

Der Journal Impact Faktor in der Soziologie: Die Gefahren eines Un-Maßes

Von Jochen Gläser

Zusammenfassung: Der Journal Impact Faktor ist aus konzeptionellen und technischen Gründen kein valides Qualitätsmaß für Zeitschriftenaufsätze. Er wird dennoch als solches eingesetzt, weil er leicht als Surrogat für inhaltliche Bewertungen genutzt werden kann sowie Objektivität und Vergleichbarkeit verspricht. Solange er als Scheinlösung für immer häufiger auftretende Evaluationsprobleme dienen kann, was ihm insbesondere in den lebenswissenschaftlichen Fächern gelingt, wird er als Evaluationsinstrument überleben. Es bleibt zu hoffen, dass die in Gang kommende Bewegung gegen die Nutzung des Journal Impact Faktors für Evaluationen schneller an Kraft gewinnt als erste Versuche einer Nutzung in der Soziologie.

1. Der Impact Faktor: schlecht als Maß, interessant als Sonde

Eugene Garfield hatte den *Journal Impact Faktor* (JIF) ursprünglich entwickelt, um Zeitschriften für den *Science Citation Index* auszuwählen. Später sah er darin vor allem ein Instrument, das Bibliotheken bei der Auswahl von Zeitschriften unterstützen kann (Garfield 2006). Der JIF sollte als Maß für die Aufmerksamkeit, die einer Zeitschrift in einer oder mehreren Fachgemeinschaften zuteil wird, dienen. Seine Eignung selbst für diese Funktion ist umstritten; auch kann er sie nur für bestimmte Fächer überhaupt erfüllen. Für die Aufgabe, für die er heute am häufigsten genutzt wird – individuelle Forschungsleistungen zu bewerten –, ist er überhaupt nicht geeignet. Der JIF sollte aus ganz grundsätzlichen Erwägungen heraus nicht für die Evaluation von Forschungsleistungen eingesetzt werden. Das alles ist unumstrittener Teil des Standes der Forschung in der Bibliometrie (Moed et al. 1996; Amin / Mabe 2000; Glänzel / Moed 2002; Moed 2005: 91ff).

Damit könnte ich diesen Beitrag abschließen. Wenn sich der JIF in den Sozialwissenschaften als Bewertungsinstrument durchsetzt, dann kann das nichts mit seiner *technischen* Eignung zu tun haben. Seine Verbreitung in den Naturwissenschaften und in der Forschungsevaluation – und eine mögliche Diffusion in die Sozialwissenschaften – speisen sich vielmehr aus seiner *sozialen* Eignung. Der JIF verspricht die Lösung dreier Probleme. Er verspricht eine Leistungsbewertung, die effizient, unabhängig von den wissenschaftlichen Inhalten und vergleichbar ist. Solange die Nachfrage nach diesen Problemlösungen wächst und der JIF sein Versprechen nicht wirklich einlösen muss, hat er gute Chancen, sich zu behaupten.

Interessanter als die Diskussion technischer Mängel des JIF ist deshalb seine Benutzung als Sonde für Problemdynamiken in Wissenschaft und Wissenschaftsmanagement. Ich werde mich auf diese Problemdynamiken konzentrieren. Zuvor aber werde ich in das *ceterum censeo* der Bibliometrie einstimmen und die wichtigsten Argumente gegen den JIF vorbringen – nicht zuletzt, weil sich einige davon auf die in den Geistes- und Sozialwissenschaften immer beliebter werdenden Zeitschriften-Rankings durch Meinungsumfrage ausdehnen lassen.

2. Das Un-Maß Journal Impact Faktor

Der JIF wird wie folgt berechnet:

$$IF_y = \frac{Z_y(p_{y-1}, p_{y-2})}{a_{y-1} + a_{y-2}}$$

Der JIF des Jahres y ist der Quotient aus der Zahl aller Zitierungen Z_y , die Publikationen der beiden Vorjahre (p_{y-1} , p_{y-2}) im Jahre y erhalten haben, und der Anzahl zitierfähiger Items in diesen beiden Jahren ($a_{y-1} + a_{y-2}$). „Citable items“ sind

- Artikel,
- Reviews
- Konferenzpapiere
- „Technical notes“ und
- Artikel in Sonderbänden

Weitere Publikationen wie z.B. Editorials und Rezensionen werden nicht als „citable items“ mitgezählt, obwohl ihre Zitierungen im Zähler mitgerechnet werden (Thomson Reuters 2014).

Aus dieser Berechnung ergeben sich zwei inhärente technische Grundprobleme für alle Verwendungen des JIF. Erstens ist der JIF von den Praktiken des Publizierens und Zitierens abhängig. Diese Praktiken sind spezifisch für Fachgebiete, d.h. für relativ kleine soziale Einheiten. Die Grenzen zwischen Fachgebieten, innerhalb derer die Publikations- und Zitierpraktiken als homogen angenommen werden können, verlaufen durch die Zeitschriften. Es gibt vermutlich nur wenige Zeitschriften, die Publikationen nur eines Fachgebietes enthalten.

Daraus folgt aber, dass der JIF einer Zeitschrift mit dem einer anderen Zeitschrift nicht wirklich vergleichbar ist, weil man nicht weiß, worauf man die Differenz kausal zurechnen soll. Ist es die Qualität der Artikel, die Zahl von Zitierungen, die Artikel in den in der Zeitschrift repräsentierten Fachgebieten typischerweise erhalten, oder die Verzögerung, mit der Publikationen in den Fachgebieten typischerweise rezipiert werden? Um den JIF vergleichbar zu machen, müsste man ihn normieren. Obwohl diesbezügliche Vorschläge nicht ohne Probleme sind (Glänzel / Moed 2002; Moed 2005: 102), wären sie dem *status quo* vorzuziehen. Sie werden aber vom heutigen Produzenten des JIF Thomson Reuters nicht implementiert. Thomson Reuters hat lediglich auf die Kritik, der JIF sei auf schnelllebige Fächer wie die Biomedizin ausgerichtet, mit der Einführung eines zusätzlichen JIF reagiert, der ein Fünf-Jahres-Fenster zugrunde legt. Das Problem der fehlenden Vergleichbarkeit bleibt unlöst.

Zweitens kann der JIF überhaupt nur in den Gebieten eine Aussage machen, in denen die in der Publikationsdatenbank (hier: dem Web of Science, im folgenden WoS) erfassten Zitierungen von Artikeln repräsentativ für das Zitierverhalten des Fachgebietes sind. Das trifft bei Fächern wie der Molekularbiologie und Biochemie, bei denen mehr als 90% des Publikationsaufkommens im WoS erfasste Zeitschriftenaufsätze sind, mit großer Wahrscheinlichkeit zu. Ist dieser Anteil deutlich geringer, stellt sich sofort die Frage, ob die im WoS gesammelten Zeitschriftenaufsätze nicht einen ganz spezifischen und damit eben nicht repräsentativen Ausschnitt des Publikationsaufkommens eines Fachgebietes bilden. Die Liste der Gebiete, für die man sich das fragen muss, ist überraschend lang. Berechnungen von Moed zufolge sind in der Chemie 84%, in den Geowissenschaften 62%, in der Mathematik 53% und in der Soziologie 27% aller Publikation im WoS erfasst (Moed 2005: 125ff).

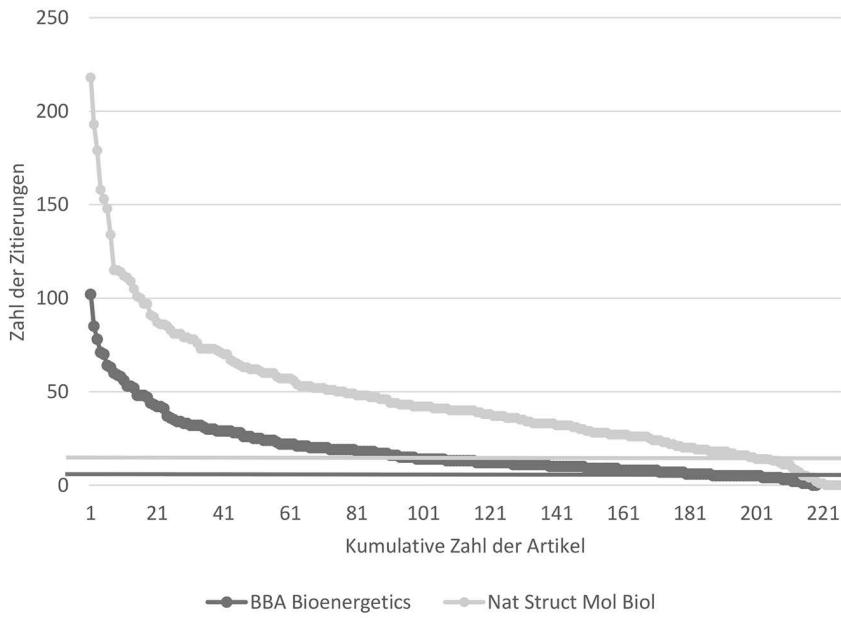
Die 27% soziologischer Publikationen sind keine repräsentative Stichprobe. Dass in Zeitschriftenaufsätzen anderes Wissen kommuniziert wird als in Monographien oder Buchkapiteln, hat z.B. Oltersdorf (2013) für die deutsche Kunstgeschichte gezeigt.

Neben diesen technischen Problemen, die den JIF sowohl für seinen ursprünglichen Zweck als auch für Evaluationen ungeeignet erscheinen lassen, steht ein prinzipielles Problem, dass den JIF als Evaluationsinstrument untauglich macht. Die Evaluation erfolgt ja dadurch, dass der JIF einer Zeitschrift, in der ein Aufsatz erscheint, als Qualitätsmerkmal der Publikation gewertet wird. Damit wird ein gelinde gesagt gewagter Schluss von der Verpackung auf den Inhalt gezogen: Wieso sollen Zitierungen, die *andere* Aufsätze in derselben Zeitschrift in der jüngeren *Vergangenheit* erhalten haben, etwas über die Qualität eines Artikels aussagen, der *jetzt* in dieser Zeitschrift veröffentlicht wird?

Die der Verwendung des JIF als Qualitätsindikator für Publikationen unterliegende Logik sagt, dass Zeitschriften mit einem höheren JIF im Mittel besser zitierte, d.h. bessere Publikationen enthalten, was die Möglichkeit und Fähigkeit der Zeitschrift impliziert, bessere Artikel auszuwählen. Aufsätze in Zeitschriften mit höherem JIF sind bessere Aufsätze, weil sie einen strengerer Auswahlprozess passiert haben.

Der Pferdefuß dieser Argumentation ist der Schluss vom Trend auf den Einzelfall. Was in der Tendenz richtig sein mag, kann ja im Einzelfall stets falsch sein. Ich möchte das an einem kleinen Beispiel demonstrieren (Abbildung 1). Die Zitierungen von Artikeln in jeder Zeitschrift sind extrem schief verteilt. In der Bibliometrie wird in der Regel von der Lotka-Verteilung ausgegangen, der zufolge die Zahl der Artikel mit n Zitierungen dem Quadrat von n umgekehrt proportional ist (Bookstein 1990 a, 1990 b). Abbildung 1 zeigt die Verteilungen von Zitierungen in zwei Zeitschriften mit deutlich verschiedenem JIF.

Abbildung 1: Vergleich der Zitierungen von Artikeln aus dem Jahre 2010 für die Zeitschriften *Nature Structural Molecular Biology* ($JIF_2(2010) = 13,685$) und *Biochimica et Biophysica Acta – Bioenergetics* ($JIF_2(2010) = 5,132$)¹



Das Diagramm macht erstens deutlich, wie wenig der JIF über die Zahl der Zitierungen aussagt, die ein Artikel erhalten wird. Zweitens wird klar, dass der Vergleich der beiden Werte (wenn sie denn überhaupt vergleichbar wären) für ganz viele Publikationen in die Irre führen würde. Die 14 meistzitierten Aufsätze der Zeitschrift *BBA Bioenergetics* haben jeweils mindestens 52 Zitierungen erhalten (mehr als das Zehnfache des JIF ihrer Zeitschrift). Diese vierzehn Artikel haben auch jeweils mehr Zitierungen erhalten als 155 Aufsätze in *Nature Structural Molecular Biology*, deren JIF mehr als doppelt so groß ist. Insgesamt widerspricht in 19,6% der möglichen Vergleiche von Artikeln aus beiden Zeitschriften das Verhältnis der erhaltenen Zitierungen der Aussage, die man aus dem Vergleich der JIF der Zeitschriften ableiten würde.² Bei einem Vergleich der *Zeitschrift für Soziologie* (ZfS) mit der *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (KZfSS) im Jahre 2010 läge dieser Anteil bei 30,4%, würde hier nicht sofort der ganze Unsinn von Zitierungsanalysen in den Sozialwissenschaften offenbar:

1. Welche Zeitschrift hat den höheren JIF? Im Jahre 2010 hatte die ZfS bei einem Zeitfenster von 2 Jahren den höheren JIF (0,611 gegenüber 0,536), bei einem Zeitfenster

1 Eigene Berechnungen auf der Grundlage von Downloads aus dem WoS am 5.12.2014.

2 Ich füge hastig hinzu, dass ich die meiner Kritik implizit zugrunde liegende Annahme, Zitierungsindikatoren würden überhaupt als Qualitätsmaß einzelner Artikel taugen, gar nicht teile. Ich habe sie lediglich für das Experiment voraussetzen müssen. Der Stand der bibliometrischen Forschung besagt, dass Zitierungen *im statistischen Aggregat* etwas über *einen Aspekt* von Qualität aussagen, nämlich den internationalen Einfluss (Van Raan 1996). Diese Aussage ist durch Vergleiche von Zitierungsanalysen mit Peer Reviews in den Naturwissenschaften und durch wissenschaftssoziologische Argumente relativ gut abgesichert (Gläser and Laudel 2007: 103ff).

von 5 Jahren die KZfSS (0,754 gegenüber 0,641). Wahrscheinlich bedeuten die Unterschiede überhaupt nichts.

2. Von den 25 im WoS enthaltenen Artikeln der Zeitschrift für Soziologie wurden bis zum Erhebungszeitpunkt sechs Artikel mehr als dreimal zitiert. Bei der KZfSS sind es 14 von 22 Artikeln. Sind das Häufigkeiten, auf denen Qualitätsurteile beruhen sollten?

Diese Werte belegen, dass die Kommunikation der deutschen Soziologie durch das WoS nicht adäquat erfasst wird. Das WoS erfasst nur ausgewählte deutsche Zeitschriften und ignoriert die meisten deutschen Sammelbände und Monographien. Nun könnte man es schlecht finden, dass deutsche Soziologinnen und Soziologen diese Publikationskanäle nutzen. Sie dafür mit invaliden Evaluationen zu bestrafen, scheint mir aber der falsche Weg.

Dass man daraus, dass ein Behälter früher einmal gute Dinge enthielt, nicht unbedingt auf die Qualität dessen schließen kann, was heute in ihm ist, gilt genauso für die vielerorts mit großem Aufwand erstellten Zeitschriftenrankings. Mit solchen Rankings sollen die Beschränkungen der Zeitschriftendatenbanken umgangen werden und alle in den Geistes- und Sozialwissenschaften genutzten Zeitschriften (und möglichst auch noch Buchverlage) eine vergleichbare Qualitätseinstufung erhalten. Das Prinzip ist dasselbe wie beim JIF: Aus der ‚Qualität‘ des Kommunikationskanals sollen Aussagen über die Qualität des gegenwärtig kommunizierten Inhalts abgeleitet werden. Die Argumente gegen den JIF lassen sich sinngemäß übertragen.

Es überrascht nun nicht mehr, dass von vielen Seiten ernsthaft und mit Nachdruck von der Nutzung des JIF in Evaluationen abgeraten wird. Dafür will ich nur drei prominente Beispiele bringen. In den Richtlinien für die Evaluationsrunde 2014 in Großbritannien, dem „Research Excellence Framework“, lesen wir: „53. No sub-panel will make use of journal impact factors, rankings or lists, or the perceived standing of the publisher, in assessing the quality of research outputs.“ (REF 2012: 8) Das gleiche gilt sinngemäß für die Pilotstudie des Wissenschaftsrates zum Forschungsrating in der Soziologie, in der die Evaluationskommission die Verwendung von bibliometrischen Indikatoren verwarf und sich die Mühe machte, Publikationen zu lesen (Bewertungsgruppe Soziologie 2008: 23, 38, 47).

Im Jahre 2012 entstand eine Initiative zur Forschungsevaluation, die dann im Jahre 2013 veröffentlichte „San Francisco Declaration on Research Assessment“. Diese Erklärung enthält eine allgemeine und zahlreiche spezifische Empfehlungen. Die allgemeine Empfehlung lautet:

„General Recommendation

1.

Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.“ (DORA 2013: 2)

Mehr muss man dazu nicht schreiben.

3. Nutzungen des JIF und ihre Ursachen

Warum all diese Aufregung um ein schlechtes Maß? Weil kein anderer bibliometrischer Indikator es geschafft hat, sich trotz offensichtlicher Nichteignung so fest in Wissenschaft und Forschungsevaluation zu etablieren.³ Das gilt bislang zwar nur für einige naturwissen-

3 Der keinesfalls auf sichereren messtheoretischen Fundamenten stehende Hirsch-Index könnte ihm allerdings bald den Rang ablaufen. Zu den Problemen dieses Indikators siehe z.B. Waltman / van Eck (2012).

schaftliche und wirtschaftswissenschaftliche Fächer. Wo der JIF aber etabliert ist, ist er es, weil die Wissenschaftlerinnen⁴ ihn ganz selbstverständlich nutzen. Publikationslisten auf Internetseiten, in Bewerbungen und in Projektanträgen enthalten ganz selbstverständlich die JIF der Zeitschriften, in denen die Wissenschaftlerinnen publiziert haben. Entscheidungen über Publikationsorte beziehen den JIF als wichtiges Kriterium ein.

Wissenschaftlerinnen nutzen den JIF, um im Reputationswettbewerb mithalten zu können. Sie müssen das tun, weil sie selbst in ihrer Rolle als Evaluatorinnen von Bewerberinnen und Projektanträgen den JIF nutzen, um Forschungsleistungen zu beurteilen. Seit Merton wissen wir, dass dieser Zyklus nur lange genug existieren muss, damit dann alle an den JIF glauben.

Die Ausgangsfrage lautet deshalb zunächst, warum Wissenschaftlerinnen den JIF als Qualitätsmaß nutzen. Darauf gibt es drei interessante Antworten. Erstens entstehen immer häufiger Situationen, in denen die ‚Peers‘, die den ‚Peer review‘ durchführen sollen, sich für die zu bewertende Forschungsleistung nicht kompetent fühlen, d.h. im strengen Sinne gar keine ‚Peers‘ mehr sind. Das folgende Zitat entstammt einem Interview mit einem Biowissenschaftler über seine Arbeit als Gutachter:

„We decided to ask all the reviewers we will talk to what kind of result is regarded as normal in the field, as normal scientific result, and what kind of result counts as outstanding.“

Our measure comes from impact factors of journals. You know, if you are looking for a global measure, that's how you do it: is your work published in ‚Science‘ or ‚Nature‘. They are the gold standard. And they are the gold standard because they accept even for review a very small number of the submitted materials. I think it is much more difficult to apply another standard because standards vary within certain branches of disciplines. So, I am an experimental plant biologist, essentially a plant physiologist and you know there are twenty or thirty branches. So, for me to decide on my own what reaches the highest standard, let's say for a synthesis – it takes me out of the province of understanding. If I see something in let's say in ‚Science‘ or ‚Nature‘ then I consider it to be exceptional. If I saw it in ‚Plant Cells‘, very high impact factor in plant biology (impact factor 13), it would also be a high measure.“

Im weiteren Verlauf des Interviews konnte der Wissenschaftler für sein unmittelbares Fachgebiet durchaus eine inhaltliche Beschreibung durchschnittlicher und außergewöhnlicher Leistungen geben. Im oben stehenden Ausschnitt reflektiert er die Beurteilung von Leistungen außerhalb dieser ‚Kompetenzzone‘. Dass solche Begutachtungen immer häufiger erwartet werden, ist der Binnendifferenzierung der Wissenschaft, der damit fortschreitenden Spezialisierung und der durch die Zunahme von Begutachtungsprozessen entstehenden ‚Gutachterinnenknappheit‘ geschuldet.

Zweitens und in engem Zusammenhang damit entlastet die Pseudo-Objektivität der Zahlen Wissenschaftlerinnen von der Verantwortung, die sie für ihre inhaltlichen Bewertungen übernehmen müssten. Dass sie die Verantwortung für den Einsatz problematischer Indikatoren übernehmen, spielt zumindest in den Fächern, in denen der JIF allgemein akzeptiert ist, keine Rolle.

Drittens verspricht der JIF etwas, was der Peer Review nur unter Aufgabe eines seiner zentralen Vorzüge leisten kann, nämlich Vergleichbarkeit.⁵ Man sieht den Zahlen ja nicht

⁴ Ich verwende in diesem Artikel durchgängig die weibliche Form. Die männliche ist jeweils eingeschlossen.

⁵ Der Peer Review kann innerhalb der ‚Kompetenzzone‘ der Gutachterinnen Vergleichbarkeit herstellen, indem er die Gutachterinnenurteile auf Zahlen oder Ranglisten reduziert (Gläser et al. 2010 a). Dies geschah bei der britischen *Research Assessment Exercise* und beim Forschungsrating des Wissenschaftsrates. Allerdings kann man sich schon fragen, ob die Reduzierung der mit großem Aufwand erarbeiteten komplexen Gutachterinnenurteile auf einige wenige Zahlen die beste Art ist, die knappe Zeit der Elite eines Faches zu nutzen.

an, dass sie gar nicht miteinander vergleichbar sind. Wann immer Qualitätsurteile aggregiert oder miteinander verglichen werden müssen, scheinen quantitative Maße den verbalen inhaltlichen Beurteilungen auch dann überlegen, wenn sie wegen der Fachgebietsspezifik der gemessenen Praktiken gar nicht vergleichbar sind (Gläser et al. 2010 a). In der leistungsorientierten Mittelvergabe an deutschen medizinischen Fakultäten wird – wie von der DFG empfohlen – der JIF zur ‚Qualitätsbewertung‘ von Publikationen eingesetzt:

„Für die Berechnung der LOM der jeweiligen leistungserbringenden Einheit (Abteilung bzw. Klinik) kann im Sinne einer stufenweisen Einführung von Qualitätskriterien die Bewertung erfolgter Original-Publikationen unter Verwendung des ungewichteten Impact Faktor der jeweiligen Publikationsorgane (JIF) erfolgen.“ (DFG 2004: 15)

Die Situation, die den JIF zu einem festen Bestandteil der Wissenschaftspraxis macht, entsteht also aus einer wachsenden Spezialisierung derjenigen, die inhaltliche Urteile über Forschung fällen können, sowie aus der steigenden Nachfrage nach Begutachtungen im Allgemeinen und vergleichenden Begutachtungen im Besonderen. Hinter der steigenden Nachfrage nach (vergleichenden) Begutachtungen stehen vor allem ein sich verschärfender Wettbewerb um Forschungsmöglichkeiten (Stellen und Drittmittel) und die Verwendung von Evaluationen als Instrument des ‚New Public Management‘. Der JIF verspricht unter diesen Bedingungen in einigen Fächern die Lösung eines immer häufiger auftretenden Problems, nämlich der effizienten Durchführung von *durch die Evaluierter akzeptierten* (nicht: validen) Evaluationen. Hinzu kommt eine Stabilisierung durch ‚Evaluationsgewinnerinnen‘: Unabhängig von der Validität von Evaluationsverfahren gibt es ja jeweils Gewinnerinnen (als leistungsfähig bewertete Einheiten) und Verliererinnen (als weniger leistungsfähig bewertete Einheiten). Erstere haben ein Interesse an der Beibehaltung der Evaluationsprozedur, das von deren Validität relativ entkoppelt ist.

4. Folgen

Welche Konsequenzen hat nun die Orientierung am JIF, und welche Konsequenzen hätte sie in der Soziologie?⁶ Die Orientierung von Autorinnen am JIF der Zeitschriften hat das Potential, die wissenschaftliche Kommunikation zu behindern. Das geschieht immer dann, wenn der JIF die Bedeutung der Zeitschrift für eine spezifische Zielgruppe nicht korrekt abbildet oder die Autorinnen Erwartungen der Zeitschriften erfüllen, um dort publizieren zu können. Der erstgenannte Fall scheint im Zeitalter der elektronischen Recherche weniger bedeutsam, da man davon ausgehen kann, dass relevante Publikationen in jedem Falle gefunden werden, dass sich Zeitschriften also allmählich in Kollektionen separater recherchierbarer Artikel auflösen. Wenn Autorinnen aber wissen, dass sie durch die Betonung des sensationellen Charakters ihrer Ergebnisse ihre Publikationschance in Zeitschriften mit hohem JIF vergrößern (Franzen 2011: 138ff) oder gezwungen werden, Artikel aus dieser Zeitschrift zu zitieren, um dadurch deren JIF zu erhöhen (Wilhite / Fong 2012), dann kann der JIF Verzerrungen der Inhalte wissenschaftlicher Kommunikation auslösen.

Die Orientierung von Evaluationen am JIF hat das Potential, Fehlentscheidungen herbeizuführen, d.h. die Vergabe von Stellen, Drittmitteln oder – im Falle der leistungsorientierten Mittelvergabe – Grundmitteln an Wissenschaftlerinnen, die in Zeitschriften mit hohem JIF publiziert haben, ohne dass der Inhalt des Titels dem aufgrund des JIF angenommenen

6 Der JIF wird meines Wissens in der deutschen Soziologie bislang kaum für Evaluationen genutzt. Am stärksten ist sein Einfluss vermutlich in dem vor allem durch die Verlage initiierten Wettbewerb der Zeitschriften um hohe JIF. Meine ad hoc – Recherche im Internet hat nur auf wenige Publikationslisten geführt, in denen zu Zeitschriftenaufsätzen der JIF der Zeitschrift angegeben wurde. Es gibt aber anekdotische Evidenz für die Berücksichtigung des JIF durch Bewerberinnen und Kommissionen in Berufungsverfahren.

Niveaus entsprach. Wie oben exemplarisch vorgeführt, können solche Differenzen häufig entstehen. Rafols und Kollegen haben einen Mechanismus in der Anwendung von Zeitschriftenrankings identifiziert, der die stärker interdisziplinär arbeitende und in Zeitschriften geringeren Ranges publizierenden Institute für Innovationsforschung an britischen Business Schools benachteiligen könnte (Rafols et al. 2012). Auch diese Möglichkeit spielt für die Soziologie bislang keine Rolle, da der JIF in unserem Fach kaum als Qualitätsindikator genutzt wird.

Eine dritte mögliche Konsequenz kann aber gerade in der Soziologie auftreten und sollte deshalb misstrauisch beäugt werden. Wenn das Bedürfnis nach (vergleichenden) Evaluierungen hinreichend stark wird, dann kann die Notwendigkeit, Publikationen mit einem JIF aufweisen zu können, ein starker Impuls zur Veränderung von Publikationspraktiken sein. Publikationen in im WoS gelisteten Zeitschriften wären dann viel wertvoller als alle anderen, weil sie überhaupt in die Bewertung einflössen. Dieser Impuls würde die Kommunikation auf wenige (zumeist englischsprachige) im WoS enthaltene Zeitschriften drängen. Monographien und insbesondere Buchkapitel würden relativ entwertet.

Damit werden zwei interessante wissenschaftssoziologische Fragen aufgeworfen. Die erste Frage bezieht sich auf den Zusammenhang von Forschungs- und Publikationspraktiken. Ist die Soziologie als eine ‚Zeitschriftenwissenschaft‘ vorstellbar? Das wäre nur möglich, wenn alle soziologischen Argumente in den räumlichen Begrenzungen von Zeitschriftenaufsätzen formuliert und dem Peer Review unterworfen würden. Ersteres scheint wegen des geringen Grades der Kodifizierung des Wissens der soziologischen Theorie und der qualitativen Sozialforschung unmöglich.⁷ Argumente in diesen Fachgebieten können sich keiner abkürzenden Symbolsprache bedienen, und sind oft zu umfangreich, um in Aufsätzen dargestellt werden zu können. Vergleichende qualitative Studien zum Beispiel stoßen ganz leicht an die Längenbegrenzungen von Zeitschriftenartikeln.

Auch ein ubiquitärer Peer Review scheint unserem Fach nicht angemessen. Die Soziologie hat mitunter mit den Geisteswissenschaften gemeinsam, dass persönliche Perspektiven der Wissenschaftlerinnen großen Einfluss darauf haben, welche Forschungsfragen als relevant wahrgenommen werden, was adäquate Bearbeitungsformen sind und was als empirische Evidenz gilt.⁸ In solchen Situationen birgt der Peer Review die Gefahr, soziologische Versuche ‚im Keim zu ersticken‘, bevor sie zu ihrer vollen Überzeugungskraft ausgearbeitet werden.

Die zweite Frage, die sich an eine ‚JIF-gerechte‘ Umwandlung der Soziologie stellen lässt, bezieht sich auf die Sprache, in der wir publizieren. Gewinnen würden wir eine viel bessere Einbindung in die internationale Soziologie, in der in der deutschen Soziologie entwickelte Ideen schneller und stärker zur Geltung kämen. Verlieren könnten wir bestimmte Stärken der deutschen Soziologie, die sich in nationalen Kommunikationsräumen halten, weil diese als Nischen fungieren.

Diese kurzen Überlegungen sollen nicht den Status quo verteidigen. Nicht jede soziologische Monographie enthält ein Argument, das einer Monographie bedarf. Nicht jeder deutsche Sammelband hätte unbedingt als solcher publiziert werden müssen. Vielleicht müssen wir auch unsere Publikationsstrategien stärker reflektieren. Ich wollte lediglich verdeutli-

⁷ Der Grad der Kodifizierung des Wissens bezeichnet das Ausmaß, in dem wissenschaftliches Wissen in prägnanten interdependenten theoretischen Formulierungen vorliegt (Zuckerman / Merton 1973[1972]: 507). Die Standardisierung der Sprache eines Fachgebietes ist ein guter – aber nicht perfekter – Indikator.

⁸ Diesen Faktor und seine Variation zwischen Fächern diskutieren Gläser et al. (2010 b: 311ff).

chen, dass ein aus dem Drang zum Un-Maß JIF resultierender Wandel der Publikationspraktiken reale Gefahren birgt.

5. Regeln

Was folgt aus all dem? Vielleicht die Hoffnung, dass die deutsche Soziologie den JIF „aus-sitzen“ kann, d.h. dass die Gegenbewegung hinreichende Kraft gewinnt und den JIF aus dem Evaluationsgeschehen vertreibt, noch ehe er in der Soziologie Fuß fasst. In jedem Fall gibt es einige ganz einfache Regeln für den Umgang mit dem JIF in der Soziologie:

1. Der JIF soll nicht verwendet werden (auch nicht für die Eigenwerbung der Zeitschriften).
2. Wenn eine Situation entsteht, in der die Verwendung des JIF unausweichlich scheint, dann soll der JIF nicht verwendet, sondern die Situation geändert werden.
3. Wenn die Verwendung des JIF günstig scheint, weil er der eigenen Person oder der eigenen Organisation(seinheit) nutzt, dann soll der JIF trotzdem nicht verwendet werden. Die langfristigen Gefahren aus der Institutionalisierung des JIF sind größer, als es irgendein kurzfristiger Nutzen sein kann.

Die Gründe, aus denen der JIF so populär ist, verweisen auf eine generelle Frage, die ich am Ende aufwerfen möchte. Natürlich ist es schwer, gegen Leistungsbewertungen zu argumentieren, weil man sich sofort dem Verdacht aussetzt, die eigene Faulheit schützen zu wollen. Dennoch: Brauchen wir wirklich alle Evaluationen, die wir gegenwärtig durchführen, und brauchen wir sie in den Zyklen, in denen sie stattfinden? Wären nicht weniger, aber validere Evaluationen die bessere Lösung?

Literatur

- Amin, Mayur / Michael Mabe (2000): Impact Factors: Use and Abuse, in: Perspectives in Publishing 1, S. 1-6.
- Bewertungsgruppe Soziologie (2008): Pilotstudie Forschungsrating Soziologie. Abschlussbericht der Bewertungsgruppe, Drs. 8422-08, Köln: Wissenschaftsrat, online abrufbar unter: <http://www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Grundlegende%20Dokumente%20zum%20Forschungsrating/8422-08.pdf>, letztes Abrufdatum: 2.2.2015.
- Bookstein, Abraham (1990 a): Informetric Distributions, Part I: Unified Overview, in: Journal of the American Society for Information Science 41 / 5, S. 368-375.
- Bookstein, Abraham (1990 b): Informetric Distributions, Part II: Resilience to Ambiguity, in: Journal of the American Society for Information Science 41 / 5, S. 376-386.
- DFG (Deutsche Forschungsgemeinschaft) (2004): Empfehlungen zu einer „Leistungsorientierten Mittelvergabe“ (LOM) an den Medizinischen Fakultäten – Stellungnahme der Senatskommission für Klinische Forschung der Deutschen Forschungsgemeinschaft, Bonn: DFG, online abrufbar unter: http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2004/stellungnahme_klinische_forschung_04.pdf, letztes Abrufdatum: 2.2.2015.
- DORA (2013): San Francisco Declaration on Research Assessment, online abrufbar unter: <http://www.ascb.org/dora-old/files/SFDeclarationFINAL.pdf>, letztes Abrufdatum: 2.2.2015.
- Franzen, Martina (2011): Breaking News: Wissenschaftliche Zeitschriften im Kampf um Aufmerksamkeit. Baden-Baden.
- Glänzel, Wolfgang / Henk F. Moed (2002): Journal impact measures in bibliometric research, in: Scientometrics 53 / 2, S. 171-193.

- Gläser, Jochen / Stefan Lange / Grit Laudel / Uwe Schimank (2010 a): Informed Authority? The Limited Use of Research Evaluation Systems for Managerial Control in Universities, in: Richard Whitley / Jochen Gläser / Lars Engwall (Hrsg.), Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and Their Consequences for Intellectual Innovation, Oxford, S. 149-183.
- Gläser, Jochen / Stefan Lange / Grit Laudel / Uwe Schimank (2010 b): The Limits of Universality: How field-specific epistemic conditions affect Authority Relations and their Consequences, in: Richard Whitley / Jochen Gläser / Lars Engwall (Hrsg.), Reconfiguring Knowledge Production: Changing authority relationships in the sciences and their consequences for intellectual innovation, Oxford, S. 291-324.
- Gläser, Jochen / Grit Laudel (2007): The social construction of bibliometric evaluations, in: Richard Whitley / Jochen Gläser (Hrsg.), The Changing Governance of the Sciences: The Advent of Research Evaluation Systems, Dordrecht, S. 101-123.
- Moed, H.F. (Hrsg.) (2005): Citation Analysis in Research Evaluation, Dordrecht.
- Moed, H.F. / T.N. vanLeeuwen / J. Reedijk (1996): A critical analysis of the journal impact factors of Angewandte Chemie and the Journal of the American Chemical Society – Inaccuracies in published impact factors based on overall citations only, in: Scientometrics 37 / 1, S. 105-116.
- Oltersdorf, Jenny (2013): Publikationen: Funktion und Repräsentation – Präsenz von Kommunikationskanälen der deutschen Kunstgeschichte in bibliographischen Nachweisinstrumenten. Dissertation, Humboldt-University of Berlin, online abrufbar unter: <http://edoc.hu-berlin.de/dissertationen/oltersdorf-jenny-2013-07-09/PDF/oltersdorf.pdf>, letztes Abrufdatum: 2.2.2015.
- Rafols, Ismael / Loet Leydesdorff / Alice O'Hare / Paul Nightingale / Andy Stirling (2012): How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management, in: Research Policy 41 / 7, S. 1262-1282.
- REF (2012): Panel criteria and working methods, online abrufbar unter: http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf, letztes Abrufdatum: 2.2.2015.
- Thomson Reuters (2014): Journal Impact Factor – Snapshot of a 3 Year Window, online abrufbar unter: http://wokinfo.com/media/pdf/JCR_Elements2013data-2014.pdf, letztes Abrufdatum: 2.2.2015.
- Van Raan, Anthony F.J. (1996): Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises, in: Scientometrics 36 / 3, S. 397-420.
- Waltman, Ludo / Nees Jan van Eck (2012): The inconsistency of the h-index, in: Journal of the American Society for Information Science and Technology 63 / 2, S. 406-415.
- Wilhite, Allen W. / Eric A. Fong (2012): Coercive Citation in Academic Publishing, in: Science 335 / 6068, S. 542-543.
- Zuckerman, Harriet / Robert K. Merton ([1972]1973): Age, Aging, and Age Structure in Science, in: Robert K. Merton (Hrsg.), The Sociology of Science, Chicago / IL, S. 497-559.

PD Dr. Jochen Gläser
 Technische Universität Berlin
 Zentrum Technik und Gesellschaft
 Hardenbergstr. 16-18
 10623 Berlin
jochen.glaeser@ztg.tu-berlin.de