zu stark ist offenbar der Wunsch nach einer automatischen Leistungsbeurteilung ohne Fachkenntnis, so dass Kritik an den Verfahren schnell als Nörgelei am Detail abgetan wird und die Meinung vertreten wird, dass bibliometrische Methoden jedenfalls besser als nichts seien und schon ‚irgendetwas‘ aussagen würden.“ [11]

### Der falsche Anstoß?

Neben der nur scheinbaren Qualität eines *peer reviews* sind andere, aus bibliometrischen Messungen gewonnene Werte ebenso sehr kritisch einzuschätzen: die Zitationsanalyse und der *impact factor* einer Zeitschrift, mit dem ihr Rang im Forschungswettbewerb belegt werden soll. Während solche Arbeiten früher mühsam mit der Hand und durch das Auszählen zahlreicher Quellen angegangen wurden, kann man sich heute auf das umfangreiche Datenmaterial stützen, das das *Institute of Scientific Information (ISI)* auch online zur Verfügung stellt. Inzwischen haben viele Untersuchungen gezeigt, dass die bibliometrische Analyse auf Basis einer statistischen Auswertung einer begrenzten Zahl von wissenschaftlichen Publikationen trotz des „Bradfordschen Gesetzes“ ein äußerst fragwürdiges Instrument ist, um wissenschaftliche Leistung, sei es einzelner Forscher, Forschergruppen oder Forschungseinrichtungen qualitativ zu messen. Um hier Solla Price zu zitieren:

„One such scale – the traditional one used by deans and other employers as a measure of scientific success – is the number of publications produced by each man in accepted scientific journals. Let it be freely admitted at the outset that this is a bad scale. Who dares to balance one paper of Einstein on relativity against even a hundred papers by John Doe, Ph. D., on the elastic constant of the various timbers (one to a paper) of the Forest of Lower Basutoland?“ [15, S. 40]

Friedemann Mattern hat dargestellt, dass die bibliometrische Analyse, die sich etwa an der Zitationsanalyse aufgrund des Datenmaterials des *ISI Science Citation Index* orientiert, für das Fach Informatik völlig fehlgeht, weil sich die Ergebnisse informatischer Forschung nicht in den vom *Science Citation Index*

fast ausschließlich ausgewerteten Zeitschriften finden, sondern in Konferenzbeiträgen und Forschungsberichten, die jedoch vom *ISI Science Citation Index* als Quellen völlig unberücksichtigt bleiben [11]. Für viele weitere Fächer, deren Schwerpunkt außerhalb des engen Kreises der im *Citation Index* in seinen verschiedenen Ausprägungen berücksichtigten Zeitschriften liegt oder in einer anderen Forschungs- und Publikationskultur arbeiten, ist demnach die Zitationsanalyse aufgrund der Nachweise im *Citation Index* ein völlig unzulängliches Instrument, um auch nur den quantitativen Output an Forschungsergebnissen zu messen. Über die Qualität von Forschungsergebnissen können solche Analysen sowieso keine Aussage treffen oder unterstützen.

Wer also die ‚wissenschaftliche Bedeutung‘ eines Autors anhand der Zitate seiner Aufsätze bestimmen will, wird in vielen Fällen falsch liegen, zumal weitere Forschungen gezeigt haben, dass die monographische Literatur weiterhin für die wissenschaftliche Arbeit in vielen Fächern von ausschlaggebender Bedeutung ist als die Vielzahl von Aufsätzen, die publiziert werden. So verwundert in der Praxis, bei der Erstellung der Universitätsbibliographie, dass mehr als die Hälfte der Meldungen bibliographisch nachrecherchiert werden muss. Die Autoren scheinen also keinen großen Wert auf der Wahrnehmung ihrer Arbeit in der *scientific community* Wert zu legen.

Wenn beim *ISI Science Citation Index* noch berücksichtigt wird, dass das Unternehmen ISI seit 1992 zum Thomson Konzern gehört, der unter anderem mit wissenschaftlichen Publikationen sein Geld verdient, mögen Zweifel an der Unabhängigkeit des Instruments für bibliometrische und szientometrische Arbeiten nicht ungerechtfertigt sein, zumal die Positionierung der Verlagszeitschriften in den ISI-Datenbanken auch ein Marketing-Argument ist [2; 9; 30]. Andererseits können die bibliometrischen Ergebnisse, die sich durch das Auszählen von Zitaten und den Vergleich statistischer Reihen ergeben, auch durch die Gestaltung der Zeitschriften und durch bestimmte Zitationsverhaltensweisen manipuliert werden.

Wenn zum Beispiel für die Editionspolitik bestimmter Zeitschriften gilt, dass dort zur Prüfung eingereichte Aufsätze möglichst viele Aufsätze aus dieser Zeitschrift zitieren sollen, kann das ein Zeichen dafür sein, dass diese Zeitschrift tatsächlich die maßgeblichen Forschungsergebnisse publiziert und man sich um keine andere mehr im Rahmen des *current awareness* kümmern muss. Ein Beispiel dazu ist die Darstellung der Zeitschrift *Health affairs*, die ausdrücklich auf ihre Position in den *Journal Citation Reports* hinweist [9]. Es kann aber genauso gut sein, dass man damit den Rangwert der Zeitschrift, falls sie in den *Journal Citation Reports* des ISI gelistet wird, erhöhen will. Nicht ohne Grund enthalten die Auswertungsroutinen von ISI Verfahren, mit denen

die Selbstzitationen von Autoren und Zeitschriften ausgeblendet werden können.

Wolfgang G. Stock hat in einem umfassenden Beitrag zu den *Journal Citation Reports* (JCR) des ISI (er selbst nennt ihn „ein umfangreiches Kompendium zu diesem wichtigen Produkt“) [22] resümierend festgestellt, dass trotz aller methodischen Schwächen und der doch geringen Menge an ausgewerteten Zeitschriften derzeit nicht Besseres verfügbar ist, dass sie zwar kein alleiniges Instrument bei der Entscheidungsfindung sein können, aber eine Entscheidung ohne Berücksichtigung der Daten der *Journal Citation Reports* suboptimal gefällt werden muss [21, S. 37]. In der fachwissenschaftlichen Diskussion kann aber die Kritik an den JCR, die in seinem Beitrag und auch in vielen Fachbeiträgen zum Thema enthalten ist, wichtige Anregungen geben, sich weiterhin mit Fragen der Wissenschaftsmessung auf massenstatistischer Basis zu beschäftigen.

Auch hier wäre zu wünschen, dass die fachwissenschaftliche Diskussion sich mit ihren Ergebnissen so deutlich artikulieren kann, dass sie auch von Entscheidungsträgern nicht überhört werden kann [21]. So würde es nicht vorkommen, dass ein Berliner Hochschulpräsident einem Wissenschaftler gegenüber die Erwartung äußert, er möge zum Nachweis seiner hohen Qualität Beiträge in wissenschaftlichen Zeitschriften veröffentlichen, die mindestens einen Impact-Factor von „2“ aufweisen.

Wer die Bildung von Impact-Faktoren auch nur einigermaßen durchschaut, wird über eine solche Forderung nur lächeln können. Dabei empfiehlt sich besonders die Ko-Autorenschaft, die in manchen Publikationszirkeln auch als ‚Ehren-Autorenschaft‘ vergeben werden kann [8]. Hier kann nach dem Prinzip der *least publishable unit* relativ leicht diese Forderung erfüllt werden. Ein (aus meiner Praxis in der Bibliographie an der Universität bekannter) Aufsatz von R. J. Mural u. a., in *Science* 2002 [12] enthält bei einem Wortumfang von ca. 6 200 Wörtern nicht weniger als 103 Autoren, deren durchschnittlicher Beitrag sich damit auf 60 Wörter reduziert. In der Psychologie liegt dieser Wert nach Fröhlich inzwischen bei einem Drittel einer Seite pro Autor [6, S. 32].

Der Impact-Factor kann nur die Rangfolge einer Zeitschrift innerhalb einer Gruppe fachlich gleichorientierter Zeitschriften (bzw. von ISI einer oder mehreren der dort gebildeten ca. 160 Gruppen zugeordneter Zeitschriften) anzeigen, niemals aber die durchschnittliche Qualität der in dieser Zeitschrift publizierten Beiträge und schon gar nicht die Qualität eines einzelnen Beitrags! Andererseits zeigt eine solche vom Impact-Factor einer Zeitschrift ausgehende Forderung, wie missbräuchlich der szientometrische Ansatz, mit dem Eugene Garfield, der ‚Erfinder‘ dieser Form von Zitationsanalyse, einst begann (Garfield wollte keineswegs die Qualität, sondern nur die

Nutzenstiftung einer Arbeit für die laufende Forschung messen) verwendet werden kann.

Vielleicht ist es an der Zeit, wieder an die Anfänge der Szientometrie, Bibliometrie und Informetrie zurückzudenken und aus der Grundabsicht, Verbindungen der wissenschaftlichen Forschung aufzuzeigen, neue Forschungsansätze stärker zu kultivieren, die über das Messen und Zählen hinaus Wissenschaft nicht nur als Management-Prozess, sondern wieder als soziokulturellen Prozess begreift. Gerhard Fröhlich fordert auf, die von ihm so bezeichnete „evaluative Szientometrie“ zu ergänzen und plädiert

„für eine wissenschaftstheoretisch aufgeklärte und sozialwissenschaftlich geläuterte Szientometrie, welche die WissenschaftlerInnen, ihre Institutionen, Prozesse und Produkte und die Beziehungen zwischen ihnen nicht mehr als Gase, Heizplatten, Erkältungskrankheiten und Parasiten modelliert [hier nimmt Fröhlich Bezug auf sein einleitendes Kapitel, in dem er die Modellansätze der szientistisch orientierten Szientometrie beschreibt], sondern als menschliche, kulturelle, soziale – und sich daher auch mit sozial- und kulturwissenschaftlichen Konzepten und Verfahren anzufreunden vermag.“ [6, S. 36]

Zweifellos ist bei solchen Forschungen auch die jahrzehntelange Expertise von Walther Umstätter gesucht und gefragt, und so soll dieser kleine Beitrag auch ein „Impact“ für weitere Arbeiten von ihm sein.

#### Literatur und Internetquellen

- 1 AMERICAN MEDICAL ASSOCIATION. (2005). The Fifth International Congress on Peer Review and Biomedical Publication, *Peer Review Congress*. URL: <http://www.ama-assn.org/public/peer/peerhome.htm>.
- 2 BLACKWELL PUBLISHING. Top Journals in their Fields *News. January/February 2006*. URL: [http://www.blackwellpublishing.com/librarians/newsletter\\_html/January06-Newsletter29.htm](http://www.blackwellpublishing.com/librarians/newsletter_html/January06-Newsletter29.htm).
- 3 BRUGBAUER, R. (1998). Bibliothekarische Erfahrungen mit dem Impact Factor. *Bibliotheksdienst* 32, 506-512.
- 4 *Cybermetrics, international journal of scientometrics, informetrics and bibliometrics*, 1.1997.
- 5 THE EVERGREEN STATE COLLEGE LIBRARY. Definition of Peer-Reviewed. URL: <http://www.evergreen.edu/library/catalog/peerreviewed.htm>.

- 6 FRÖHLICH, G. (1999). Das Messen des leicht Meßbaren: Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? In: J. Becker & W. Göring (Hrsg.). *Kommunikation statt Markt: zu einer alternativen Theorie der Informationsgesellschaft* (S. 27-38). St. Augustin: Gesellschaft für Mathematik und Datenverarbeitung.
- 7 FRÖHLICH, G. (2003). Anonyme Kritik: Peer Review auf dem Prüfstand der Wissenschaftsforschung. *medizin – bibliothek – information* 3, Nr. 2, 33-39. URL: [http://www.akh-wien.ac.at/agmb/mbi/2003\\_2/froehlich33-39.pdf](http://www.akh-wien.ac.at/agmb/mbi/2003_2/froehlich33-39.pdf).
- 8 Fröhlich, G. (Gasthrsg.) (2006). [Themenheft mit dem Schwerpunkt Plagiate & unethische Autorenschaften]. *Information: Wissenschaft und Praxis*. 75, H. 2.
- 9 HEALTH AFFAIRES (2005), About us. About the Journal, *The Policy Journal of the Health Sphere*. URL: [http://www.healthaffairs.org/1500\\_about\\_journal.php](http://www.healthaffairs.org/1500_about_journal.php).
- 10 Hochschulverträge 2006 bis 2009 gemäß Artikel II § 1 Abs. 1 und 4 des Haushaltsstrukturgesetzes 1997. URL: [http://www.science.berlin.de/2\\_hochschulen/inhalt/2\\_hochpolitik/3\\_vertraege/8GesamtVertraegePublikation.pdf](http://www.science.berlin.de/2_hochschulen/inhalt/2_hochpolitik/3_vertraege/8GesamtVertraegePublikation.pdf).
- 11 MATTERN, F. (2002). Zur Evaluation der Informatik mittels bibliometrischer Analyse. *Informatik Spektrum* 25, 22-32. URL: <http://www.springerlink.com/media/1b22aprxwh0rpy2gnt2l/contributions/p/b/3/7/pb37nel8ymleednu.pdf>.
- 12 MURAL, R. J. [u. a.] (2002). A comparison of whole genome shotgun derived mouse chromosome 16 and the human genome. *Science*, 296, 1661-1617.
- 13 NAUMANN, U. (1997). Der Beitrag bibliometrischer Methoden zur Messung von Forschungsleistungen: kurzgefaßte Überlegungen zur Präzisierung des Leistungsindikators „Publikationen“ durch Berücksichtigung des „impact factor“. In *Hochschulinterne Steuerungsmodelle an der Freien Universität Berlin – Erfahrungen und Ausblick* (S. 35-38). Berlin: Freie Universität, Projekt Hochschulinterne Steuerungsmodelle.
- 14 PARLIAMENTLIVE.TV. Broadcasting live from Westminster. URL: <http://www.parliamentlive.tv>.
- 15 PRICE, D. J. DE SOLLA. (1963). *Little Science, Big Science*. New York, London. Columbia University Press.

- 16 PRICE, D. J. DE SOLLA. (1974). *Little Science, Big Science*. URL: <http://www.ib.hu-berlin.de/%7Ewumsta/price14.html>.
- 17 PRÖMEL, H. (2005). Editorial. *CMS Journal* 27, S. 1. URN: urn:nbn:de:kobv:11-10044097.
- 18 *Science*. 20. Januar 2006, Vol. 311. no. 5759, S. 335.
- 19 *Science Express*. 12. Januar 2006.
- 20 *Scientometrics. An international journal for all quantitative aspects of the science of science, communication in science and science policy / Akadémiai Kiadó*. Oxford: Elsevier Science [u. a.], 1.1978/79 -
- 21 SEITER, M. & STIRZEL, M. (2005). Messung von Forschungsleistungen: State-of-the-Art. *Wissenschaftsmanagement*, 11, 3, 25-29.
- 22 STOCK, W. G. (2001). Journal Citation Reports: Ein Impact Factor für Bibliotheken, Verlage und Autoren? *Password* 5, 24-39.
- 23 STOCK, W. G. (2001). *Publikation und Zitat, die problematische Basis empirischer Wissenschaftsforschung*. Köln: Fachhochschule Köln, Fachbereich Bibliotheks- und Informationswesen.
- 24 UMSTÄTTER, W. (2004). Szientometrie. In Rainer Kuhlen ... (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (S. 237-243), München: Saur.
- 25 UMSTÄTTER, W. (2005). *Bradford's Law of Scattering*. URL: <http://www.ib.hu-berlin.de/%7Ewumsta/infopub/pub2001f/Bradford05fold.pdf>.
- 26 UMSTÄTTER, W. & NOURMOHAMMADI, H. A. (2006). Über den jährlichen Anstieg des Impact Factors und der „Garfieldschen Konstante“. *Information: Wissenschaft & Praxis*, 57, 113-118.
- 27 Universitätsbibliographie, FREIE UNIVERSITÄT BERLIN. URL: <http://www.ub.fu-berlin.de/unibibliographie>.
- 28 VICIANO, A. (2006). Amerikanisches Klonopfer: Der US Mediziner Gerald Schatten hat maßgeblich am beispiellosen Forschungsbetrug des Koreaners Hwang mitgewirkt doch seine Universität in Pittsburgh spricht ihn frei von Schuld. *DIE ZEIT* 16.02.2006, 8.
- 29 WEINGART, P. (2006). Hwang und Rang: Die betrügende, betrogene Wissenschaft. *Süddeutsche Zeitung* 2.1.2006, 14.

- 30 ZIMMERMANN, K. F. (2005). *Standpunktbestimmung nach der Vorlage des Bewertungsberichts der Bewertungsgruppe der Leibniz-Gemeinschaft über das Deutsche Institut für Wirtschaftsforschung Berlin (DIW Berlin) vom 8. April 2005 für den Senatsausschuss Evaluierung der WGL 23. Mai 2005.* URL:  
[http://www.diw.de/deutsch/dasinstitut/info/2005/050523\\_diw\\_berlin\\_standpunktbestimmung.pdf](http://www.diw.de/deutsch/dasinstitut/info/2005/050523_diw_berlin_standpunktbestimmung.pdf).

Die Internetadressen wurden zuletzt am 10.07.2006 aufgerufen.