

Alle Macht den Zahlen! Zur Soziologie des Zitationsindexes

Von Richard Münch

Zusammenfassung: Zitationsindizes sind wie alle quantitativen Indikatoren höchst selektiv und werden der Vielfalt wissenschaftlicher Forschung nicht gerecht. Durch ihre Reaktivität erzeugen sie eine eigene, die wissenschaftliche Praxis bestimmende Wirklichkeit mit der fatalen Konsequenz der Einengung der Wissensevolution auf ein schmales Spektrum möglichen Wissens. Das soll in diesem Essay dargelegt werden.

Die Nutzung quantitativer Indikatoren bei der Evaluation wissenschaftlicher Forschung scheint nicht mehr aufzuhalten zu sein (Hornbostel 1997). Das gilt auch für Zitationsindizes wie den Journal Impact Faktor zur Messung des Einflusses einer Fachzeitschrift auf die wissenschaftliche Praxis oder den h-Index zur Messung des Einflusses eines einzelnen Wissenschaftlers oder einer einzelnen Wissenschaftlerin. Trotzdem stellt sich der Nutzung solcher Indizes für die Evaluation wissenschaftlicher Forschung eine wachsende Front der Kritik entgegen. Im Folgenden sollen drei Elemente dieser Kritik expliziert werden: (1) die Selektivität von Zitationsindizes, (2) ihre Reaktivität und (3) ihre Fatalität in der Einengung der Erkenntnissuche auf den sicheren Weg zum Zitat.

Selektivität

Zum irreführenden Charakter von Zitationsindizes und zu deren negativen Effekten auf den Erkenntnisfortschritt ist nahezu alles gesagt. Im Vordergrund steht dabei deren der Vielfalt wissenschaftlicher Leistungen nicht gerecht werdende Selektivität. Eine Reihe wissenschaftlicher Vereinigungen hat mit allem Nachdruck davor gewarnt, Zitationsindizes als „objektives“ Maß der Qualität wissenschaftlicher Forschung zu betrachten. Ein 2009 veröffentlichter Bericht der International Mathematical Union verweist darauf, dass Statistiken irreführenden Charakter haben können, wenn die Hintergründe ihrer Erstellung nicht reflektiert werden, dass die „Objektivität“ von Zahlen eine Illusion sein kann, dass das alleinige Verlassen auf Zitationsdaten zu keinem angemessenen Verständnis der zu bewertenden Forschung führt (Adler et al. 2009).

Eine zentrale Rolle kommt inzwischen dem jährlich veröffentlichten Journal Impact Faktor zu, der 1975 vom Institute for Scientific Information (ISI) mit dem Journal Citation Report (JCR) als Ergänzung zum Science Citation Index (SCI) eingeführt wurde (Fleck 2013). Das seit 1992 von Thomson Reuters betriebene Institut wurde 1960 von Eugene Garfield zu dem Zweck gegründet, das unmittelbar nach dem Zweiten Weltkrieg schon als exponentiell identifizierte Wachstum wissenschaftlicher Publikationen insbesondere für die Akquisitionstätigkeit der Bibliothekare überschaubar zu machen. Das Institut verfügt mit dem Web of Science (WoS), das den Science Citation Index (SCI), den Social Science Citation Index (SSCI) und den Arts and Humanities Citation Index (AHC) umfasst, über ein Informationsmonopol, auf das insbesondere Rankings zurückgreifen. Der Journal Impact Faktor eines Jahres gibt an, wie häufig die in den zwei Jahren zuvor in einer Fachzeitschrift erschienenen Artikel in diesem Jahr in den im Web of Science registrierten Fachzeitschriften im Durchschnitt zitiert wurden. Dabei wird die Gesamtzahl der in diesem Jahr von den in den zwei vorausgegangenen Jahren erschienenen Artikeln eines Journals erreichten Zitate durch die Zahl der in den vergangenen zwei Jahren in dem Journal erschienenen Artikel dividiert. Natürlich haben die zentralen Zeitschriften eines Faches die höchsten Impact-Faktoren, während die Spezialzeitschriften randständiger Fachgebiete nur auf einen sehr niedri-

gen Impact Faktor kommen. Den höchsten Impact Faktor haben amerikanische Kernzeitschriften, schon wegen des größten Publikationsmarktes mit weltweiter Ausstrahlung und der hegemonialen Stellung der US-Wissenschaft. So lag der Impact Factor der *American Sociological Review* im Jahr 2013 bei 4.266, des *British Journal of Sociology* bei 1.62 und der *Kölner Zeitschrift für Soziologie und Sozialpsychologie* bei 0.51. Die Qualität einer Fachzeitschrift anhand ihres Impact Factors zu messen, wird von der International Mathematical Union als ein viel zu einfaches Instrument zurückgewiesen, weil dadurch ein viel zu geringes Maß an Informationen erfasst wird, die man benötigt, um die Qualität einer Fachzeitschrift beurteilen zu können. Das gilt erst recht für die häufig vollzogene Schließung vom Impact Faktor einer Fachzeitschrift auf die Qualität eines einzelnen darin erschienenen Aufsatzes. Allein schon dessen Zitationen können stark vom Durchschnittswert aller in der Fachzeitschrift in einem Zweijahreszeitraum publizierten Aufsätze abweichen, und zwar viel häufiger nach unten als nach oben, weil der Durchschnittswert in der Regel von wenigen viel zitierten Aufsätzen nach oben gezogen wird.

Auch die unmittelbare Zahl der Zitationen eines Aufsatzes kann nicht ungeprüft als Maß für dessen Qualität oder Originalität verwendet werden. Erst recht geht das nicht im Vergleich zwischen Aufsätzen, weil Faktoren die Zitationshäufigkeit eines Aufsatzes beeinflussen, die nicht unmittelbar etwas mit dessen Qualität oder Originalität zu tun haben. So gibt es große Unterschiede in der Häufigkeit von Zitationen zwischen den wissenschaftlichen Disziplinen und zwischen den Teilgebieten der Disziplinen. Mit der Zahl der Autoren eines Aufsatzes wachsen die Zitationschancen, weil sich die Netzwerke der Autoren summieren. Die schiere Größe und die Zentralität des amerikanischen Publikationsmarktes verschaffen Publikationen in amerikanischen Fachzeitschriften oder führenden amerikanischen und englischen Universitätsverlagen einen großen Vorsprung in den Zitationschancen. Das gilt für amerikanische Autoren in diesen Publikationsorganen weit mehr als für nicht-amerikanische, weil diese nochmals von der festen Inklusion in amerikanische Netzwerke und von der Zugehörigkeit zu Universitäten profitieren, die in den – wie die Zitationsindizes viel kritisierten – internationalen Rankings die vorderen Rangplätze besetzen. Weiterhin ist der Matthäus-Effekt in der Erzeugung wachsender Ungleichheit in Zitationen wirksam (Merton 1968 a). Ein neuer Aufsatz eines vielzitierten Autors aus einem großen Netzwerk von Kollegen und Kolleginnen, einem zentralen Land mit großem Publikationsmarkt und einer in internationalen Rankings weithin sichtbaren Universität hat ganz unabhängig von seiner Qualität oder Originalität eine vielfach höhere Zitationschance als jeder andere Aufsatz, der einer weniger zentralen Quelle entstammt, und sei er bei genauer Betrachtung noch so gut und/oder originell.

Auch Zitationsindizes, die der „Qualitätsmessung“ einzelner Wissenschaftler oder Wissenschaftlerinnen dienen sollen, werden in dem Bericht der Internationalen Mathematischen Union als zu große Vereinfachung in der Beurteilung der Qualität oder Originalität der Forschungsleistungen einzelner Wissenschaftler gesehen. Sehr häufig verwendet wird der von dem argentinischen Physiker Jorge E. Hirsch (2005) entwickelte h-Index. Die Formel dazu lautet: Sind h Publikationen eines Wissenschaftlers von seinen insgesamt N Publikationen mindestens h -mal zitiert und alle seine restlichen Publikationen ($N-h$) höchstens h -mal, dann hat er einen h-Index von h . Die Zitationshäufigkeiten der 10 Publikationen eines Wissenschaftlers A können sich in absteigender Zahl z.B. wie folgt darstellen: 78, 61, 52, 15, 3, 3, 2, 2, 0, 0. Der Wissenschaftler hat einen h-Index von 4, weil vier Publikationen mindestens viermal, alle weiteren höchstens viermal zitiert wurden. Wie man an dem Beispiel erkennen kann, prämiert der Index eine relativ hohe Zahl von Publikationen, die eine gute Zitationshäufigkeit erreichen, und kappt die Spitzen. Die entscheidende Zahl ist 15 an vierter Stelle, wodurch der h-Index 4 erreicht wird. Die davor liegenden Zitationen könnten auch weit höher liegen oder auch deutlich niedriger, trotzdem wäre der h-Indexwert 4. Ein h-Index von

4 wird demnach im Minimum auch mit den folgenden Werten eines Wissenschaftlers erzielt: 15, 12, 11, 10, 3, 3, 2, 2, 0, 0. Für die Berechnung des h-Indexes hat der erste Wissenschaftler demnach eine Menge überflüssiger Zitationen in der Spalte. Die Rezeption, die er mit seinen Spitzenergebnissen erfahren hat, kommt im h-Index nicht zum Ausdruck. Man sieht daran, dass es sich um einen Index handelt, der keineswegs die Einzelbetrachtung ersetzen kann. Der Index steigt außerdem kontinuierlich mit dem Alter eines Autors, sodass Vergleiche nur innerhalb einer Alterskohorte möglich sind. Innerhalb einer Kohorte machen sich aber wieder die Besonderheiten von Disziplinen und disziplinären Teilgebieten bemerkbar. Insgesamt zeigen sich an allen Ecken und Enden Unzulänglichkeiten, sodass der Index keineswegs geeignet ist, als Maß für die Beurteilung der Qualität oder Originalität eines Wissenschaftlers zu dienen.

Der h-Index begünstigt massiv Publikationsstrategien, die auf viele kleine Aufsätze mit einer Vielzahl von Autoren in Fachzeitschriften zielen und von einem weitreichenden Netzwerk von in regelmäßigen Austausch stehenden Wissenschaftlern getragen werden. In den Sozialwissenschaften ist das z.B. typisch für die quantitative empirische Sozialforschung, die sich auf der Basis leicht verfügbarer und auswertbarer Massendaten mit Sozialstruktur, Lebenslauf, Familie, Bildung oder Arbeitsmarkt beschäftigt. Dagegen ist es völlig untypisch für die Arbeit an theoretischen, historisch-vergleichenden oder zeitdiagnostischen Fragen, die mehrere Jahre dauert, bis sie in die Publikation einer Monographie mündet, mit der man nicht mehr Punkte im h-Index sammeln kann als mit einem zehnseitigen Aufsatz. Es gilt auch für die Arbeit an allen empirischen Fragen, für die es keine von einer Vielzahl von Autoren verwendeten Massendaten gibt. Man könnte diesem Umstand dadurch Rechnung tragen, dass Monographien das zehnfache Gewicht eines Aufsatzes gegeben wird oder empirischen Arbeiten ohne Massendaten das dreifache Gewicht eines Aufsatzes mit Massendaten. Außerdem könnte konsequent nach der Zahl der Autoren einer Publikation relativiert werden. Das würde z.B. der massiven Begünstigung von kollaborativ produzierten Aufsätzen aus der quantitativen empirischen Sozialforschung mit Massendaten ein wenig entgegenwirken.

Wegen der Unzulänglichkeiten des h-Indexes sind eine Reihe anderer Indizes entwickelt worden, die allerdings noch nicht dieselbe Beachtung gefunden haben (Bornmann / Daniel 2005; Bornmann et al. 2008, 2010). So hat z.B. Leo Egghe (2006) einen g-Index entwickelt, bei dem hochzitierte Artikel zu Buche schlagen. Ordnet man die Artikel eines Autors absteigend nach ihrer Zitationshäufigkeit, dann ist der g-Index die höchste Zahl, bei der die top-g-Artikel zusammen mindestens g^2 Zitationen erreichen. Die höher zitierten Artikel eines Autors bekommen dadurch ein höheres Gewicht und ziehen den g-Index im Vergleich zum h-Index nach oben. Der oben erstgenannte Wissenschaftler A mit höher zitierten Artikeln in der Spalte als der zweitgenannte Wissenschaftler B schneidet dementsprechend beim g-Index besser ab. Die zehn Publikationen von Wissenschaftler A erreichen in der Summe 216 Zitate. Das ist mindestens so viel wie das Quadrat der Anzahl der zehn Artikel, d.h. 100. Wissenschaftler A hat damit einen g-Index von 10. Die zehn Publikationen von Wissenschaftler B haben in der Summe nur 58 Zitationen erhalten. Die Anzahl 7 der ersten sieben Artikel ergibt im Quadrat 49, was noch unterhalb 58 liegt. Bei acht kämen wir auf 64, was über 58 hinausgehen würde. Also hat Wissenschaftler B einen g-Index von 7. In diese Richtung geht auch die von Google Scholar errechnete Gesamtzahl an Zitationen. Sie bildet hochzitierte Publikationen in vollem Umfang ab. Genau in die entgegengesetzte Richtung weist der ebenfalls von Google Scholar ausgewiesene i10-Index. Es handelt sich dabei schlicht um die Zahl aller Publikationen eines Autors, die mindestens zehnmal zitiert wurden. Ein Versuch, die Beurteilungsbasis der Forschungsleistungen eines Wissenschaftlers auf eine breitere Basis als Zitationen in Fachzeitschriften zu stellen, ist Altmetrics. Dabei

werden Sichtung, Downloads und Nennung in Social Media berücksichtigt (Priem et al. 2010).

Natürlich haben alle derartigen Versuche, einen Index zu bilden, der die Nachteile des h-Indexes vermeidet, wieder ihre eigenen Nachteile. Sie rücken immer nur eine von mehreren Seiten des komplexen Prozesses der Erkenntnisproduktion ins Rampenlicht und stellen andere, nicht minder wichtige Aspekte in den Schatten, sodass sie nicht mehr gesehen werden. Dieser Mangel der Messung von Forschungsleistungen mit Hilfe von Indizes ist grundsätzlich nicht zu beseitigen (Quack 2015). Das ist der Grund, weshalb nicht nur die Internationale Mathematische Union empfiehlt, sich bei der Beurteilung von Forschungsleistungen nicht auf solche Indizes zu verlassen. Dazu zählen u.a. das International Council for Science Committee on Freedom and Responsibility in the Conduct of Science (ICSU 2008) und die Deutsche Forschungsgemeinschaft (DFG 2010). Im Dezember 2012 hat eine Initiativgruppe auf der Tagung der American Society for Cell Biology die San Francisco Declaration on Research Assessment verabschiedet, die empfiehlt, den Journal Impact Faktor nicht zur Beurteilung von Forschungsqualität und als Grundlage von Entscheidungen bei Berufungen, Promotionen und der Bewilligung von Förderanträgen zu verwenden. Im Vordergrund steht dabei die unzureichende Abbildung der Vielfalt von Forschungsleistungen durch Indizes (Bornmann / Daniel 2008; Welpe et al. 2015). Die San Francisco Declaration hat 155 individuelle und 82 institutionelle Erstunterzeichner, weitere 12055 individuelle und 547 institutionelle Unterzeichner haben sich bis Januar 2015 angeschlossen (DORA 2012).

Reaktivität

Über diesen methodisch grundsätzlich nicht zu beseitigenden Mangel der unangemessenen Selektivität hinaus gilt es aus soziologischer Perspektive, auch die enorm verändernde Wirkung der Indizes auf die wissenschaftliche Praxis in den Blick zu nehmen. Hier geht es nicht mehr nur um Unzulänglichkeiten, sondern um einen massiven Schaden, den der Einsatz derartiger Indizes verursacht. Der entscheidende soziale Mechanismus, der diesen Schaden hervorruft, ist die Reaktivität von Indikatoren. Der amerikanische Sozialpsychologe Donald Campbell (1957: 298) hat diesen Mechanismus auf den Punkt gebracht. Demnach können Indikatoren, die ein Verhalten messen sollen, einen direkten Einfluss auf das Verhalten ausüben. Das gilt umso mehr, je mehr Belohnungen an den Indikator geknüpft sind (Campbell 1976: 54).

Es handelt sich hier um eine Spezifikation des berühmten Thomas-Theorems, das sagt, dass Situationen real wirken, wenn sie als real definiert werden (Thomas / Thomas 1928: 572). Es wird in diesem Zusammenhang deutlich, dass Statistiken Realität nicht einfach abbilden, sondern konstruieren. Je mehr wir uns an ihnen orientieren, umso mehr wird die statistisch konstruierte Realität zur dominanten soziale Praxis bestimmenden Realität (Porter 1995; Desrosières 2008 a, 2008 b). Wendy Espeland und Michael Sauder (2007) haben den Mechanismus der Reaktivität sehr prägnant mit einer Studie zum Ranking amerikanischer Law Schools durch *US News and World Report* demonstriert. Viele beklagen sich über die unzulängliche Abbildung der Leistungen der Law Schools durch das Ranking und über dessen verheerende Wirkungen, gerade auch die Deans, die sich dem Druck ausgesetzt sehen, mit allen verfügbaren, mehr oder weniger legitimen Mitteln die für das Halten oder die Verbesserung der Position im Ranking erforderlichen Zahlen zu liefern.

Die Protagonisten derartiger Rankings sehen darin lediglich bestätigt, dass sie mit Rankings eine positive Wirkung erzeugen; alle strengen sich mehr an und erbringen bessere Leistungen. Wer sich in der Spitzengruppe befindet, lässt nicht nach, um die Position zu halten, die mittlere Gruppe kämpft um kleine Verbesserungen, die untere Gruppe sieht an einzelnen Fällen, wie man sich ein wenig verbessern kann. Am Ende werden alle von Jahr zu

Jahr etwas besser, so wie Usain Bolt einen Schwanz an Sprintern nach sich zieht, die zwar niemals seinen Weltrekord brechen werden, aber anscheinend glücklich sind, wenn sie ihre persönliche Bestmarke von 9,99 auf 9,98 oder von 10,09 auf 10,08 verbessern können. Genau so stellen sich die Protagonisten der indikatorgesteuerten Forschung die ständige Verbesserung der Forschungsleistungen von der Spitze über die Mitte bis zu den niederen Rängen vor. Das Problem dabei ist nur, dass sich Forschungsleistungen nicht so leicht messen lassen wie die Zeit, die für einen Sprint über 100 m benötigt wird. Genau das suggerieren jedoch Indizes. Die Folge ihres Einsatzes ist dann der typische Effekt der übermäßigen und einseitigen Produktion der prämierten Dinge, während in allen anderen Dingen großer Mangel herrscht, wie wir es von der sozialistischen Zentralverwaltungswirtschaft kennen. In unserem Fall können es z.B. massenhaft Artikel sein, die uns auf der Basis leicht verfügbarer Massendaten immer wieder dieselben Botschaften übermitteln, z.B. dass die soziale Herkunft maßgeblich den Schulerfolg bestimmt oder Partnerschaften von Berufspendlern weniger stabil sind als Partnerschaften von Menschen, die zusammen an einem Ort leben, oder der Berufseinstieg umso besser gelingt, je mehr Ressourcen wie Bildungstitel, Praktika und Netzwerke auf der Seite der Berufsanfänger und je weniger diskriminierende Vorurteile auf der Seite der Arbeitgeber vorhanden sind. Es ist grundsätzlich ein Fortschritt, dass die Soziologie über solche Zusammenhänge auf guter Datenbasis methodisch gesicherte Aussagen machen kann. Diese Forschung hat ihren berechtigten Platz in der Soziologie. Wenn sie aber so dominant wird, dass andere Formen der Soziologie in den Fachzeitschriften keinen Platz mehr finden, dann ist das doch eine deutliche Einschränkung von Diversität. Wir finden diese Einschränkung typischerweise in den Fachzeitschriften mit höherem Impact Faktor (*Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *Zeitschrift für Soziologie*) mehr als in den Fachzeitschrift mit niedrigerem Impact Faktor (*Soziale Welt*, *Berliner Journal für Soziologie*). Dieser Zusammenhang erklärt sich dadurch, dass Fachzeitschriften, die sich im Zuge der alleinigen Auswahl der veröffentlichten Artikel durch ein rigoroses Peer Review-Verfahren zwangsläufig auf das dominante Geschäft der Auswertung von Massendaten konzentrieren, auch von den höheren Zitationsquoten dieser Art von Forschung profitieren. Die vergleichsweise höheren Zitationsquoten erzielt diese Forschung, weil sie mit mehr Ko-Autoren arbeitet, weitreichender vernetzt ist und die breit geteilte Datenbasis die wechselseitige Bezugnahme auf Forschungsergebnisse erleichtert. Diese Forschung hat eine größere Affinität zu den Erfolgsbedingungen des Social Science Citation Index als Forschung singulärer und damit oft originellerer Art ohne breit geteilte Massendaten. Die Fachzeitschriften, die noch größere Diversität repräsentieren, veröffentlichen mehr singuläre und weniger breit vernetzte Forschung, bezahlen das aber unter dem Regime des Impact Faktors mit der Herabstufung in die Nachrangigkeit.

Die Daten des Sozioökonomischen Panels (SOEP) sind ein besonders ergiebiges Arsenal, aus dem sich in Deutschland mit gutem Computereinsatz massenhaft Artikel der standardisierten quantitativen Sozialforschung produzieren lassen. Sie füllen in zunehmendem Maße die Fachzeitschriften der Soziologie mit dem höheren Impact Faktor. Alber, Fliegner und Nerlich (2009) konnten deshalb feststellen, dass sich unter den mit dem Thyssen-Preis für die besten Aufsätze eines Jahrgangs der deutschsprachigen Fachzeitschriften für Sozialwissenschaften prämierten Aufsätzen ein eindeutiger Trend zum quantitativen Artikel zeigt. Beim seit 1996 durchgeführten Research Assessment Exercise in Großbritannien, das über die Vergabe von Forschungsmitteln bestimmt, ist eine stetig wachsende Zahl der zur Evaluation eingereichten Publikationen in einer höher gerankten Fachzeitschrift erschienen, obwohl auch dort die offizielle Verlautbarung sagt, dass es nicht auf den Impact des Publikationsorgans, sondern auf die eigene Qualität der Publikation ankommt. Das zeigt sich sogar in einem noch recht wenig standardisierten Fach wie der Geschichte (Hamann 2014).

Die Protagonisten des Indikatorenregimes sehen in dieser Entwicklung hin zum standardisierten Produkt einen Fortschritt der Qualitätssicherung, eine Objektivierung der Evaluation von Forschungsergebnissen und eine Annäherung an das von ihnen in der Regel hochgehaltene Ideal der „Einheitswissenschaft“, die über einheitliche methodische Standards verfügt und die Beurteilung der Qualität einer bei einer Fachzeitschrift eingereichten Publikation unabhängig von den jeweiligen Gutachtern bzw. Gutachterinnen macht. Das wird auch für alle sekundären Evaluationsverfahren, für Promotionen, Berufungen und die Bewilligung von Förderanträgen angenommen. Als Schreckgespenst auf der Gegenseite wird gern darauf verwiesen, dass ohne verbindliche Indikatoren in der Vergangenheit die Herrschaft der Patronen und ihrer idiosynkratischen Präferenzen praktiziert wurde. Das trifft sicherlich auf die deutsche oder französische Tradition der feudalen Lehrstuhlstruktur und der Platzierung des wissenschaftlichen Nachwuchses auf Professorenstellen durch den Einfluss des „Doktor-“, oder „Habitationsvaters“ zu (Bourdieu 1992: 149-180). Trotzdem ist der Anspruch der Objektivierung durch die Orientierung an verbindlichen Indikatoren wegen der weithin festgestellten zu weitgehenden Simplifizierung der Qualitäts- bzw. Originalitätsprüfung durch Indikatorenmodelle gar nicht einlösbar.

Fatalität

Der Grad der Übereinstimmung von Gutachtern bei der Beurteilung von Forschungsleistungen ist nicht groß genug, um von einer Objektivierung sprechen zu können, am wenigsten ist das bei den Geistes- und Sozialwissenschaften der Fall. Der Trend zum quantitativen Artikel, der Massendaten auswertet, ist kein Zeichen für zunehmenden Konsens, sondern ein Ausdruck der Einigung auf dem kleinsten gemeinsamen Nenner. Entsprechend einfach gehalten sind oft die in diesen Aufsätzen behandelten Fragen und entsprechend informationsarm sind die Ergebnisse. Für die Wissenschaft ist das fatal, weil es ihr Erkenntnispotential schrumpfen lässt. Es setzt sich das durch, was Michael Burawoy (2005) in seiner Presidential Address bei der Jahrestagung 2004 der American Sociological Association als „Professional Sociology“ bezeichnet hat, eine Soziologie, die den höchsten methodischen Standards entspricht, aber so nichtssagend geworden ist, dass sie nur noch interessant für die Gemeinde der Artikelproduzenten selbst ist. Dagegen lassen sich die von Burawoy für mindestens ebenso wichtig gehaltenen Varianten der Policy-oriented Sociology, der Critical Sociology und der Public Sociology nicht in das Standardformat eines im Web of Science von Thomson Reuters gelisteten Fachzeitschriftenaufsatzes zwingen. Außerhalb der Gemeinde der „Professionellen“ Soziologie würde aber eine Einschränkung der Soziologie auf genau diese Publikationssorte unisono als nicht tragbare Einschränkung des Erkenntnis- und Erklärungspotentials der Soziologie beklagt. Trotzdem sind wir im Fahrwasser des Indikatorenregimes auf dem besten Wege, genau diese Einschränkung vorzunehmen. Dass das selbst beim besten Vorsatz der Multidimensionalität in der Bewertung von Forschungsleistungen geschieht, hat das Forschungsrating Soziologie des Wissenschaftsrats in aller Deutlichkeit gezeigt. Der begutachtete Fachzeitschriftenaufsatz hat sich als dominanter Maßstab herausgeschält und überwiegend diejenigen Forschungseinrichtungen ganz oben in der Rangfolge platziert, die sich ganz und gar dieser Art der Soziologie verschrieben haben (Münch / Baier 2009). Dass es daneben ebenso wichtige andere Arten gibt, ist weitgehend untergegangen. An die Stelle des Nebeneinanders unterschiedlicher Soziologien ist eine Hierarchie getreten, bei der die positivistische Variante den Maßstab setzt. An dieser allgemeinen Tendenz können Einzelfälle der expliziten Berücksichtigung singulärer Forschungsleistungen nichts ändern, zumal das nur dann geschehen ist, wenn breit ausgewiesene Reputation als Entscheidungshilfe zur Verfügung stand.

Der damit einhergehende Verlust an Erkenntnis- und Erklärungspotential erscheint schon so selbstverständlich, dass noch nicht einmal beim DGS-Kongress darüber gestritten wird.

Da indessen die positivistische Soziologie auf den Kongressen – wie im Vorstand und in den Sektionen – weit unterrepräsentiert ist, zeigt sich das merkwürdige Bild einer in zwei Teile gespaltenen Soziologie: auf der einen Seite die zunehmend positivistisch verfahrende Soziologie in den dominanten Fachzeitschriften, auf der anderen Seite eine davon kaum Notiz nehmende Vielfalt an Soziologien, die aber in den unter dem Indikatorenregime dominant gewordenen Fachzeitschriften gar nicht zu sehen ist. Man könnte etwas sarkastisch sagen, „der Kongress tanzt, bemerkt aber nicht, dass er gar nichts mehr zu sagen hat“. Auch diese Seite der Soziologie kreist um sich selbst und gibt nicht gerade ein Bild ab, das großes Interesse von außerhalb des jeweiligen Kreises weckt. Es fehlt dieser Forschung, wie auch der positivistischen Variante, häufig das, was Max Weber unter der Kulturbedeutung und Wertbeziehung der Forschung verstanden hat.

Man könnte vermuten, dass diese Belanglosigkeit auch deshalb gut gedeihen kann, weil alles, was von den quantitativen Indikatoren nicht erfasst wird, der völligen Beliebigkeit überlassen bleibt. Dem Fach fehlt eine Plattform, auf der um Antworten auf die Fragen unserer Zeit gerungen wird. Die Fachzeitschriften mit dem höchsten Impact leisten das nicht, weil sie sich dem standardisierten Artikel verschrieben haben, der DGS-Kongress leistet es nicht, weil er sich nahezu vollständig in der Hand der Sektionen befindet, die ihre Steckenpferde reiten, für die sich außerhalb ihres Zirkels niemand interessiert. Und die durch das Regime des Impact Factors in die Nachrangigkeit gestoßenen Fachzeitschriften haben nicht den dafür erforderlichen Einfluss. Die Inthronisierung des begutachteten Fachzeitschriftenaufsatzes mit Chancen der Wahrnehmung und Zitierung im Web of Science als Königsweg der Soziologie hat auch die Kehrseite, dass außerhalb dieser sehr eingeschränkten Erkenntnisproduktion ein Feld der Erkenntnisproduktion besteht, das mangels einflussreicher Publikationsplattform überhaupt nicht mehr wahrgenommen wird und der völligen Beliebigkeit anheimgegeben ist. Das lehrt uns eine relationale Betrachtung des Indikatorenregimes, die auch die andere, von diesem Regime mit produzierte Seite beleuchtet.

Für den Fall der Soziologie in Deutschland können wir demnach sagen, dass der explizite oder vielfach auch nur implizite Gebrauch quantitativer Publikations- und Zitationsindizes zunächst einmal die Diversität der in den Fachzeitschriften erscheinenden Soziologie enorm reduziert und den quantitativ-methodischen, Massendaten auswertenden Aufsatz zum nahezu allein herrschenden Standard gemacht hat. Diese Beobachtung liegt auf der Linie der von Espeland und Sauder durchgeföhrten Studie zum Ranking amerikanischer Law Schools durch *US News and World Report*. Neben dem schon erwähnten Aufsatz zur Reaktivität von Rankings haben die beiden Autoren in einem weiteren Aufsatz die Diversität einschränkende Wirkung dieses Rankings herausgearbeitet (Espeland / Sauder 2009). Die Dominanz der positivistischen Soziologie in den zentralen Fachzeitschriften bedeutet in Deutschland jedoch nicht, dass alle anderen Arten der Soziologie völlig verschwinden. Im Gegenteil, auf dem DGS-Kongress sind sie noch sehr zahlreich vertreten. Die soziologische Diversität lebt also noch. Die Frage ist allerdings, wie sie sich präsentiert. Sie tut das zwar auf dem DGS-Kongress und in den in die Nachrangigkeit gestoßenen Fachzeitschriften, und sie wird in den Universitätsseminaren gepflegt. Es geschieht aber weitgehend unbemerkt und trägt den Makel des anscheinend Überflüssigen, ist demnach auf lange Sicht in der Existenz bedroht. Ein Vorgeschnack davon, wohin die Entwicklung geht, gibt die Situation in Großbritannien und in den Niederlanden oder auch in etwas anderer Form in Schweden. Nach dem neoliberalen Prinzip „Value for Money“ muss dort die Soziologie entweder den Weg der positivistisch verfahrenden Professionellen Soziologie gehen (die Niederlande) oder sich wenigstens Policy-orientiert als gesellschaftlich nützlich erweisen (Schweden) oder beides zusammen tun (wie neuerdings in Großbritannien im Rahmen des Research Excellence Framework).

Wir können grundsätzlich zwischen primären und sekundären Faktoren unterscheiden, die auf die Zitationschancen einer Veröffentlichung einwirken. Dazu zählen die folgenden Faktoren:

Primäre Faktoren des Beitrags zum Erkenntnisfortschritt

- Durchbruch vs. Beitrag zum vorhandenen Wissen (Naturwissenschaften)
- Originelles, die Situation neu definierendes Werk vs. Kommentar zu vorhandenen Werken (Geisteswissenschaften).

Sekundäre Faktoren der Forschungsorganisation

Kollektive vs. singuläre Forschung

- breite, gemeinsam geteilte vs. singuläre Datenbasis
- große Zahl von Ko-Autoren vs. Einzelautor
- breite Vernetzung vs. Singularität der Forschung

Sekundäre Faktoren der symbolischen Macht

Zentrale vs. periphere Position

- des Forschungsgebietes
- des Landes
- der Institution
- der Sprache
- des Forschers/der Forscherin

Wir können annehmen, dass Durchbrüche in den Naturwissenschaften grundsätzlich gute Chancen haben, irgendwann auch wahrgenommen und zitiert zu werden, jedoch durch sekundäre Faktoren beschleunigt oder gebremst werden. In den Geisteswissenschaften dürfen sekundäre Faktoren ein größeres Gewicht als in den Naturwissenschaften haben, um originellen, die Situation neu definierenden Werken auch breite, in Zitaten zum Ausdruck kommende Anerkennung zu verschaffen. In diesen Wissenschaften findet sich eine extrem große Ungleichheit der Konzentration von Zitaten auf wenige, oft längst verstorbene Meister, während sich das aktuelle Geschehen im Sinne von Foucaults (1991) Ausschlussregeln des Diskurses weitgehend im Kommentar zu den Werken der Meister erschöpft. Für die Geisteswissenschaften ist der Fachzeitschriftenaufsatz ohnehin weitgehend auf den Kommentar beschränkt. In diesen Disziplinen lässt sich Nachhaltiges nur schwer in die Form eines Fachzeitschriftenaufsatzes zwängen. Da das Web of Science primär aber nur Fachzeitschriften registriert, ergibt sich aus der darauf basierenden Herrschaft der Zahlen eine Degradierung der aktuellen geisteswissenschaftlichen Forschung ausschließlich auf den Kommentar. Die Soziologie ist wie die Sozialwissenschaften insgesamt von Anfang an und bis heute zwischen den Anforderungen der Natur- und Geisteswissenschaften zerrissen. Das macht sich unter dem Regime des Journal Impact Faktors in neuer Weise bemerkbar. Durchbrüche wie in den Naturwissenschaften gibt es in den Sozialwissenschaften kaum, originelle Werke, die die Situation neu definieren nur in ihrer geisteswissenschaftlichen Abteilung, und das nur in ganz großen Abständen. Das Setzen auf den Peer Review im neuen Evaluationsregime bevorteilt den standardisierten Fachzeitschriftenaufsatz, während die Monographie mit dem Makel in die Bedeutungslosigkeit geschickt wird, dass sie nicht demselben rigorosen Begutachtungsverfahren unterworfen ist. Die Aufmerksamkeit für Monographien wird in hohem Maße von einem Oligopol weniger Verlage beherrscht. Im internationalen Geschäft sind das

die Universitätsverlage der Universitäten, die an der absoluten Spitze der internationalen Universitäts-Rankings zu finden sind. Über den Zugang zu diesen Verlagen wachen die Lektoren, die vor allem ökonomisch denken und nur solche Werke aufnehmen, von denen sie sich einen ordentlichen Impact und Absatz versprechen. Weil das in den Geistes- und Sozialwissenschaften anhand der wissenschaftlichen Qualität eines Werkes kaum zu entscheiden ist, halten sich die Lektoren zwangsläufig an die sekundären Faktoren der Rezeptionschancen einer Monographie. Daraus ergibt sich eine enge Verflechtung dieser Verlage mit dem Oligopol der dominanten Institutionen. Die extrem ungleiche Verteilung der Aufmerksamkeit für Monographien wird dadurch noch verstärkt und verfestigt. Das verdammt den größten Teil der geisteswissenschaftlich verfahrenden Soziologie erst recht zum Kommentar (im Wesentlichen zu Marx, Weber, Durkheim, Parsons, Habermas, Luhmann, Bourdieu). Die vielen Kommentatoren zitieren die großen Meister, aber nur sehr begrenzt andere Kommentatoren. Wir haben es hier mit einem hochgradig zentralisierten Zitationsnetzwerk zu tun. Da fast alle Meisterdenker tot sind, gibt es Platz für eine umfangreiche Priesterschaft, die sich wieder in wenige Oberpriester und viele einfache Priester unterteilt. Die Oberpriester bilden ein Oligopol der Auslegung der kanonischen Werke und bringen es auf eine ansehnliche Zitationsquote. Dagegen hat die breite Masse der Priesterschaft darum zu kämpfen, überhaupt wahrgenommen zu werden.

Mangels Durchbrüchen und mangels Chancen auf Originalität sind die sekundären Faktoren bei der Bestimmung der Zitationschancen in der Soziologie und in den anderen Sozialwissenschaften besonders bedeutsam. Und diese sekundären Faktoren sind ganz anders als Durchbrüche und originelle Werke in ihrer Verteilung im wissenschaftlichen Feld hochgradig stabil, nahezu unverrückbar und reproduzieren sich selbst mit jedem Zitat immer wieder von Neuem. In der Soziologie ist zu beobachten, dass die im theoretischen Diskurs ohnehin vorherrschende extrem ungleiche Verteilung der Aufmerksamkeit durch das Regime des Impact Faktors nochmals verstärkt wird, zumal der ohnehin stark wirkende Matthäus-Effekt dadurch befeuert wird. Auf der geisteswissenschaftlichen Seite der Soziologie üben die Klassiker und die ganz wenigen modernen Theoretiker eine so absolute Herrschaft aus, dass kaum Neues entstehen und Beachtung erlangen kann. Nahezu das gesamte Feld erschöpft sich im Kommentar. Deshalb ist es auch nahezu unmöglich geworden, einen theoretischen Beitrag in einer der dominanten Fachzeitschriften unterzubringen. Nahezu alles scheitert an dem Verdikt der Unausgegorenheit und der Redundanz, weil alles schon gesagt ist. Das Regime des Impact Faktors verbannt deshalb die Arbeit an theoretischen Fragen an den äußersten Rand der Disziplin, während die großen Meister umso unumstößlicher die Herrschaft im Feld innehaben. Dabei sind neben dem Kommentar auch die anderen Ausschließungsmechanismen nach Foucaults Ordnung des Diskurses wirksam: Was als vernünftig gilt, wird im inneren Kreis des theoretischen Diskurses bestimmt, das Wahre kann nur in der Perspektive der vorgegebenen Theorien vom Falschen geschieden werden, die dominanten Theorien wirken wie Denkverbote, die Konzentration auf die Autorschaft weniger Meister unterwirft alles Gesagte einer herrschenden Interpretationslinie und degradiert alle weiteren Beiträge zum Kommentar der Meister, um die sich Diskursgesellschaften scharen, die über den Zugang zum Diskurs wachen. Das entsprechende Zitationsnetzwerk stellt sich wie ein Oligopol dar, das allein durch seine Existenz dazu zwingt, sich seiner Herrschaft zu unterwerfen, um überhaupt am Diskurs teilnehmen zu dürfen. Wenn wir jetzt dieses Oligopol im Zuge der Inthronisierung des Impact Faktors als alles entscheidendes Statusmerkmal aus der Latenz in die offene Herrschaft überführen, dann ist der Gedanke nicht fern, den Diskurs komplett zu schließen, weil es anscheinend nichts mehr zu sagen gibt (Münch 2011, 2014). In der Politikwissenschaft ist das schon sehr weit gediehen. Reihenweise werden Lehrstühle für Politische Theorie nicht mehr besetzt. Die Soziologie „hinkt“ dieser Entwicklung noch etwas hinterher, es zeigen sich aber auch in diesem Fach genau dieselben strukturellen Schwierigkeiten, Lehrstühle für Soziologische Theorie neu zu besetzen.

Schlussbemerkungen

Wie wir sehen, sind Zitationsindizes wie andere quantitative Indikatoren kein Instrument der Evaluation wissenschaftlicher Forschung, das deren höchst vielfältiger Natur gerecht wird. Zu viele sekundäre Faktoren, die nichts mit der Originalität oder Qualität einer wissenschaftlichen Publikation zu tun haben, bestimmen die Häufigkeit, in der sie zitiert wird. Aber nicht nur das, durch den Mechanismus der Reaktivität schaffen sie eine eigene Realität der Zahlen, an der sich die wissenschaftliche Praxis orientiert, sodass die Steigerung von Zitationsquoten zum alleinigen Ziel wird. Damit verliert die wissenschaftliche Forschung die Offenheit der Suche nach Erkenntnis, die notwendig ist, um Neues hervorbringen zu können. Erkenntnisfortschritt ist nur auf der Grundlage vieler Misserfolge möglich. Diese kann sich die Wissenschaft aber nicht mehr leisten, wenn sie unter ständiger Beobachtung steht. Die Folge ist die Prämierung sicherer Wege zum Zitat und des Einsatzes der sekundären Faktoren der Aufmerksamkeitszeugung. Die Wege des Erfolgs werden in einem Maße vorgezeichnet, das die Hervorbringung von Neuem erheblich behindert. Für die Wissenschaft ist das in dem Sinne fatal, dass sie auf einen Weg der Normalisierung gebracht wird, der dem Erkenntnisfortschritt abträglich ist.

Literatur

- Adler, Robert / John Ewing / Peter Taylor (2009): Citation Statistics. A Report from the International Mathematical Union (IMU) in Cooperation with the International council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics, in: Statistical Science 24 / 1, S. 1-14.
- Alber, Jens / Florian Fliegner / Torben Nerlich (2009): Charakteristika prämieter Forschung in der deutschsprachigen Sozialwissenschaft. Eine Analyse der mit dem Preis der Fritz-Thyssen Stiftung ausgezeichneten sozialwissenschaftlichen Aufsätze 1981-2006, WZB Discussion Paper No. SPI 2009-201.
- Bormann, Lutz / Hans Dieter Daniel (2005): Does the h Index for Ranking of Scientists Really Work?, in: Scientometrics 65 / 3, S. 391-392.
- Bormann, Lutz / Hans Dieter Daniel (2008): What do citation counts measure? A review of studies on citing behavior, in: Journal of Documentation 64 / 1, S. 45-80.
- Bormann, Lutz / Rüdiger Mutz / Hans-Dieter Daniel (2008): Are there Better Indices for Evaluation Purposes than the h Index? A Comparison of Nine Different Variants of the h Index Using Data from Biomedicine, in: Journal of the American Society for Information Science and Technology 59 / 5, S. 830-837.
- Bourdieu, Pierre (1992): *Homo academicus*, Frankfurt / Main.
- Burawoy, Michael (2006): For Public Sociology, in: American Sociological Review 70 / 1, S. 4-28.
- Campbell, Donald T. (1957): Factors Relevant to the Validity of Experiments in Social Settings, in: Psychological Bulletin 54, S. 297-312.
- Campbell, Donald T. (1976): Assessing the Impact of Planned Social Change. Occasional Papers Series No. 8, Western Michigan University Kalamazoo, Evaluation Centers, online abrufbar unter: www.erk.ed.gov/PDFS/Ed303512, letztes Abrufdatum: 17.7.2012.
- Desrosières, Alain (2008 a): Pour une sociologie historique de la quantification. L'argument statistique I, Paris.
- Desrosières, Alain (2008 b): Pour une sociologie historique de la quantification. L'argument statistique II, Paris.
- DFG (2010): „Quality not Quantity“ – DFG Adapts Rules to Counter the Flood of Publications in Research, online abrufbar unter: www.dfg.de/en/service/press/press_release/2010/pressemitteilung_nr07/index.html, letztes Abrufdatum: 2.1.2015.

- DORA (2012): San Francisco Declaration on Research Assessment, online abrufbar unter: am.asb.org/dora/, letztes Abrudftdatum: 31.12.2014.
- Egghe, Leo (2006): Theory and Practice of the g-index, in: *Scientometrics* 69 / 1, S. 131-152.
- Espeland, Wendy N. / Michael Sauder (2007): Rankings and Reactivity. How Public Measures Recreate Social Worlds, in: *American Journal of Sociology* 113 / 1, S. 1-40.
- Espeland, Wendy N. / Michael Sauder (2009): Rankings and Diversity, in: *Southern California Review of Law and Social Justice* 18 / 3, S. 587-608.
- Fleck, Christian (2013): The Impact Factor Fetishism, in: *European Journal of Sociology* 54 / 2, S. 327-356.
- Foucault, Michel (1991): Die Ordnung des Diskurses, München.
- Hamann, Julian (2014): Stratifikation und Standardisierung. Effekte externer Leistungsbewertung in der Wissenschaft, unveröffentlichtes Manuskript, Warwick / UK.
- Hirsch, Jorge E. (2005): An Index to Quantify an Individual's Scientific Research Output, in: *Proceedings of the National Academy of Sciences of the United States of America* 102 / 46, S. 16569-16572.
- Hornbostel, Stefan (1997): Wissenschaftsindikatoren. Bewertungen in der Wissenschaft, Opladen.
- ICSU (International Council for Science) (2008): Publication Practices and Indices and the Role of Peer Review in Research Assessment, online abrufbar unter: www.icsu.org/publications/cFrs-statements/, letztes Abrudftdatum: 2.1.2015.
- Merton, Robert K. (1968): The Matthew Effect in Science, in: *Science* 159 / 3810, S. 56-63.
- Münch, Richard / Christian Baier (2009): Die Konstruktion der soziologischen Realität durch Forschungsrating, in: *Berliner Journal für Soziologie* 19 / 2, S. 295-319.
- Münch, Richard (2011): Akademischer Kapitalismus. Über die Politische Ökonomie der Hochschulreform, Berlin.
- Münch, Richard (2014): Academic Capitalism. Universities in the Global Struggle for Excellence. London – New York / NY.
- Porter, Theodore (1995): Trust in Numbers. The Pursuit of Objectivity in Science and Public Life. Princeton / NJ.
- Priem, Jason / Dario Taraborelli / Paul Groth / Cameron Neylon (2010): altmetrics: a manifesto, online abrufbar unter: altmetrics.org/manifesto/, letztes Abrudftdatum: 2.1.2015.
- Quack, Martin (2015): Myths, Challenges, Risks and Opportunities in Evaluating and Supporting Scientific Research, in: Isabelle M. Welpe / Jutta Wollersheim / Stefanie Ringelhan / Margit Osterloh (Hrsg.), Incentives and Performance. Governance of Research Organizations, Heidelberg, S. 223-239.
- Thomas, William I. / Dorothy Thomas (1928): The Child in America: Behavior Problems and Programs, New York / NY.
- Welpe, Isabelle M. / Jutta Wollersheim / Stefanie Ringelhan / Margit Osterloh (Hrsg.) (2015): Incentives and Performance. Governance of Research Organizations, Heidelberg.

Prof. Dr. Richard Münch
 Universität Bamberg
 Lehrstuhl für Soziologie
 Feldkirchenstr. 21
 96052 Bamberg
 richard.muench@uni-bamberg.de

