

Über Zitate zu Zensuren und Zaster

Immer stärker hängt die Karriere eines Wissenschaftlers von der Beachtung seiner Publikationen ab. Ob damit der Forscherleistung wirklich Genüge getan wird, ist fraglich.

Spanische Professoren erhalten Gehaltszuschläge, wenn sie Artikel in angesehenen Zeitschriften platzieren. Eine Reihe deutscher Fakultäten - beispielsweise die medizinischen der Universitäten Düsseldorf und Bochum - verteilt Forschungsgelder an Institute und Arbeitsgruppen nach deren Publikationsleistung. Finnland vergibt einen Teil der Zuschüsse für die Universitätskliniken danach, wie viele Veröffentlichungen die dortigen Wissenschaftler verfasst haben und wie hoch das Ansehen der Beiträge im Kollegenkreis ist. Im Slang der Wissenschaft heißt das „Impact“ (englisch für „Wirkung, Resonanz“). Impact-Faktoren geben an, wie häufig Artikel aus einer Zeitschrift im darauf folgenden Jahr in anderen Veröffentlichungen zitiert wurden. Im Jahr 2000 war es der finnischen Regierung jedes Mal 7000 US-Dollar wert, wenn ihre Mediziner einen Artikel in einer Zeitschrift mit dem Impact-Faktor 3 statt in einer mit dem Faktor 2 unterbrachten.

„Bibliometrische Daten und daraus erstellte Rankings gewinnen in den Industrieländern enorm an Bedeutung“, stellt Prof. Peter Weingart fest, Leiter des Instituts für Wissenschafts- und Technikforschung der Universität Bielefeld. Zwar haben diese auf der Zählung von Zitaten basierenden Bewertungen von Forschungsarbeiten nur selten unmittelbaren Einfluss auf die Zuweisung von Geldern. Doch über die Medien beeinflussen solche Ranglisten, die an die Fußballbundesliga-Tabelle erinnern, den Ruf von Forschungseinrichtungen durchaus. So veröffentlichte das Gütersloher Centrum für Hochschulentwicklung Ende 2003 in Kooperation mit der Zeitschrift „Stern“ Rankings der Universitäten, die zum Teil auf bibliometrischen Daten basieren.

Dass sich die Leistung von Wissenschaftlern in ihren Publikationen widerspiegelt, ist weitgehend unbestritten. Doch Veröffentlichungen sind natürlich nicht die einzige Messlatte: Gerade Geisteswissenschaftler profilieren sich auch in

Vorträgen und Debatten, für wirtschaftsnahe Forscher sind vor allem Patente wichtig. Klar ist auch, dass ein Wissenschaftler mit nur 30 Veröffentlichungen, wie der Physik-Nobelpreis-träger Richard Feynman (1918 bis 1988), nicht unbedingt weniger geleistet hat als einer mit 500 - auf so viele bringen es heute etliche Forscher.

Deshalb wurde die Zitationsanalyse entwickelt: Gezählt wird - meist mit Hilfe der Datenbank „Science Citation Index“ (SCI) -, wie häufig andere Forscher auf die jeweilige Veröffentlichung Bezug nehmen. Dem liegt die Annahme zugrunde, dass wesentliche Arbeiten besonders häufig zitiert werden. Allerdings: „Mit Zitationsanalysen lässt sich nicht unmittelbar die Qualität von Publikationen messen, sondern nur ihre Beachtung - das ist ein Unterschied, gibt es doch auch in der Forschung Moden“, sagt Weingart.

Zitate zu zählen, klingt einfach - und lässt ein exaktes Ergebnis erwarten. Doch in der Praxis gibt es viele Fallstricke: „Fehlerhafte oder mehrdeutige Zitate machen die größten Schwierigkeiten“, sagt Prof. Anthony van Raan von der Universität Leiden, einer der weltweit führenden Bibliometriker. So stellte sein Kollege Dr. Wolfgang Glänzel fest: Eine von ihm mitverfasste Veröffentlichung wurde bis September 2003 insgesamt 141-mal zitiert, davon 26-mal mit falschem Autorennamen, falschem Veröffentlichungsjahr oder einem anderen Fehler: das heißt, bei nahezu jedem fünften Zitat stimmte die Quellenangabe nicht. Mehrdeutige Institutionsbezeichnungen sind eine andere Fehlerquelle: So können sich hinter der „Freien Universität Brüssel“ zwei verschiedene Universitäten verbergen: die flämische „Vrije Universiteit Brüssel“ und die valonische „Universite Libre de Bruxelles“. Oder Kliniken veröffentlichen mal unter Namen wie „St. Marienhospital“ und mal unter „Universitätsklinik“.

Auch wenn die Zahl richtig ermittelt wurde, die Größenordnung der Zitate alleine hat noch keine große Aussagekraft. Beispielsweise werden Artikel zur biomedizinischen Grundlagenforschung rund sechsmal häufiger zitiert als solche aus der Mathematik. Ein Spitzenmathematiker hätte somit gegen einen mittelmäßigen Molekularbiologen im Zitationswettkampf keine Chance, weshalb der fachübergreifende Vergleich wertlos ist. Andererseits bringt auch die Einordnung von Wissenschaftlern und ihren Publikationen in Fachgebiete Probleme mit sich: Die Grenzziehung ist angesichts zunehmender Verquickung verschiedener Disziplinen schwer, beeinflusst aber die Ergebnisse der Zitationsanalyse.

Damit nicht genug: Der Science Citation Index - die Datenbank, auf der die meisten bibliometrischen Analysen weltweit beruhen - deckt englischsprachige Journale und die Lebenswissenschaften besser ab als deutschsprachige Zeitschriften und die Ingenieurwissenschaften. Außerdem hängt das Ergebnis der Analyse auch davon ab, ob nur echte Artikel oder auch Zuschriften dazu und Konferenzbeiträge einbezogen werden.

Dennoch argumentiert Anthony van Raan: „Trotz solcher Probleme wäre es falsch, wenn Politiker und Wissenschaftsmanager bibliometrische Daten überhaupt nicht nutzen würden. Denn nur mit guten bibliometrischen Analysen lassen sich die Urteile kontrollieren, die Gutachterkommissionen über die Leistung von Forschergruppen fällen.“

Eine wichtige Kontrolle: Denn Wissenschaftler sind nicht zwangsläufig objektiv, wenn sie als Gutachter Kollegen bewerten. Häufig erhobener Vorwurf: Konkurrenzdenken und Seilschaften würden die Begutachtung durch sachverständige Kollegen - das Peer-Review - beeinflussen. „Das Peer-Review ist sicher nicht fehlerfrei, hinzu kommt, dass die Kritik daran nur schwer an den Mann gebracht werden kann: Die Gutachter sind notwendigerweise oft anonym, um sie vor den Begutachteten in Schutz zu nehmen“, sagt der Hirnforscher Karl Zilles, Professor an der Universität Düsseldorf und Leiter des Instituts für Medizin im Forschungszentrum Jülich. Bewertungen, die aus bibliometrischen Analysen resultieren, sind dagegen überprüfbar - wenn auch nur mit viel Aufwand. „Außerdem lässt sich ein Wissenschaftsgebiet mit Hilfe der Bibliometrie aus einem größeren Blickwinkel betrachten, als ihn der einzelne Forscher oder Gutachter haben kann“, ist Weingart überzeugt.

Wird die Analyse nach allen Regeln der bibliometrischen Kunst durchgezogen, ist sie sogar gegen Strategien immun, mit denen Wissenschaftler versuchen, ihre bibliometrischen Werte zu beeinflussen. Dazu gehört, sich in Veröffentlichungen unnötig häufig auf eigene ältere Publikationen zu berufen. „Daneben gibt es eine Minderheit von Wissenschaftlern, die sich

vorwiegend untereinander zitieren, obwohl es wichtige, manchmal bessere Ergebnisse außerhalb der Zirkels gibt“, weiß Prof. Werner Bergholz, Professor of Electrical Engineering an der International University Bremen. Ob solche Kartelle immer bewusst aufgebaut werden, bezweifelt Zilles allerdings: „Manchmal steckt einfach Bequemlichkeit dahinter: Man kennt die Arbeiten der Kollegen und kann sie ohne große Recherche zitieren.“ Selbstzitationen und Zitationskartelle können bei der bibliometrischen Analyse erkannt und die Ergebnisse korrigiert werden, erklären van Raan und Weingart unisono. Auch die Strategie von Wissenschaftlern, vorwiegend so genannte Reviews zu schreiben, verfängt bei einer richtig zugeschnittenen Analyse nicht. Reviews geben einen Überblick über ein Forschungsgebiet, enthalten aber keine originären Ergebnisse. Dennoch werden gerade sie besonders gerne zitiert.

Manche Wissenschaftler lehnen Qualitätsurteile über Forschung, die auf bibliometrischen Zahlen beruhen, bis heute kategorisch ab. Nach Ansicht des Frankfurter Sportmediziners Prof. Gerd Hoffmann lässt sich wissenschaftliche Qualität nicht durch Zählstatistiken erfassen. Erforderlich sei die Bewertung des Inhalts von Publikationen. „Wenn man wissen will, wie kalt es ist, taugt ein Höhenmesser nicht als Messinstrument - auch wenn es einen schwachen Zusammenhang zwischen der Höhe und sinkenden Temperaturen gibt“, betont Hoffmann.

Meist richtet sich die Kritik freilich nicht gegen ausgefeilte bibliometrische Analysen, die zudem noch mit dem Peer-Review kombiniert werden, sondern gegen simple Verfahren zur Messung der Publikationsleistung. „Wenn es um die Berufung eines Professors geht, wird leider meist nur gezählt, wie viele Artikel er geschrieben hat, wie viele davon als Erstautor und wie viele pro Zeitschrift“, berichtet Mathematiker Dr. Patrick Bangert von der International University Bremen. Ein zweifellos fragwürdiger Leistungsmaßstab - allein schon deshalb, weil diese Methode jene begünstigt, die zusammengehörige Forschungsergebnisse auf mehrere Publikationen verteilen oder die gleichen Ergebnisse in mehreren Zeitschriften unterbringen.

Besonders Mediziner sind versucht, die Messlatte auf einfache Weise zu verbessern: Lehrstuhlkandidaten bewerben sich häufig mit dem „kumulierten Impact-Faktor“, der sich aus der Zahl der Veröffentlichungen und aus den jeweiligen Impact-Faktoren der Zeitschriften ergibt, in denen diese erschienen sind. So offenbaren die Bewerbungsvorschriften der Uni Lübeck, dass Universitäten die Bewerbung mit den kumulierten Impact-Faktoren standardmäßig verlangen. Medizinische Fakultäten verteilen ihre Forschungsgelder oft nach ähnlichen Berechnungen. Dabei wird der kumulierte Impact-Faktor der verschiedenen Arbeitsgruppen ermittelt: Wer den höchsten hat, bekommt das meiste Geld - allerdings wird wohl nur ein kleiner Teil des Gesamtbudgets nach diesem Verfahren verteilt. „Quick and dirty“ - schnell und schmutzig - nennt van Raan diese Vorgehensweise. Mediziner Zilles charakterisiert sie als „Missbrauch von Impact-Faktoren“. Hauptkritikpunkte: Natürlich veröffentlichen auch Zeitschriften mit hohem Impact-Faktor Beiträge, die in der Fachwelt nicht auf Resonanz stoßen. Hinzu kommt, dass Bücher bei dieser Methode unter den Tisch fallen. Lehrstuhlbewerber bezeichnen sie deshalb gerne als Jobkiller, denn sie haben keinen Impact-Faktor.

Wie wichtig Impact-Faktoren mittlerweile sind, lässt sich daran ablesen, dass die jeweiligen Zeitschriften nicht nur mit ihnen werben, sondern sie manchmal sogar manipulieren wollen. So hat der Bielefelder Bibliometriker Weingart einen Versuch der Herausgeber der etablierten medizinischen Zeitschrift SHOCK öffentlich gemacht. Sie wollten einen Autor dazu bringen, in eine eingereichte Veröffentlichung besonders viele auf SHOCK verweisende Zitate einzubauen.

Weingart und seine Kollegen sehen die Gefahr, dass Universitätsdekane, Klinikchefs, Forschungszentren-Vorstände und Politiker die Objektivität bibliometrischer Werte überschätzen und nicht berücksichtigen, wie sie zustande gekommen sind. So warnt Weingart davor, bibliometrische Werte laienhaft zu bestimmen. Dies werde aber gerne getan, weil man die Kosten einer fachmännischen Analyse scheue, die 20- bis 30-mal teurer ist. Was solche Analysen kosten, offenbart der Leidener Bibliometriker van Raan: Für rund 170 000 Euro hat er im Sommer dieses Jahres acht medizinische Fakultäten in den Niederlanden unter die Lupe genommen, wobei die jeweiligen Universitäten vollständige Publikationslisten lieferten. Einzelne Forscher nach ihrer Publikationsleistung einzuordnen, lehnen Weingart und van Raan ab: Zu klein seien die Unterschiede, die darüber entscheiden, ob ein Wissenschaftler Platz 1 oder Platz 5 einnimmt - und zu groß die Fehlermöglichkeiten. Anders sehe das Ranking von Forschungseinrichtungen aus. Dort flößen vergleichsweise viele Veröffentlichungen und Zitate in die Bewertung ein, wodurch sich statistische Unwägbarkeiten nivellieren würden.

Die Forscher, die für diesen bild-der-wissenschaft-Beitrag direkt auf die bisher veröffentlichten Rankings in den Medien angesprochen wurden, urteilten fast einhellig: „Schwachsinn“. „Oft ist schon die zugrunde liegende Fragestellung unsinnig. So kann es die beste Universität nicht geben, allerhöchstens den besten Fachbereich“, urteilt Weingart. Vielfältig sind die Vorschläge, was zu tun ist, damit bibliometrische Werte von Medien, Politikern, Wissenschaftsmanagern und Wissenschaftlern nicht falsch gebraucht werden. „Die Bibliometriker sollten ein knappes Heftchen mit einfachen Erklärungen und klar formulierten Empfehlungen herausbringen und überall verteilen“, rät Zilles. „Außerdem sollten sie nicht selbstverliebt ihre methodischen Werkzeuge verbessern, sondern die Anwendung der Bibliometrie in der Praxis zum Forschungsgegenstand machen.“

Van Raan glaubt, Zeitschriften und Zeitungen den Verzicht auf die nur scheinbar genauen Ranglisten schmackhaft machen zu können. Seine Empfehlung: Universitäten und andere Forschungseinrichtungen sollten lediglich in die drei Kategorien top, gut und mittelmäßig eingeordnet werden. Und die Physikerin Hildegard Meyer-Ortmanns, Professorin in Bremen, hat die Berufungspraxis im Blick: „Ein Teil der Lehrstühle, beispielsweise zehn Prozent, sollte an Kandidaten vergeben werden, die durch besondere wissenschaftliche Originalität in ihren Veröffentlichungen aufgefallen sind - unabhängig von der Zahl ihrer Publikationen oder von anderen bibliometrischen Werten.“ So bekämen auch diejenigen Wissenschaftler eine Chance, die ihre Forschungsthemen unabhängig von Erfolgsaussichten und Modeströmungen wählen.

Ein Ende der Diskussion um die Bibliometrie ist mangels besserer Alternativen vorerst nicht in Sicht. „Konstruktive Kritik an der Bibliometrie fällt schwer“, stellt denn auch Mathematiker Bangert fest. Mutig, wenn es jemand trotzdem versucht: LUB-Professor Bergholz, zuvor beim Chiphersteller Infineon in einem der Produktionsbetriebe verantwortlich für „Quality Management and Business Excellence“, hat die Idee, zwei Datenbanken aufzubauen. In die eine stellt jeder Forscher seine neuen Publikationen und gibt ihnen gleichzeitig selbst mehrere Noten - für die wissenschaftliche, wirtschaftliche und gesellschaftliche Bedeutung der Ergebnisse. Andere Wissenschaftler können Kommentare dazu eingeben. In einer parallelen Datenbank benotet sich der Autor später erneut und führt für seine Bewertung jeweils aktuelle Belege an - etwa Arbeiten, die sich auf seine Veröffentlichung beziehen.

Berufungskommissionen oder die Geber von Forschungsgeldern könnten die Datenbanken so abfragen, dass der Computer die bestbenoteten Wissenschaftler für ein Spezialgebiet ausspuckt.

Dass darunter möglicherweise Aufschneider sind, würde nicht stören: Anhand der Differenzen zwischen ursprünglicher und späterer Benotung und der Kommentare der anderen Wissenschaftler wären sie vergleichsweise schnell zu enttarnen. „Dieses System käme weitgehend ohne Fremdbeurteilung aus, die aufwendig und anfällig für Missbrauch ist“, so Bergholz. Warten wir doch einfach ab, welche Kommentare und Noten dieser Vorschlag demnächst von anderen Wissenschaftler bekommt.

DAS GEHEIMNIS DES IMPACT-FAKTORS

NEHMEN WIR AN, eine Zeitschrift hat 150 Originalpublikationen in den Jahren 2001 und 2002 abgedruckt. Und diese Publikationen wurden 2003 weltweit 300-mal zitiert. Dann hat die Zeitschrift im Jahr 2003 einen Impact-Faktor von 2,0 (300/150). Bei 450 Nennungen in anderen Zeitschriften hätte sie einen Impact-Faktor von 3,0.

Einen Faktor von 2,482 erzielte 2003 beispielsweise die Zeitschrift „European Journal of Inorganic Chemistry“. Kein schlechter Wert: Zahlreiche Fachzeitschriften, darunter auch solche der Chemie, haben einen Faktor unten. Andererseits kam die „Angewandte Chemie“ im letzten Jahr auf 8,427. Unter den Top 100 der wissenschaftlichen Zeitschriften ist sie damit noch nicht. Der derzeitige Impact-Spitzenreiter kommt auf den Faktor 52,28. Sein Name ist nicht „Nature“ (Rang 8) oder „Science“ (Rang 11), wie man vermuten könnte, sondern „Annual Review of Immunology“.