

# Der vermessene Nutzen der Forschung Impact als neuer Faktor im britischen Evaluationssystem

Martina Franzen, Judith Hartstein und Stefan Priester

**Summary:** While the utility of science had been considered to be self-evident for a long time, the increased competition for public funds brings about the need to practically prove its societal benefits. For the first time, the British Research Excellence Framework 2014 bases the allocation of public research funds on the evaluation of the societal impacts of research. Despite the overall very considerable results, questions remain about the effects of such an instrument on the conduct of research. These questions concern the expectation that research, instead of solely meeting scientific standards, is supposed to be of measurable practical relevance.

**Kurz gefasst:** Galt die Nützlichkeit von Wissenschaft lange als selbstevident, heißt es im Zuge des Wettbewerbs um knappe Fördermittel, den gesellschaftlichen Nutzen konkret zu belegen. Das britische Research Excellence Framework 2014 hat die Vergabe öffentlicher Forschungsmittel erstmalig auch an die Bewertung des gesellschaftlichen Impacts von Forschung geknüpft. Trotz der insgesamt guten Ergebnisse bleiben Fragen nach den Auswirkungen eines solchen Instruments auf die Forschung, wenn diese nicht allein wissenschaftsspezifischen Standards genügen, sondern zugleich von messbarer und nachweisbarer praktischer Relevanz sein muss.

Im Zuge der leistungsorientierten Mittelallokation wächst die Erwartung an Wissenschaft, den gesellschaftlichen Nutzen ihrer Forschung konkret unter Beweis zu stellen. Vorreiter bei der *Impact*-Messung für die Vergabe öffentlicher Mittel ist Großbritannien. Dessen nationales Forschungsrating zielt darauf ab, staatliche Forschungsausgaben so effektiv wie möglich zu gestalten. In wiederkehrenden Assessments werden deshalb die Forschungsleistungen aller Hochschuleinrichtungen evaluiert und nur die am besten bewerteten Einheiten mit staatlichen Zuwendungen bedacht – bis zur nächsten Evaluationsrunde, in der die Karten neu gemischt werden.

In den vergangenen drei Jahrzehnten britischer Forschungsevaluation ist das Verfahren Runde um Runde modifiziert worden. Dabei wurde stets das Ziel verfolgt, ein möglichst transparentes, auf Leistungsgerechtigkeit ausgelegtes Verfahren zu schaffen, das der wissenschaftlichen Diversität Rechnung trägt: Eingeführt in der Ära Thatcher 1986 unter der Bezeichnung *Research Selectivity Exercise* und weitergeführt als *Research Assessment Exercise*, erlebt das Forschungsrating nun unter der neuen Bezeichnung *Research Excellence Framework* (REF) in seiner siebten Runde eine gravierende Änderung: Erstmals wurde mit dem REF 2014 nicht mehr allein die wissenschaftliche Relevanz, sondern der gesellschaftliche Nutzen auf Basis von narrativen *Impact Case Studies* einer vergleichenden Bewertung unterzogen.

Die Einführung des *Impact*-Elements in das nationale Forschungsrating hatte in der Vorbereitungsphase bereits Kontroversen in der Wissenschaft ausgelöst. Die Hauptkritikpunkte zur Neuausrichtung bezogen sich auf vier Aspekte: Erstens würde die langfristige Wirkung von Wissenschaft gering geschätzt zugunsten des Kriteriums kurzfristiger Nutzenmaximierung. Zweitens werde Grundlagenforschung zugunsten anwendungsorientierter Forschung strukturell benachteiligt. Drittens würden kleinere gegenüber größeren Forschungseinheiten potenziell ungleich behandelt. Viertens entstünde ein systematischer Nachteil für die Geistes- und Sozialwissenschaften. Um diesen Einwänden zu begegnen, wurde *Impact* bewusst breit gefasst und schließt neben den ökonomischen Auswirkungen auch den politischen, kulturellen und sozialen Nutzen wissenschaftlicher Forschung mit ein.

In einer Pilotstudie, an der sich 23 Hochschuleinrichtungen in fünf Fachgebieten in den Jahren 2009 und 2010 beteiligten, wurde die Bewertung des *Impacts* erprobt. Auf Basis der Ergebnisse wurden einige Verfahrensänderungen vorgenommen, um den oben genannten möglichen Verzerrungen entgegenzuwirken. Im Zentrum der Diskussionen stand in den jeweiligen Begutachtungspanels die fachspezifische Operationalisierung der beiden Bewertungsdimensionen des *Impacts*: Reichweite und Signifikanz. Die Erhebung des *Impacts* erfolgte in Form von Fallstudien (*Case Studies*). Dabei handelt es sich um vierseitige narrationsbasierte Darstellungen, die durch ein Formular strukturiert und standardisiert sind: Im ersten Teil wird die zugrunde liegende Forschung präsentiert, der zweite Teil beschreibt den konkret erzielten *Impact* und listet Belege auf, die die Qualität hinsichtlich der Bewertungsdimensionen von Reichweite und Signifikanz untermauern sollen. Genau hierin

zeigen sich deutliche Unterschiede, die in den jeweiligen Bewertungen zum Ausdruck kommen.

Das REF-Verfahren sieht des Weiteren vor, dass jede teilnehmende Hochschuleinrichtung mindestens zwei *Impact Case Studies* (Anzahl in Relation zum Personalschlüssel) und ein sogenanntes *Impact Template* einreicht, das weniger die Inhalte, sondern vielmehr die (zukünftigen) Strategien zur Erreichung von *Impact* zur Darstellung bringt. Beide Dokumententypen, *Impact Case Studies* und *Impact Template*, werden einem Rating durch wissenschaftliche Experten und vereinzelte Anwender unterzogen. Mit einer Gewichtung von 4:1 werden die Bewertungen der *Impact Case Studies* und des *Impact Templates* zu einer Note zwischen eins und vier Sternen zusammengefasst.

Für die Verteilung der Zuwendungen auf Basis der Leistungsmessung wird das neu eingeführte *Impact*-Element nur mit 20 Prozent gewichtet. Den Schwerpunkt der Leistungszurechnung bilden auch im REF 2014 die wissenschaftlichen *Outputs*, die mit 65 Prozent in die Gesamtbewertung eingehen. Das dritte Element des Verfahrens ist das sogenannte *Environment*-Profil, das zum einen die Strategien umfasst, mit der eine Einrichtung ihre Forschungsleistung optimiert, und zum anderen die konkreten Arbeitsbedingungen der Forscherinnen und Forscher zum Gegenstand hat und zu 15 Prozent in der Gesamtwertung berücksichtigt wird.

Zum Verständnis des Verfahrens sollen die Bewertungsmechanismen am Beispiel der *Unit of Assessment* „Sociology“ genauer erläutert werden. In dieser haben 29 verschiedene Hochschuleinrichtungen insgesamt 100 *Impact Case Studies* eingereicht. Zu berücksichtigen ist allerdings, dass die im REF vorgenommene Klassifizierung nach *Units of Assessment* nicht einer strikten disziplinären Logik folgt, sondern zahlreiche thematische Überlappungen aufweist. Die erzielten Ratingergebnisse für das *Impact*-Element für die „Soziologie“ verteilen sich dabei wie folgt:

#### Ratingergebnisse für das *Impact*-Profil der Soziologie im REF 2014

| Grades  | Results |
|---|---------|
| 4 Outstanding impacts in terms of their reach and significance.   | 43.2 %  |
| 3 Very considerable impacts in terms of their reach and significance.   | 39.4 %  |
| 2 Considerable impacts in terms of their reach and significance.  | 13.6 %  |
| 1 Recognized but modest impacts in terms of their reach and significance.   | 3.3 %   |
| U/C The impact is of little or no reach and significance; or the impact was not eligible; or the impact was not underpinned by excellent research produced by the submitted unit. | 0.5 %   |

Wie aus der Tabelle der Ratingergebnisse ersichtlich wird, sind über drei Viertel der Einreichungen im Bereich Soziologie in Bezug auf den *Impact* mit den beiden bestmöglichen Bewertungen versehen worden. Bezogen auf die Grundgesamtheit der Einreichungen in den Facheinheiten rangiert die Soziologie damit im Mittelfeld. Vergleicht man die erzielten Ergebnisse mit den anderen *Units of Assessment* desselben *Main Panels C* (Sozialwissenschaften) liegt die Soziologie leicht über dem Durchschnitt. Zwar fällt die Streuung der Noten zwischen den Bewertungseinheiten äußerst gering aus, doch selbst minimale numerische Differenzen erfahren eine soziale Dynamik durch öffentliche Rankings. Die in Bezug auf den *Impact* vorderen Plätze unter den Einreichungen für die Soziologie im REF 2014 nehmen die Universitäten von York, Cardiff und Lancaster ein, das Schlusslicht bilden in absteigender Reihenfolge die Universitäten Queen's Belfast, Abertay und Winchester.

Diese Rangfolge deckt sich in etwa mit den Evaluationsergebnissen für den wissenschaftlichen *Output*: Mitglieder der sogenannten Russell Group, einem Zusammenschluss der selbst ernannten 24 führenden Universitäten des König-



Martina Franzen ist wissenschaftliche Mitarbeiterin der Forschungsgruppe Wissenschaftspolitik. Die Soziologin forscht über das Verhältnis zwischen Wissenschaft und Gesellschaft, speziell über wissenschaftliche Bewertungspraktiken und neue Inklusionsformen im (digitalen) Wandel.

[Foto: David Ausserhofer]

[martina.franzen@wzb.eu](mailto:martina.franzen@wzb.eu)

reichs, befinden sich eher an der Spitze und alle anderen eher am unteren Ende des allgemeinen Qualitätsprofils. Was den gemessenen *Impact* betrifft, durchbrechen nur Lancaster und Queen's Belfast (Russell Group) dieses Muster, indem sie hier unerwartet hohe beziehungsweise niedrige Werte erzielen. Lassen diese Ergebnisse nun den Schluss zu, dass die Bewertung von Forschung nach wissenschaftsexternen Relevanzkriterien am Ende zu den gleichen Resultaten führt?

Zur Interpretation des *Impact*-Assessment ist aus wissenschaftssoziologischer Perspektive zunächst empirisch zu klären, um welche Arten von *Impact* es sich konkret handelt. Wurde in den Leitlinien des REF-Verfahrens der Begriff von *Impact* bewusst auf alle gesellschaftliche Sphären ausgedehnt, zeigen sich in den einzelnen *Units of Assessment* doch gewisse Konzentrationseffekte eines bestimmten *Impact*-Typs. In der Soziologie ist *societal impact* wenig überraschend die dominante Ausprägung von *Impact*. Auffällig ist in dieser *Unit of Assessment* vor allen Dingen, dass *technological impact* so gut wie keine Rolle spielt, obwohl dieser Typ von *Impact* im Gesamtverfahren überaus prominent vertreten ist und in circa einem Fünftel aller *Case Studies* im Zentrum steht.

Seitens der Politik wurden die *Impact*-Messungen als großer Erfolg gewertet, bildeten sie doch scheinbar den insgesamt großen Wirkungsradius wissenschaftlichen Wissens jeder Fachkultur in den jeweiligen Sektoren ab. Die Ergebnisse der soziologischen Forschungseinheiten scheinen diesen generellen Befund zu bestätigen. Der verschwindend kleine Anteil von 0,5 Prozent der Einreichungen, die als nicht bewertbar eingestuft wurden, scheint zunächst dafür zu sprechen, dass das Verfahren tatsächlich alle relevanten Leistungen soziologischer Forschung adäquat zu erfassen vermochte.

Allerdings zeigt das Beispiel der Soziologie auch, welche Spannungen zwischen Verfahrens- und disziplinärer Logik im Zuge der Evaluation von *Impact* auftreten können. So kommt bereits das entsprechende Begutachtungspanel 23 in seinem Überblicksreport zu dem Schluss, dass soziologische Forschungseinheiten trotz der hohen Relevanz, die diese Disziplin insbesondere für langfristige gesellschaftliche Veränderungen hat, in den *Case Studies* kurzfristige Auswirkungen der eigenen Forschung in den Vordergrund stellen.

Die Krux der *Impact*-Messung besteht vor allem darin, überzeugende Evidenzen beispielsweise in Form von Patenten, klinischen Richtlinien, Videos oder, empirisch am meisten verbreitet, *testimonials* für die erzielte gesellschaftliche Wirkung zu erbringen. Die unterschiedliche Häufigkeit, mit der bestimmte Typen von *Impact* im gesamten Assessment der REF wie auch in einzelnen *Units of Assessment* vertreten sind, ist ein weiterer Hinweis darauf, dass das Verfahren möglicherweise die Einreichung bestimmter, in ihrer Kausalität gut darstellbarer Formen gesellschaftlichen Nutzens präferiert. Die konkrete Bewertungspraxis ist zudem gekoppelt an die Aktivitäten, die zur Erreichung des *Impacts* geführt haben. Zufällige *wissenschaftliche Breitenwirkungen*, wenn zum Beispiel ein spezielles Thema im politischen Diskurs kurzzeitig aufflammt, erhalten mit Blick auf die Pilotstudienresultate kaum die Bestnote. Und obwohl sich zwischen Verbreitungsstrategien oder Policy-Engagement und erzieltm gesellschaftlichem *Impact* empirisch nicht unbedingt ein direkter Zusammenhang zeigt, unterstellt das Aufbaumuster der Fallstudien eine solche Linearität.

Dieses Modell ist gerade auf die Sozialwissenschaften nicht immer übertragbar, wie sich beispielhaft an einem prominenten *Impact*-Erfolg der jüngeren deutschen Soziologie, Ulrich Becks 1986 publizierter „Risikogesellschaft“, aufzeigen lässt. Beck postuliert hier einen Bruch der Moderne, weg von der Industrie- und hin zur Risikogesellschaft. Soziologische Zeitdiagnosen dieser Art versprechen in der Regel einen hohen Absatz, der über die Grenzen der Fachgemeinschaft hinausgeht. Für den weltweiten Erfolg der „Risikogesellschaft“ kam jedoch ein weiterer Umstand hinzu, der sich nicht antizipieren ließ: Mitten in der Drucklegungsphase ereignete sich die atomare Katastrophe von Tschernobyl. Eine so enorme Verbreitung des Reflexionswissens, das Beck dem öffentlichen Diskurs bot, hätte nicht zuvor strategisch geplant werden können. Vielmehr war es die besondere „diagnostische Sensibilität“, wie es der Münchner Soziologe Armin

Nassehi einmal ausdrückte, die Becks Monografie zu einem zentralen Bezugspunkt öffentlicher Debatten und die „Risikogesellschaft“ zu einem geflügelten Wort werden ließ.

Ohne Frage stellt dieser Fall eine besondere Breitenwirkung soziologischer Forschung unter Beweis. Ob dem REF-Verfahren gemäß jedoch hinreichend Belege für einen „outstanding impact“ hätten angeführt werden können, zum Beispiel ein messbarer Effekt auf die politische Entscheidungsfindung, ist dagegen fraglich. Ein Blick in die im REF-Verfahren eingereichten Fallstudien im Bereich „Soziologie“ suggeriert die Dominanz eines bestimmten Forschungstypus. Legt man zwecks Klassifizierung der Soziologie das Vier-Felder-Schema von Michael Burawoy an, so scheint sich das Gros der *Impact Case Studies* im Feld der Angewandten Soziologie („Policy“) zu konzentrieren, wohingegen die Felder der Öffentlichen Soziologie, der Kritischen Soziologie und der Professionellen Soziologie unterrepräsentiert sind. Zu vermuten ist, dass sich in anderen Fachgebieten eine ähnliche Verteilung von eingereichten Forschungstypen konstatieren lässt.

Eine Klassifikation der Case Studies im REF-Verfahren eröffnet nun die Möglichkeit, noch einen Schritt weiterzugehen und zu fragen, wie sich die Orientierung am *Impact* auf die epistemische Diversität selbst auswirkt. Für die weitergehende Reflexion des Verfahrens der *Impact*-Messungen stellt sich somit eine generelle Frage: Können die beschriebenen Prozeduren, die darauf abzielen, einen ebenso eindeutigen wie in den Details nachprüfbaren Kausalzusammenhang zwischen Forschung und Praxis zu konstruieren, wirklich der langfristig wirkenden Forschung gerecht werden, die ja von Erkenntnisinteresse geleitet wird?

Und schließlich: Welche organisatorischen Anpassungsleistungen und welche fachlichen Selektionseffekte zeichnen sich in der hochschulinternen Governance ab, wenn es darum geht, den *Impact* von Forschungsleistungen in den Hochschuleinrichtungen systematisch zu erhöhen? Mit diesen Fragen werden wir uns im von der Mercator Stiftung geförderten Forschungsprojekt „Wissenschaft und Gesellschaft: ‚Research Impact‘ als Governance-Instrument von Forschung“ näher beschäftigen. Die vorliegenden rund 7.000 *Impact Case Studies* liefern für die empirische Beantwortung dieser Fragen reichhaltiges Material, das es mit Hilfe einer Kombination aus qualitativen und quantitativen Methoden der Textanalyse zu erschließen gilt.

#### Literatur

Burawoy, Michael: „For Public Sociology. American Sociological Association Presidential Address 2004“. In: *American Sociological Review*, 2005, Vol. 70, No. 1, pp. 4–28.

*Digital Science: The Societal and Economic Impacts of Academic Research. International Perspectives on Good Practice and Managing Evidence. Digital Research Reports: March 2016.*

Research Excellence Framework 2014: Search REF Impact Case Studies. Online: <http://impact.ref.ac.uk> (Stand 13.05.2016)

Research Excellence Framework 2014: Overview report by Main Panel C and Sub-panels 16 to 26, UOA 23, pp. S. 91–95. Online: <http://www.ref.ac.uk/media/ref/content/expanel/member/Main%20Panel%20C%20overview%20report.pdf> (Stand 11.05.2016).



Stefan Priester ist wissenschaftlicher Mitarbeiter der Forschungsgruppe Wissenschaftspolitik im Projekt Wissenschaft und Gesellschaft. ‚Research Impact‘ als Governance-Instrument der Forschung.

[Foto: Martina Sander]

[stefan.priester@wzb.eu](mailto:stefan.priester@wzb.eu)

Judith Hartstein ist Studentin im Masterstudiengang Wissenschaftsforschung an der Humboldt-Universität zu Berlin. Bis Mitte März 2016 war sie Praktikantin in der Forschungsgruppe Wissenschaftspolitik.

[hartstey@hu-berlin.de](mailto:hartstey@hu-berlin.de)