Offener Brief zur Evaluierung der wissenschaftlichen Forschung

Giacomo Becattini, Enrico Bellino, Pierangelo Garegnani, Giorgio Lunghin, Sergio Parrinello, Luigi Pasinetti, Pierluigi Porta, Piero Tani, Gianni Vaggi, Alessandro Vercelli

In verschiedenen akademischen Bereichen (sowohl italienischen als auch europäischen) und hauptsächlich in der Wirtschaftswissenschaft verbreiten sich Forschungs-Bewertungssysteme, die auf indirekten Faktoren basieren: die einzelnen Beiträge der Grundlagenforschung werden anhand der Zeitschrift, in der diese veröffentlicht wurden, bewertet, und diese Bewertungen sind deshalb nicht unmittelbar vom wissenschaftlichen Wert des Inhaltes abhängig.

Die Qualität des Inhaltes wird anhand von Indikatoren wie dem „Impact Factor“ und durch die Einteilung der Artikel in Klassen gemessen, die verschiedene Elemente widerspiegeln (wie den Impact Factor selber und andere, die in Bezug auf die einzelnen Fälle unterschiedlich sind).

Die Vorteile dieser Methode sind differenziert: z.B. wird die Schnelligkeit und die Einfachheit, Meinungen über Arbeiten auszudrücken - selbst wenn es sich um Arbeiten handelt, die nicht im engeren Kompetenzbereich des Beurteilers liegen -, genannt. Dieser Methode wird vor allem der Vorteil der Unvoreingenommenheit unterstellt.

Insbesondere mit dem letzten Punkt stimmen wir nicht überein. Am wichtigsten scheint es uns, in diesem Zusammenhang zu erwähnen, dass diese Beurteilungsmethode vor allem die Wissenschaftler benachteiligt, die sich nicht auf die im Moment für die Mehrzahl der Wissenschaftler aktuellsten Themen und Methodologien beschränken.

Forscher, die sich mit weniger gängigen Themen befassen, die die Voraussetzungen des Ansatzes, der vom Großteil der internationalen Scientific Community verfolgt wird, kritisieren oder die eine Alternative zum gängigen Ansatz verfolgen, haben praktisch keine Möglichkeit, einen Platz in den gewichtigeren Zeitschriften zu erringen.

Die Review-Prozedur der Referees wird oft gar nicht erst eingeleitet; die Managing Editors der Zeitschriften selbst raten den Beitragenden, die einen Artikel zur Bewertung einreichen, diesen an spezialisiertere Zeitschriften zu schicken. Dies ist eine legitime Entscheidung für eine Zeitschrift; aber es gibt keinen Grund, weshalb der Verdienst eines Forschers mit der Linie des Herausgebers der meistquotierten Zeitschrift übereinstimmen sollte. Somit wird den Editors und den Referees einiger Zeitschriften eine allgemeine Bewertungsmacht zugeteilt, die deren natürliche Funktion überschreitet. Nicht umsonst haben die Erfinder des Impact Factors (welcher ursprünglich für die Bedürfnisse der Bibliothekare entwickelt wurde) vor deren Nutzung zur Forschungsbewertung gewarnt.

Neben der unverdientermaßen zugewiesenen Macht der Editors und Referees, über die Karriere der Wissenschaftler zu entscheiden, entstehen potentielle Konsequenzen, die für die Entwicklung der Disziplin gravierend sein können. Diese beinhalten:

a) Die Unabhängigkeit der Forschung: für die jungen Wissenschaftler entsteht ein Anreiz, sich mit Themen zu befassen, welche eine Publikation in den Zeitschriften in Aussicht stellt, statt sich mit Themen zu befassen, für die ein spontanes Interesse bestehen würde.
b) In Folge dessen entsteht ein Anreiz, sich mit den von der Mehrheit geteilten Prämissen konform zu zeigen, und dies führt zu einer Demotivierung für eine unabhängige und wirklich innovative Arbeit.
c) Die Ethik der Forschung: die Ermunterung zum intellektuellen Konformismus hat die Entfernung von der Disziplin der wirklichen Forschungs-Interessierten zur Folge.

Unter anderem sind die meist zitierten Zeitschriften angelsächsischen Ursprungs und dadurch entsteht die Gefahr, originelle Beiträge, die aus verschiedenen Kulturen stammen, zu unterdrücken. (Walras oder Pareto fanden z.B. keine offenen Türen in Zeitschriften und bei Herausgebern dieser Länder, die zwar für Ausländer offen erschienen, aber nur für diejenigen, die ihnen ähnliche kulturelle Tendenzen aufwiesen.)

Außerdem sollte nicht vergessen werden, dass für den Forscher der Zugang zu den meist zitierten Zeitschriften eher von der Fähigkeit abhängt, ein Netzwerk zu knüpfen und zu pflegen, und dass dies wichtiger wird als seine realen Fähigkeiten.

Dazu muss es eine Alternative geben, bei der man zwischen zwei Möglichkeiten unterscheiden kann: einerseits die Bewertung des Werkes einzelner Forscher (üblicherweise zu Wettbewerbszwecken, für Bestätigungen und Rufe), zum anderen die Bewertung der Institutionen, wie Universitäten und Forschungszentren (üblicherweise zum Zweck der Zuteilung von Ressourcen). Wir werden im Folgenden hauptsächlich auf die erste Bewertungsart eingehen.

Bei der Bewertung des einzelnen Forschers besteht die bevorzugte Alternative immer in der direkten Beurteilung durch die effektive Lektüre des Beurteilers. Die praktischen Probleme, wie die Schnelligkeit oder die Nicht-Übereinstimmung der zu bewertenden Arbeiten mit den Kompetenzen des Beurteilers, sind lösbar: Die Probleme der Schnelligkeit oder der Nicht-Kongruenz sind dadurch lösbar, dass einerseits die unbestreitbare Zunahme der Zahl der Wissenschaftler, welche zu bewerten sind, zu einer automatischen Erhöhung der Zahl der potentiellen Beurteiler führt. Andererseits könnte die Aufgabe der Selektion zum Teil den Kandidaten überlassen werden, indem sie aufgefordert werden, eine begrenzte Anzahl von Arbeiten zu präsentieren. Das Problem der Kompetenz der Beurteilenden ist dadurch lösbar, dass sie auf einen Experten zurückgreifen können, der vom Bewertenden selbst ausgewählt wird und dessen Bewertung zusammen mit der des Beurteilers öffentlich bekanntgemacht werden sollte.

Die Annahme der Verantwortung von Seiten derer, die zur Beurteilung berufen werden, und die Transparenz der Prozedur, durch welche Bewertungen schnell öffentlich gemacht werden (z.B. durch das Internet) und wodurch die Beurteiler selbst beurteilt werden können, erschwert in gewisser Weise die Anwendung unkorrekter Kriterien: von diesen hängt nämlich die Reputation der Beurteiler selbst ab.

Wir sind uns über die Missbräuche, welche sich in der Vergangenheit ergeben haben, bewusst und wir verurteilen sie.

Es scheint uns auf der anderen Seite illusorisch zu denken, dass diese Missbräuche sich durch mechanische Regeln verhindern lassen. Ihre Wurzeln liegen in den universitären Strukturen, im Zerfall des Verhältnisses Hochschullehrer/Student, in der Machtkonzentration, welche vom System ermöglicht wird. Auch wenn die Missbräuche durch traditionellere Formen verhindert werden könnten, finden sie andere Formen, wie z.B. die Manipulation der bibliometrischen Ergebnisse, über deren Möglichkeiten glaubwürdig berichtet wurde. Die traditionellen Missbräuche sind im Allgemeinen ein Ausdruck der individuellen Macht, und sie tendieren deshalb zu einer gewissen Interdependenz untereinander: leider beeinflussen sie auf negative Weise das Ergebnis einiger Wettbewerbe. Doch der Automatismus der vorgeschlagenen Lösungen ist aus den oben erwähnten Gründen ein noch schlechteres Gegenmittel. Eine solche Diskriminierung kann die Entwicklung der ökonomischen Disziplin auf eine schwer wieder gutzumachende Art schädigen, da die Disziplin wie jede andere Wissenschaft Entscheidungsfreiheit über die Argumente und Methoden benötigt, vor allem in einer Zeit des verbreiteten Unbehagens über den allgemeinen Stand der Disziplin, wie man es seit mehreren Jahren registrieren muss. Man sollte sich fragen, ob sich hinter einer oftmals zu selbstgefälligen Seriosität nicht die Absicht verbirgt, einige theoretische Orientierungen oder, noch schlimmer, gewisse Orientierungen der Wirtschaftspolitik, zu begünstigen.

Wir wissen alle, dass die Ökonomie eine Disziplin ist, die von der Macht der Interessen, die sie eigentlich analysieren sollte, stark negativ beeinflusst ist. Deswegen ist ein Umfeld, in dem eine totale intellektuelle Freiheit herrscht, umso notwendiger.

Unterschiedliche Betrachtungen in der Bewertung der Forschungsinstitutionen scheinen möglich zu sein. Zu diesem Ziel ist es verständlich, dass teilweise auf die bibliometrischen Indikatoren zurückgegriffen wird, zumal in diesem Umfeld die spezifische Bewertung der einzelnen Forscher zu Gunsten durchschnittlicher Bewertungen ersetzt wird. In diesem Fall scheint die Anwendung bibliometrischer Faktoren weniger schädlich, wenn diesen

Indikatoren jede Parteilichkeit entzogen ist und sie neben den Beiträgen in Zeitschriften auch andere Publikationen beinhalten wie Monographien, Essays usw.

Schlussfolgerung: viele Befürworter des Impact Factor nennen als Beispiel guter ökonomischer Forschung diejenige, die aus den USA kommt. Das zeigt eine Untersuchung, die von der italienischen Gesellschaft der Wirtschaftswissenschaftler bei den Verantwortlichen für die Auswahl der besten amerikanischen wirtschafts-wissenschaftlichen Fakultäten durchgeführt wurde. Es ging dabei darum, die bibliometrischen Kriterien zu analysieren, um die Forschung zu bewerten. Einer dieser Forscher hat erklärt, falls die wissenschaftlichen Arbeiten nicht auf eine solch mechanische Art beurteilt würden, dann wären die USA nicht in der Führungsposition der wirtschaftswissenschaftlichen Aktivitäten. Und wir, müssen wir das wirklich so anders beurteilen?

Übersetzung: Miriam de Blasi/Wolfram Elsner, Universität Bremen
Februar 2009