

Die Konstruktion soziologischer Exzellenz durch Forschungsrating

Von Richard Münch

Zusammenfassung: Das von einer Bewertungsgruppe im Auftrag des Wissenschaftsrates als Pilotstudie durchgeführte und 2008 veröffentlichte Forschungsrating Soziologie soll sich gegenüber den herkömmlichen Rankings durch bessere Qualität und Multidimensionalität auszeichnen. Wie in dem Aufsatz dargelegt wird, hatte jedoch auch dieses Evaluationsverfahren erhebliche Probleme, der Vielfalt der Forschungsrealität gerecht zu werden und treffsichere sowie zuverlässige Indikatoren der Forschungsaktivitäten zu bilden. Auch das Forschungsrating hat selektiv soziologische Exzellenz konstruiert, die sich bei anderer Messung anders darstellen würde. Ein bislang offenes Feld gleichberechtigter Diskursteilnehmer wird einer Stratifikation unterworfen, die zur weiteren Selbstreproduktion tendiert, weil die erworbenen Wettbewerbsvorteile in weitere Vorteile umgesetzt werden können. Die Soziologie entfernt sich so weiter als bisher von der idealen Sprechsituation. Die Evolution des soziologischen Wissens wird durch zunehmende Vermachtung geschlossen.

1 Forschungsrating: Auf der Suche nach wissenschaftlicher Exzellenz

Die Suche nach wissenschaftlicher Exzellenz ist zum Mantra der Forschungspolitik geworden. Die aktuellste Variante dieser Suchaktion ist das vom Wissenschaftsrat initiierte Verfahren des Forschungsratings. Mit diesem Verfahren soll anders als bei den sonst üblichen Rankings die Realität der Forschungsleistungen differenziert und objektiv erfasst werden. In Pilotstudien zur Chemie und zur Soziologie ist das neue Instrumentarium erprobt worden (Wissenschaftsrat 2007, 2008 a). Welche Konstruktionsprobleme bei der Vermessung der soziologischen Forschungsleistungen auftreten, darüber informiert der Abschlussbericht der Bewertungsgruppe Soziologie (Wissenschaftsrat 2008 a: 341-411) ganz offen und ungeschminkt (ebenso Neidhardt 2008). Die folgenden Analysen schließen daran an, um sie noch ein Stück weiter zu führen und die Probleme der Evaluation von Wissenschaft besser erkennen und verstehen zu können (vgl. Power 1997; Hirschauer 2004; Frey 2006; Espeland / Sauder 2007). Dabei geht es um die Bewertung der Forschungsleistungen der insgesamt 57 Forschungseinrichtungen (FER), die an dem Verfahren teilgenommen haben. Die im Forschungsrating zusammengetragenen Daten werden einer Sekundäranalyse unterzogen. So verdankt sich die folgende Analyse der zweijährigen Arbeit der Bewertungsgruppe, bei der alles getan wurde, was in einem solchen Rahmen überhaupt möglich ist. Die Daten wurden von der Geschäftsstelle des Wissenschaftsrates zur Verfügung gestellt. Alle Daten in den Tabellen 1 bis 9 entstammen diesem Datensatz.

Auch die hier vorgelegte Sekundäranalyse des Forschungsratings kann nicht alle bewertungsrelevanten Gesichtspunkte zur Sprache bringen. Sie kann das Spektrum der Gesichtspunkte gegenüber dem Abschlussbericht der Bewertungsgruppe lediglich ein Stück weit erweitern. Weitere Analysen müssen den Prozess der Selbstaufklärung über Verfahren der Evaluation, des Rankings und des Ratings weiter vorantreiben, um ihre Effekte auf die Evolution des Wissens zu ermitteln. Da es sich um einen Akt der Selbstaufklärung handelt, lässt sich nicht vermeiden, dass der hier vorgelegte Beitrag selbst im Machtfeld der Evaluation mitspielt und aus einer Position in diesem Feld heraus argumentiert. Dazu gehört auch, dass der Verfasser dieses Beitrags als Mitglied der Bewertungsgruppe selbst am Forschungsrating Soziologie teilgenommen hat und Mitglied einer Forschungseinrichtung (FER) ist, die als Ganzes im Vergleich zu anderen FER evaluiert wurde. Die dadurch entstandenen Verzerrungen müssen in der kritischen Debatte erkennbar gemacht und überwunden werden.

Das Forschungsrating des Wissenschaftsrates soll durch „informed peer review“ Forschungsleistungen sensibel, valide und zuverlässig messen (Gerhards / Wagner 2008). Es unterscheidet sich von den üblichen Rankings vor allem durch Multidimensionalität. Es bewertet Forschungseinheiten (FE) nach ihrer mit mehreren Indikatoren erfassten Forschungsqualität und die aus den einzelnen Forschungseinheiten (meist Lehrstühlen) zusammengesetzten Forschungseinrichtungen (FER) nach (1) der hinsichtlich Personalstärke gewichteten Forschungsqualität ihrer Einheiten, nach (2) Impact, (3) Effizienz, (4) Nachwuchsförderung, (5) Transfer in die Praxis und (6) Wissensvermittlung und Wissensverbreitung in der Öffentlichkeit. Während die Kriterien (1) bis (4) nach einer Fünferskala angewandt wurden, kam bei den Kriterien (5) und (6) mangels Differenzierungsvermögen nur eine Dreierskala zur Anwendung. Als Verfahren wurde ein „informed peer review“ gewählt. Einerseits wurden mit quantitativen Daten (Drittmittel, Publikationen, Promotionen und dgl.) relative und absolute Perzentilwerte errechnet, andererseits qualitative Daten, insbesondere die von allen FE eingereichten Texte von jeweils 50 Seiten Maximallänge pro Text, von jeweils zwei Gutachtern (aus der Bewertungsgruppe oder extern) bewertet. Auf der Basis der Textlektüre und der Perzentilwerte wurden die FE nach einer Fünferskala von 1 (nicht befriedigend) bis 5 (exzellent) eingestuft, worüber schließlich im Plenum der aus 16 Mitgliedern bestehenden Bewertungsgruppe entschieden wurde. Den Perzentilwerten waren die folgenden Noten zugeordnet: 0 bis 10 % nicht befriedigend, über 10 bis 30 % befriedigend, über 30 bis 70 % gut, über 70 bis 90 % sehr gut, über 90 bis 100 % exzellent (Wissenschaftsrat 2007, 2008 a).

2 Kriterium I: Forschungsqualität (FQ)

Bei der ersten Veröffentlichung der Ergebnisse des Forschungsratings Soziologie durch den Wissenschaftsrat am 18. April 2008 wurde aus der Forschungsqualität (FQ) der einzelnen FE ein Mittelwert gebildet. In der öffentlichen Rezeption wurde daraus ein Ranking der Standorte gemacht, das – weil es an erster Stelle der sechs Kriterien stand – das Abschneiden bei den anderen Kriterien überstrahlte. So wurde die Intention der Multidimensionalität in der öffentlichen Wirkung des Forschungsratings unterlaufen. Die Geschäftsstelle des Wissenschaftsrates hat daraus in einer tabellarischen Übersicht der Ergebnisse vom 6.8.2008 die Konsequenz gezogen, den Mittelwert aus der Kurzdarstellung der Ergebnisse wieder herauszunehmen und statt dessen die Spannweite der erzielten Ergebnisse in der FQ sowie die Gesamtzahl der gemeldeten FE der FER zu präsentieren und auf eine Reihung nach diesem Kriterium ganz zu verzichten. Das hat zu einer differenzierteren Darstellung der Ergebnisse beigetragen (Wissenschaftsrat 2008 b). Welchen Mittelwert der FQ eine FER erreicht hat, hing stark von ihrer Größe, ihrer Unterteilung in einzelne FE und der Aufgabenstellung der Einheiten ab. Zehn von 57 FER haben sich insgesamt als eine einzige FE definiert, wovon acht über mindestens zwei, einzelne davon sogar über fünf oder noch mehr Abteilungen verfügten. Durch diese Strategie war es möglich, einen höheren Durchschnittswert zu erzielen als bei der Unterteilung in einzelne FE. Die stärkeren FE haben die schwächeren mit nach oben gezogen. Das machte sich insbesondere bei der Erzielung von Spitzenwerten bemerkbar. Außerdem ergab sich bei der Selbstdefinition einer größeren FER als eine einzige FE der Vorteil, eine festgelegte Zahl von Texten aus einem größeren Pool von Personen auszuwählen als bei der Unterteilung der FER in mehrere FE. So hat beispielsweise eine aus 29 Personen bestehende und als eine einzige FE definierte FER elf Texte zur qualitativen Begutachtung eingereicht. Bei einer Unterteilung in fünf oder sechs FE wäre die Einreichung von 15 bzw. 17 Texten das in aller Regel ausgeschöpfte und dadurch zur Norm gewordene Maximum gewesen. Es liegt auf der Hand, dass sich 29 Personen leichter tun, elf statt 15 bzw. 17 durch Peer Review geweihte Fachzeitschriftenaufsätze in das Forschungsrating einzubringen. Wurde nur eine durchschnittliche Bewertung erreicht, dann konnte es aber auch vorkommen, dass sehr gute Leistungen einzelner Einheiten im Mittelwert ihrer FER untergegangen sind.

Problematisch war auch die Bindung des Ratings an den Fünfjahreszeitraum vom 1.1.2001 bis zum 31.12.2005. Alles, was davor oder danach geleistet wurde, hat nicht gezählt. Wer z.B. im Zeitraum davor Bedeutsames publiziert hatte, danach aber eine Forschungsphase durchlief, die erst nach dem 31.12.2005 zu Publikationen führte, konnte kein herausragendes Ergebnis erzielen. Auch auf die Abhängigkeit der Produktivität von der jeweiligen Lebensphase konnte keine Rücksicht genommen werden. Dazu kommt noch, dass das Eintreten in oder Ausscheiden aus einer Forschungseinheit bzw. -einrichtung vor oder nach dem Stichtag am 31.12.2005 von ganz entscheidender Bedeutung sein konnte. Für die sehr hoch gewichteten relativen Publikationsleistungen pro Personaleinsatz war außerordentlich bedeutsam, wie viele Drittmittelmitarbeiter zum Stichtag am 31.12.2005 in einer Einheit bzw. Einrichtung beschäftigt waren. Viele Drittmittelmitarbeiter – die z.B. zunächst mit der Datenerhebung und nicht mit Publikationen beschäftigt sind – haben den Nenner belastet und die relativen Perzentilwerte nach unten gezogen.

In der öffentlichen Rezeption des Forschungsratings findet jedoch keine Relativierung der verteilten Noten nach den Bedingungen statt, unter denen sie zustande gekommen sind. Die Noten stehen als ein ziemlich unumstößliches soziales Faktum da und wirken etikettierend. Auch wenn nachdrücklich betont wird, dass sich die Qualitätsbewertungen auf die Forschungseinheiten und nur auf den Fünfjahreszeitraum und nicht auf einzelne Personen und deren ganzes Lebenswerk beziehen, wird diese Trennung in der öffentlichen Rezeption de facto nicht vorgenommen. Und es ist auch nicht zu verhindern, dass die weiteren Erfolgschancen der einzelnen Personen maßgeblich von der Bewertung ihrer Forschungseinheit und auch der ganzen Forschungseinrichtung in dem untersuchten Fünfjahreszeitraum beeinflusst werden. De facto gibt es in Einheiten und ganzen Einrichtungen, die nur als durchschnittlich bewertet wurden, jedoch exzellente Forscherinnen und Forscher, wie auch in solchen Einheiten und Einrichtungen, die als sehr gut oder gar exzellent eingestuft wurden, nur durchschnittliche. Der Mittelwert der Forschungsqualität der FER verzerrt deshalb die tatsächliche Verteilung von Forschungsqualität auf Einheiten und Einrichtungen.

Eine „objektiv gültige“ Reihung der FER nach ihrer FQ ist nicht möglich. Sie sieht in Abhängigkeit von der Methode jeweils anders aus. Eine Möglichkeit, dieser Tatsache einigermaßen, wenn auch nicht vollständig, gerecht zu werden, besteht darin, mehrere Maßzahlen nebeneinander zu stellen. Das soll hier versucht werden (Tab. 1). Dabei kommen die folgenden Indikatoren zur Anwendung:

- (1) die FQ der besten vier FE einer FER
- (2) der mittlere Perzentilwert der begutachteten Fachzeitschriftenaufsätze (Peer Reviewed Articles, PR-Artikel) der besten und der mittlere Perzentilwert der besten drei FE einer FER nach Vollzeitäquivalenten des Personals gewichtet
- (3) der Perzentilwert der Monografien der besten und der mittlere Perzentilwert der besten drei FE einer FER nach Vollzeitäquivalenten des Personals gewichtet
- (4) die Spannweite der FQ der FE einer FER
- (5) die Anzahl der gemeldeten FE einer FER
- (6) der Prozentanteil des Personals der exzellenten oder sehr guten FE am gesamten Personal einer FER
- (7) der Mittelwert der FQ aller FE einer FER im Forschungsrating, gewichtet nach Vollzeitäquivalenten des Personals
- (8) die Noten des FR einer FER bei den Kriterien II (Impact), III (Effizienz), IV (Nachwuchsförderung), V (Wissenstransfer) und VI (Wissensverbreitung)
- (9) die Summe aus dem Mittelwert der FQ der besten drei FE und den Noten des FR bei den Kriterien II bis VI

Tabelle 1: Forschungsqualität der besten vier Forschungseinheiten der Forschungseinrichtungen

FER	FQ beste vier FE	PR-Artikel VZÄ-gew.		Monogr. VZÄ-gew.		Spannweite FQ	Anzahl FE	Prozent sehr gute / exzel-lente FE	Mittelwert FQ im FR	Noten Krit. II, III, IV, V, VI	Summe Mittelwert FQ beste drei FE u. Noten Krit. II bis VI
		beste FE	beste drei FE	beste FE	beste drei FE						
SOEP	5	98		77		5	1	(100)	(5)	45344	(25)
Mannheim	5	88		34		5	1	(100)	(5)	54422	(22)
MPIfG Köln	5	85		39		5	1	(100)	(5)	45534	(26)
Bamberg	5543	95	92	100	60	1-5	7 (1)	41	3,3	44444	24,7
LMU München	5444	98	95	100	73	4-5	4 (2)	78	4,3	55323	22,3
FU Berlin	5333	97	81	90	52	2-5	7 (1)	15	2,9	43433	20,7
Konstanz	5333	96	80	63	60	3-5	5	11	3,2	43222	16,7
Erfurt	53	94	(79)	93	(60)	3-5	2	45	(3,9)	34322	(18)
Düsseldorf	44	100	(84)	55	(51)	4	2	(100)	(4)	44332	(20)
Bielefeld	4444	83	82	85	77	2-4	10	40	3,3	53544	25
Frankfurt / Oder	4	76		58		4	1	(100)	4	24322	(17)
Jena	4443	90	79	86	86	2-4	6	67	3,5	44244	22
JU Bremen	4442	100	98	88	64	1-4	5	53	2,8	35-32	(20)
Heidelberg	4433	97	82	92	84	2-4	5	52	3,4	33322	16,7
Siegen	4433	90	79	65	60	1-4	5	44	3,2	33123	15,7
Köln	4433	89	77	69	59	2-4	6	37	3,2	33233	17,7
HU Berlin	4433	89	72	84	79	2-4	8 (1)	19	3	54543	24,7
Göttingen	4433	79	58	88	75	3-4	4	55	3,5	54443	23,7
Bremen	4432	92	81	78	61	2-4	4	32	3,2	44533	22,7
Leipzig	4422	84	65	94	69	2-4	4 (4)	30	3,2	-	-
Kassel	4333	91	74	98	77	1-4	6	41	3	33233	17,3
UBW München	43	78	(73)	37	(37)	3-4	2 (1)	26	(3,5)	22222	(13,5)
Trier	4333	95	59	91	62	1-4	5 (1)	12	2,8	32342	17,3
WZB	43	73	(59)	40	(36)	3-4	2	37	(3,4)	33334	(19,5)
Bochum	4333	63	54	97	70	2-4	7	29	3	33444	21,3
Frankfurt / Main	4333	87	54	99	87	1-4	9 (1)	5	2,5	42433	19,3
TU Chemnitz	4332	88	67	88	72	2-4	5	34	3	33332	17,3
TU Berlin	4332	82	65	97	76	2-4	5	30	3	33223	16,3

FER	FQ beste vier FE	PR-Artikel VZÄ-gew.		Monogr. VZÄ-gew.		Spannweite FQ	Anzahl FE	Prozent sehr gute / exzel-lente FE	Mittelwert FQ im FR	Noten Krit. II, III, IV, V, VI	Summe Mittelwert FQ beste drei FE u. Noten Krit. II bis VI
		beste FE	beste drei FE	beste FE	beste drei FE						
Hagen	4332	73	55	99	96	2-4	6	27	2,8	22122	12,3
Tübingen	4331	83	50	82	69	1-4	4 (1)	25	2,9	23332	16,3
Dresden	4322	76	67	53	37	1-4	5	21	2,5	33333	18
Mainz	4322	72	43	93	81	2-4	4	35	3	22244	17
Eichstätt	431	82	57	99	72	1-4	3	31	2,8	23222	13,7
U Rostock	42	76	(64)	43	(36)	2-4	2 (1)	17	(2,7)	33234	(18)
UBw Hamburg	422	95	40	73	44	2-4	3	36	2,7	11123	10,7
Duisburg-Essen	3333	94	79	66	63	1-3	9 (1)		2,4	33344	20
Bonn	3	74		52		3	1		(3)	23233	(16)
Aachen	3	70		41		3	1		(3)	23422	(16)
Wuppertal	3	61		81		3	1		(3)	33233	(17)
Freiburg	33	67	(60)	98	(94)	3	2		(3)	33343	(19)
Halle-Wittenberg	3333	85	59	96	92	1-3	10		2,5	43344	21
Stuttgart	33	55	(52)	76	(43)	3	2		(3)	22342	(16)
Münster	3	50		22		3	1		(3)	-	-
Erlangen-Nürnberg	3	39		49		3	1		(3)	33333	(18)
Hannover	3	39		41		3	1		(3)	21233	(14)
DSPH Köln	33	24	18	71	(61)	3	2		(3)	22242	(15)
Augsburg	3332	93	75	94	69	2-3	4		2,9	34232	17
Marburg	3332	86	73	88	74	1-3	7		2,6	33333	18
Dortmund	3332	69	61	84	78	1-3	7 (1)		2,1	32443	19
Gießen	3322	74		82	78	2-3	5		3	33334	18,7
U Ham-burg	3322	63		65	52	2-3	4		2,5	43333	18,7
Osnabrück	3321	38	21	76	60	1-3	7		1,7	12242	13,7
Kiel	3311	80	46	58	38	1-3	4		2,1	22233	14,3
TU Darmstadt	32	59	(36)	95	(64)	2-3	2		2,9	33333	(17,5)
Potsdam	322	63	53	85	60	2-3	3 (1)		2,3	11122	9,3
Magdeburg	3222	57	42	58	30	2-3	4		2,3	22333	15,3

FER	FQ beste vier FE	PR-Artikel VZÄ-gew.		Monogr. VZÄ-gew.		Spannweite FQ	Anzahl FE	Prozent sehr gute / exzel-lente FE	Mittelwert FQ im FR	Noten Krit. II, III, IV, V, VI	Summe Mittelwert FQ beste drei FE u. Noten Krit. II bis VI
		beste FE	beste drei FE	beste FE	beste drei FE						
Bayreuth	322	50	25	95	62	2-3	3		2,3	12223	12,3

Legende: FQ=Forschungsqualität laut Forschungsrating; FR=Forschungsrating; FE=Forschungseinheiten; FER=Forschungseinrichtung

Anmerkung: Werte in Klammern, z.B. (5), (4), (3) oder (100): Eine differenzierte Bewertung war nicht möglich, da die ganze FER als eine einzige FE gemeldet oder nur in zwei FE unterteilt wurde. Bei der Anzahl FE bedeuten Werte in Klammern die Anzahl nicht bewertbarer FE.

Die namentliche Reihenfolge der FER in Spalte 1 richtet sich nach der an erster, zweiter, dritter, vierter Stelle erreichten FQ der FE. Bei weniger als vier FE zählt jeweils die an erster, (zweiter), (dritter) Stelle erzielte FQ der FE. Bei Gleichstand entscheidet der Perzentilwert der PR-Artikel der besten drei FE einer FER, gewichtet nach Vollzeitäquivalenten des Personals.

Wir sehen, dass sehr gute bis exzellente Forschungsleistungen breit gestreut sind: 35 von 57 Standorten verfügen über mindestens eine sehr gute oder exzellente FE. Darüber hinaus ist zu erkennen, dass auch die drei besten FE von FER, die keine sehr gut bewertete FE vorweisen können, recht hohe Perzentilwerte bei den nach Vollzeitäquivalenten gewichteten PR-Artikeln erreichen. Zehn davon erzielen Werte zwischen 60 und 79 Prozent, sind also im sehr guten Bereich oder nahe daran. Alle FER beherbergen mindestens eine gute FE, die bei den PR-Artikeln, gewichtet nach Vollzeitäquivalenten, recht hohe Perzentilwerte erzielt.

FER, die nur eine FE gemeldet haben, führen nach wie vor zwangsläufig die Notenstufen 5 und 4 an. Auch bei der Notenstufe 3 sind sie gegenüber etlichen, in einzelne FE aufgeteilten FER im Vorteil. Dieses Ergebnis ist demnach maßgeblich durch die Zahl der bewerteten FE einer FER bestimmt. Die Reihung der FER in Tabelle 1 würde sich ändern, wenn man nur die ersten drei FE oder die Perzentilwerte der nach Vollzeitäquivalenten gewichteten PR-Artikel als Maßstab zugrunde gelegt hätte. Weitere Relativierungen ergeben sich aus den anderen Indikatoren. Man sieht daran, wie durch die Indikatorenwahl eine Rangfolge konstruiert wird. Die gewählte Vorgehensweise entspricht in etwa dem britischen Research Assessment Exercise (RAE), an dem explizit Personal einer FER nicht teilnimmt, das andere Aufgaben als professionalisierte Forschung per se erfüllt, wie z.B. Auftragsforschung oder Lehre. Auf diese Weise wird einigermaßen verhindert, dass Äpfel mit Birnen verglichen werden, zumal die Beurteilung der FQ einer FE nach einem sehr einseitigen, der Vielfalt von Forschungsleistungen nicht gerecht werdenden Maßstab erfolgte. Um in der FQ hohe Punktwerte zu erzielen, war nahezu ausschließlich das folgende Leistungskriterium entscheidend:

- Zahl der PR-Artikel, relativ zum Personal aus Grund- und Drittmittelausstattung (Professional Sociology)

Haben sich FE anderen Aufgaben als der Publikation begutachteter Fachzeitschriftenaufsätze gewidmet, dann konnten sie keine sehr guten Noten in der FQ erreichen. Das gilt insbesondere für die Aufgaben einer Policy-oriented Sociology, einer Public Sociology oder einer Critical Sociology, die Michael Burawoy (2005) von der Professional Sociology der Fachzeitschriften unterschieden hat. Die folgenden Tätigkeiten haben bei der Beurteilung der FQ vor allem dann eine negative Wirkung gehabt, wenn durch sie die Publikation von PR-Artikeln verhindert wurde:

- Wissenstransfer in die Praxis, z.B. in Gestalt von Auftragsforschung (Policy-oriented Sociology)
- Wissensverbreitung in der Öffentlichkeit durch zeitdiagnostische Werke und Zeitungsartikel (Public Sociology)
- Begriffliche und theoretische Grundlagenarbeit (Critical Sociology)
- Publikation von Monographien (eine Monographie hat im Endeffekt weniger Punkte gebracht als ein PR-Artikel), insbesondere mangels Qualitätssicherung durch Begutachtungsverfahren
- Publikation von Lehrbüchern, Handbüchern oder Handbuchartikeln
- Konzentration auf Spezielle Soziologien mit besonderer praktischer Relevanz wie z.B. Stadt- und Regionalsoziologie, Geschlechterforschung, Soziologie sozialer Probleme, Sportsoziologie oder Rehabilitationssoziologie
- Interdisziplinäre Forschung im Grenzgebiet zwischen der Soziologie und anderen Disziplinen wie z.B. Rechtssoziologie, Agrarsoziologie oder Medizinische Soziologie
- Intensives Engagement in der Methodenausbildung einschließlich Forschungspraktika bzw. Lehrforschungsprojekten mit hohem Zeitaufwand

Die Spalten 3 und 4 zu den PR-Artikeln und zu den Monographien, jeweils nach Vollzeit-äquivalenten gewichtet, zeigen die Schieflage des Forschungsratings zugunsten der PR-Artikel und zu Ungunsten der Monographien. Bestnoten in der FQ konnten FE nur mit sehr hohen Perzentilwerten bei den PR-Artikeln erreichen, relativ niedrige Werte bei den Monographien haben dies jedoch nicht durchgehend verhindert. Umgekehrt sind unter den als durchschnittlich oder unterdurchschnittlich eingestuften FE zahlreiche Einheiten mit sehr hohen Perzentilwerten bei den Monographien zu finden. Die höchsten Perzentilwerte der jeweiligen FER bei den Monographien haben in 23 von 57 Fällen FE unterhalb der höchstbewerteten erzielt, elfmal waren dies befriedigende, zwölfmal gute. Es fällt aber auch auf, dass bei den PR-Artikeln die jeweils besten FE über das genannte Feld der FER hinweg Spitzenwerte aufweisen. Diese Leistungen gehen in der Endnote der jeweiligen FE und erst recht in den aggregierten Werten der FER unter.

Das Problem der systematischen Benachteiligung von FE, die sich anderen Aufgaben gewidmet haben als der Publikation von PR-Artikeln, lässt sich kaum lösen. Eine Ausweitung des Kriterienkatalogs über den PR-Artikel hinaus erhöht die Komplexität und verringert die Vergleichbarkeit. Umso bedeutsamer wird dann das qualitative Urteil der Gutachter. Aber auch Gutachter sind überfordert, wenn sie zehn und noch mehr verschiedene Aufgabenstellungen von FE in ihrer Besonderheit würdigen sollen. Umso wahrscheinlicher übernehmen dann verfügbare quantitative Indikatoren die Funktion der Komplexitätsreduktion. Die Unsicherheit über die Validität verschiedener quantitativer Indikatoren wird dann leicht dadurch bewältigt, dass ein Indikator zur Leitschnur wird, weil er mehr Vertrauen als andere Indikatoren genießt. Diese Rolle hat der Indikator des PR-Artikels übernommen. Die Gutachter vertrauen dabei den Herausgebern von Fachzeitschriften mit hohem Impact, die wiederum den Gutachtern der Aufsätze Vertrauen schenken. Im Forschungsrating hat wiederum im Regelfall das Plenum den beiden Einzelgutachtern einer FE vertraut. Schließlich müssen die Rezipienten des Forschungsratings dem abschließenden Plenumsurteil Vertrauen entgegenbringen. Bis die Endnote beim Rezipienten ankommt, befindet man sich bei einem Vertrauen vierter Ordnung. Bei jeder der vier Stufen der Vertrauensstiftung wird Komplexität reduziert. Dadurch wird vieles ausgeblendet, was dennoch relevant sein könnte. Da indessen Wissenschaft in besonders hohem Maße auf die Ausschöpfung von Vielfalt und Kreativitätspotentialen angewiesen ist, um Fortschritte zu erzielen, kann der Prozess der mehrstufigen Komplexitätsreduktion den Spielraum für den Erkenntnisfortschritt empfindlich einengen. Je mehr Evaluationsverfahren den Prozess der Erkenntnis überlagern, umso mehr wird dieser Prozess durch den Zwang zur Komplexitätsreduktion behindert.

„Informed peer review“ verbindet quantitative Indikatoren mit qualitativem Urteil. Dadurch soll ermöglicht werden, eine größere Zahl von FE bzw. FER hinsichtlich ihrer Leistungen zu beurteilen und trotzdem nicht der einseitigen Selektion und Standardisierung nach festgelegten Indikatoren zu verfallen. Was sich auf den ersten Blick als elegante Lösung des Problems darstellt, hat auf den zweiten Blick dennoch seine Schwächen. Diese liegen vor allem in der Eichung der Urteile. Während im einfachen Peer Review z.B. eine Gutachtergruppe eine einzige FER evaluiert, mussten beim Forschungsrating 254 FE und 57 FER begutachtet werden. Bei einer solchen Zahl kann nicht sichergestellt werden, dass die einzelnen Gutachter durchgehend dieselben Maßstäbe in derselben Weise anwenden. Umso größere Bedeutung kommt dann der Seite des „informed“ peer review zu, das heißt den quantitativen Daten. Am ehesten möglich war die Korrektur von Fehlinformationen durch die Perzentilwerte anhand der Publikationslisten und der eingereichten Publikationen selbst. So konnte z.B. entdeckt werden, dass als PR-Artikel gezählte Publikationen gar keine PR-Artikel waren oder dass die vorgelegten Texte den Eindruck der quantitativen Daten nicht bestätigten. Das Problem, über den PR-Artikel hinaus eine breitere Vielfalt von Forschungsleistungen zu würdigen, konnte aber nicht gelöst werden, weil diesem Anliegen kaum in standardisierter, 254 FE und 57 FER vergleichbar machender Form entsprochen werden kann.

Die überragende Bedeutung, die dem PR-Artikel in der Qualitätsbestimmung zugekommen ist, spiegelt nicht die Publikationspraxis in der deutschen Soziologie wider. Während der PR-Artikel etwa zwei Drittel der Varianz in den Qualitätsurteilen im Forschungsrating erklärt hat, waren nur ein Drittel der erfassten Publikationen Fachzeitschriftenaufsätze, wovon wiederum nur ein Drittel in begutachteten Fachzeitschriften herauskamen, die in internationalen Datenbanken erfasst sind. Aber auch Monographien trugen nur 7 % zum gesamten Publikationsaufkommen bei (Wissenschaftsrat 2008 a: 445). Demnach waren ca. 11 % der Publikationen von entscheidender Bedeutung für die Qualitätsurteile. Statt die Publikationspraxis widerzuspiegeln, wird das Forschungsrating einen erheblichen normativen Druck in die Richtung der vermehrten Publikation von PR-Artikeln ausüben. Dasselbe gilt – zum großen Teil bei derselben Textgattung – für die Prämierung englischsprachiger Publikationen. Nur 15,6 % der im Forschungsrating erfassten Publikationen sind im nicht-deutschsprachigen Ausland erschienen (Wissenschaftsrat 2008 a: 445). In der Tat ist unübersehbar, dass die deutsche Soziologie hier Nachholbedarf hat, wenn sie international präsenter sein will, als es gegenwärtig der Fall ist. Schließlich hat sie ein großes Erbe fortzuführen (und nicht nur zu verwalten). Die Soziologie teilt dieses Schicksal mit allen anderen geistes- und sozialwissenschaftlichen Disziplinen. Als Ursache für dieses Defizit sehe ich maßgeblich die oligarchischen Strukturen, die insbesondere im Vergleich zu den USA Innovationen und deren Institutionalisierung in der Breite behindern. Gäbe es anstelle von Instituten und Lehrstühlen größere Departments, dann wäre die deutsche Soziologie auch international in referierten Fachzeitschriften präsenter. Statt im Schreiben von Drittmittelanträgen und -berichten verheizt zu werden, könnten viel mehr junge Wissenschaftlerinnen und Wissenschaftler den Aufwand betreiben, der für die Publikation eines PR-Artikels erforderlich ist. Gleichwohl wäre es der Vielfalt des soziologischen Wissens abträglich, wenn sich alles um den PR-Artikel drehen würde. Alles bedarf seines richtigen Maßes.

3 Kriterium II: Impact

Reputation

Ein Problem bei der Messung des Impacts ergab sich dadurch, dass Reputation zwar ausdrücklich ermittelt werden sollte und darüber auch Informationen vorlagen, die Bewertungsgruppe mit diesen Informationen aber im Vergleich zu den Perzentilwerten nur schwer etwas anfangen konnte. Ein Weg, das Problem zu lösen, bestünde darin, entscheidende Indikatoren in Zahlen und letztlich in einen Skalenwert umzurechnen, der dann neben die Perzentilwerte

der anderen Indikatoren gestellt werden kann. Herausgeberschaften und Beiratsmitgliedschaften bei Fachzeitschriften als Pendant zum besonderen Gewicht der PR-Aufsätze und DFG-Fachgutachter-Positionen könnten sich als Indikatoren für Reputation eignen, worauf im Abschlussbericht der Bewertungsgruppe verwiesen wird (Wissenschaftsrat 2008 a: 399-400). Auch im britischen Research Assessment Exercise (RAE) wird darauf zurückgegriffen.

Sicherlich werden wie bei allen Indikatoren bei dieser Operationalisierung von Reputation Entscheidungen getroffen, die mögliche andere Zeichen von Reputation nicht erfassen. Dazu zählen z.B. Mitgliedschaften in Vorständen und Beiräten der Forschungsorganisation wie z.B. die in der GESIS zusammengeschlossenen Einrichtungen *IZ*, *ZUMA* und *Zentralarchiv für empirische Sozialforschung* oder Mitgliedschaften in Fachbeiräten von Max-Planck-Instituten. Die Veränderungen, die sich aus der Berücksichtigung solcher Positionen ergäben, halten sich jedoch in engen, zahlenmäßig kaum ins Gewicht fallenden Grenzen. Es soll auch nur beispielhaft gezeigt werden, wie sich diese Operationalisierung von Reputation auf das Ergebnis auswirkt. Dabei wird den ermittelten Punktwerten kein auf die Gesamtmenge bezogener Perzentilwert zugeordnet, sondern ein geschätzter Skalenwert, der in etwa den Abstand der Punktwerte wiedergibt (Tab. 2).

Für Herausgeberschaften, wiss. Beiräte / Editorial Boards von Fachzeitschriften und DFG-Fachgutachterpositionen werden die folgenden Punktwerte vergeben:

Engelschsprachige Kernzeitschrift	Editor (E) Associate / Consulting Editor (AE / CE)	40 20
Engelschsprachige Spezialzeitschrift	Editor (E) Associate / Consulting Editor (AE / CE)	20 10
Deutschsprachige Kernzeitschrift	Herausgeber / Mitherausgeber (E)	30
Deutschsprachige Spezialzeitschrift	Herausgeber / Mitherausgeber (E)	10
nicht-deutsche, nicht-englischsprachige Fachzeitschrift	Herausgeber / Mitherausgeber (E)	10
Wissenschaftlicher Beirat / Editorial Board einer Kernzeitschrift (B)		5
Wissenschaftlicher Beirat / Editorial Board einer Spezialzeitschrift (B)		5
DFG-Fachgutachter (FG, DFG), DFG-Bewilligungsausschuss		30

Bei Hauszeitschriften wird E nur einmal pro FER gezählt.

Über den Status als Kern- oder Spezialzeitschrift entscheidet der allgemeine bzw. spezielle Titel der Zeitschrift. So ist z.B. die Zeitschrift für Soziologie eine Kernzeitschrift, dagegen ist Das Gesundheitswesen eine Spezialzeitschrift. Umrechnung von Punkt- in Skalenwerte:

$E + AE / CE + DFG + B$				
201	-	250	=	100
151	-	200	=	90
121	-	150	=	80

101	-	120	=	70
81	-	100	=	60
61	-	80	=	50
41	-	60	=	40
21	-	40	=	30
11	-	20	=	20
5	-	10	=	10
0	-	4	=	0

Tabelle 2: Herausgeberschaften und Beiratsmitgliedschaften bei Fachzeitschriften sowie DFG-Fachgutachter- und Ausschusspositionen

Universität	Punkt- wert	Zugeordne- ter Skalen- wert	Universität	Punkt- wert	Zugeordne- ter Skalen- wert
Bielefeld	240	100	Jena	35	30
Bamberg	205	100	Köln	35	30
Konstanz	125	80	Aachen	30	30
Kassel	120	80	Frankfurt / Main	30	30
Göttingen	100	60	Magdeburg	30	30
HU Berlin	95	60	Mannheim	30	30
MPIfG, Köln	95	60	Stuttgart	30	30
Erlangen-Nürnb.	90	60	Wuppertal	30	30
LMU München	85	60	Darmstadt	25	30
WZB	85	60	Osnabrück	25	30
Erfurt	80	50	Gießen	20	20
U Bremen	75	50	Rostock	20	20
FU Hagen	75	50	SOEP	20	20
TU Berlin	70	50	Eichstätt	15	20
Dortmund	65	50	DSPH Köln	15	20
Tübingen	65	50	Dresden	10	10
Duisburg-Essen	60	40	U Hamburg	10	10
FU Berlin	55	40	Münster	10	10
Düsseldorf	55	40	JU Bremen	5	10
Freiburg	55	40	UBW München	5	10
Leipzig	50	40	Trier	5	10
Mainz	50	40	Augsburg	0	0
Chemnitz	45	40	Frankfurt / Oder	0	0
Halle-Wittenberg	45	40	UBW Hamburg	0	0
Hannover	40	30	Kiel	0	0
Bayreuth	35	30	Potsdam	0	0
Bochum	35	30	Siegen	0	0

Universität	Punkt- wert	Zugeordne- ter Skalen- wert	Universität	Punkt- wert	Zugeordne- ter Skalen- wert
Heidelberg	35	30			

Peer Reviewed Articles

Ein gravierendes Problem bei der Feststellung des Impacts wie auch der Effizienz war die Unzuverlässigkeit des Indikators „Peer Reviewed Article“. Viele nicht begutachtete Texte sind unter diese Kategorie gefallen. Dazu zählen Kurzeassays in allgemeinen Publikumszeitschriften, Rezensionen, Kommentare und Editoriale. Diese Ungenauigkeit streut nicht gleichmäßig, sondern ungleichmäßig über die FER. Es gibt FER, deren Daten in der Tat annähernd das widerspiegeln, was unter einem Peer Reviewed Article (PR-Artikel) verstanden wird, während andere zwar zahlenmäßig mithalten, aber inhaltlich weit davon entfernt sind. Dazu kommen noch gravierende Differenzen derart, dass FER überwiegend in gewichtigen Fachzeitschriften im Kern der Soziologie publizieren, andere dagegen in sehr randständigen. Auch in dieser Hinsicht ist das Profil sehr ungleich. Es gibt FER, bei denen sich z.B. Publikationen in der *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (KZSS), der *Zeitschrift für Soziologie* (ZfS) oder der *European Sociological Review* (ESR) häufen, andere sind offensichtlich spezialisiert auf Zeitschriften wie *Das Gesundheitswesen*, *Zeitschrift für Sozialreform* und dgl. Wieder andere publizieren häufig in Publikumszeitschriften wie *Merkur*, *Internationale Politik*, *Commentaire*, *Dissent* und dgl., aber nicht in den soziologischen Kernzeitschriften. Dazu kommt noch die Publikation in der Hauszeitschrift, d.h. der Fachzeitschrift, die von der FER selbst herausgegeben wird. Was die mangelnde Trennschärfe betrifft, ist es irreführend, wenn u.a. die folgenden Zeitschriften nicht als PR zählen:

- Zeitschrift für Bevölkerungswissenschaft
- Zeitschrift für internationale Beziehungen
- Sonderhefte der KZSS, *Sozialen Welt*, *ZfS*

Dafür sind recht häufig folgende Zeitschriften dabei, die nicht im Entferntesten einen PR-Artikel veröffentlichen:

Merkur

Internationale Politik

Soziologie

Soziologische Revue

Um mehr Licht in dieses Dunkel zu bringen, wurde für die Gruppe der 16 im Impact mit der Note 4 oder 5 bewerteten FER plus 10 weitere FER eine Prüfung der Publikationslisten durchgeführt. Dabei wurden *nur* die als PR gewichteten Artikel sortiert, es wurde jedoch nicht geprüft, welche nicht als PR gewerteten Artikel eigentlich aufgenommen werden müssten. Auch dadurch ergäbe sich wahrscheinlich eine weitere Differenzierung zwischen FER, an denen viel in gewichtigen Fachzeitschriften publiziert wird, und solchen, an denen das weniger der Fall ist. Dabei zeigen sich allerdings auch hier Größeneffekte. In der Gruppe der überprüften 26 FER gibt es in dieser Hinsicht deutliche Unterschiede (Tab. 3). Das Ergebnis der Prüfung ist im Folgenden notiert:

Peer reviewed Articles

FER	= Forschungseinrichtung
K	= Kernzeitschrift (zählt ganz)
KH	= Kernzeitschrift im Haus (zählt nur zur Hälfte)
S	= Spezialzeitschrift (zählt ganz)
SH	= Spezialzeitschrift im Haus (zählt nur zur Hälfte)
BK	= Besprechung, Kommentar, Replik (zählt nicht)
P	= Allgemeine Publikumszeitschrift (zählt nicht)
E	= Editoriale, Ehrungen, Nachrufe (zählt nicht)
N.N.	= Name nicht auf Namensliste der FER (zählt nicht)
A/B/C	= Alte Zahl / Neue Zahl ohne BK, P, E / Neue Zahl ohne BK, P, E und ½, KH, ½ SH

Tabelle 3: Peer Reviewed Articles

FER	A / B / C	FER	A / B / C	FER	A / B / C
	Anzahl PR-Artikel		Anzahl PR-Artikel		Anzahl PR-Artikel
Bielefeld	81 / 72 / 69	Göttingen	39 / 37	Konstanz	31 / 25
K = 30		K = 15		K = 18	
KH = 6		S = 22		S = 7	
S = 35		B = 2		BK = 2	
SH = 1				P = 4	
BK = 5					
P = 4					
E = 0					
Düsseldorf	52 / 51	HU Berlin	59 / 39 / 33	Bochum	25 / 24
K = 6		K = 9		K = 11	
S = 45		KH = 13		S = 13	
BK = 1		S = 17		PI = 1	
		BK = 2			
		P = 12			
		E = 6			
Mannheim	61 / 51	Jena	43 / 38 / 32	Augsburg	28 / 22
K = 30		K = 11		K = 5	
S = 21		S = 14		KH = 6	
BK = 6		SH = 13		S = 14	
P = 3		BK = 5		BK = 1	
Doppelt = 2				P = 1	
				E = 1	
Hamburg	52 / 50	IU Bremen	32 / 31	Siegen	18 / 16
K = 15		K = 20		K = 10	
S = 35		S = 11		S = 6	
BK = 1		P = 1		P = 1	
P = 1				E = 1	

FER	A / B / C	FER	A / B / C	FER	A / B / C
	Anzahl PR-Artikel		Anzahl PR-Artikel		Anzahl PR-Artikel
SOEP	49 / 49	Köln	34 / 31	Gießen	17 / 15
K = 23 (davon 12 Schmollers Jahrbuch)		K = 14		K = 9	
S = 26		KH = 6		S = 6	
		S = 11		P = 2	
		BK = 2			
		E = 1			
Bamberg	49 / 47	Bremen	31 / 30	TU Berlin	15 / 14
K = 27		K = 16		K = 6	
S = 20		S = 14		S = 8	
BK = 1		BK = 0		P = 1	
		P = 1			
		E = 0			
FU Berlin	51 / 46	Halle-Wittenberg	31 / 30	Kassel	21 / 11
K = 14		K = 7		K = 4	
S = 32		S = 23		S = 7	
BK = 2		P = 1		BK = 5	
P = 3				P = 4	
E = 0				E = 1	
LMU München	66 / 48 / 43	Frankfurt / Main	34 / 27	Frankfurt / Oder	10 / 10
K = 22		K = 8		K = 4	
KH = 10		S = 19		S = 6	
S = 16		BK = 1			
BK = 9		P = 6			
P = 8					
E = 1					
MPIfG	48 / 39	Erfurt	29 / 25		
K = 24		K = 8			
S = 15		S = 17			
BK = 6		BK = 1			
E = 3		P = 3			

Nach der Auswertung kann man die in Tabelle 4 wiedergegebenen Unterschiede feststellen: In der ersten Spalte nach den Standortnamen wird die Zahl der Kern- / Spezialzeitschriften genannt. In der zweiten Spalte wird ein Durchschnittswert aus Kernzeitschriften mit doppeltem Gewicht und Spezialzeitschriften mit einfachem Gewicht ermittelt. Den Zahlen wird in der dritten Spalte ein Skalenwert zwischen 0 und 100 zugeordnet, der in etwa die Abstände wiedergibt (Tab. 4).

Tabelle 4: Präsenz der aktivsten FER in Kern- und Spezialzeitschriften

	Kern / Spezial	Kern / Kern / Spezial	Zugeordneter Skalenwert
Bielefeld	33 / 36	34	100
Mannheim	30 / 21	27	99
Bamberg	27 / 20	25	98
SOEP	23 / 26	24	97
MPIfG	24 / 15	21	94
LMU München	22 / 16	20	93
Hamburg	15 / (35) 24 (23 Med. Soz.)	18	80
FU Berlin	14 / (32) 27 (11 Med. Soz.)	18	80
JU Bremen	20 / 11	17	78
Göttingen	15 / 22	17	78
HU Berlin	16 / 17	16	77
Köln	17 / 11	15	76
Bremen	16 / 14	15	76
Konstanz	18 / 7	14	75
Jena	11 / 21	14	75
Bochum	11 / 13	12	60
Frankfurt / Main	8 / 19	12	60
Halle-Wittenberg	7 / 22	12	60
Düsseldorf	6 / (45) 23 (42 Med. Soz.)	12	60
Erfurt	8/17	11	60
Augsburg	8/14	10	55
Siegen	10/6	9	45
Gießen	9/6	8	45
TU Berlin	6/8	7	40
Kassel	4/7	5	30
Frankfurt / Oder	4/6	5	30

Anmerkung: Bei den für U Hamburg, FU Berlin und U Düsseldorf in Klammern genannten Aufsätzen in Spezialzeitschriften handelt es sich jeweils um 23, elf bzw. 42 Aufsätze in Medizinischer Soziologie. Aufgrund der im Umfang stark abweichenden Publikationspraxis dieses Fachs und des nur begrenzten Impacts auf die Soziologie wurde in Anlehnung an die Praxis im FR nur die Hälfte gezählt.

Da mehr als die Hälfte der Publikationen des *Sozioökonomischen Panels* (SOEP) am *Deutschen Institut für Wirtschaftsforschung* (DIW) in ökonomischen Fachorganen erschienen sind, ist es fraglich, ob der errechnete Impactwert uneingeschränkt auf die Soziologie bezogen werden kann. Die nur halbe Zählung der Fachzeitschriftenaufsätze der Medizinischen Soziologie bringt ein Vergleichsproblem mit sich, das letztlich nicht wirklich befriedigend gelöst werden konnte, da es ja Aufgabe der Medizinischen Soziologie ist, auf die Medizin einzuwirken.

Ergebnisse

Im Folgenden stellt Tabelle 7 die Ergebnisse dar. Es werden die Perzentilwerte der folgenden Indikatoren aus dem Forschungsrating übernommen:

- (1) Publikationen gesamt
- (2) Aufsätze in PR-Zeitschriften (ggfs. für die ersten 15 plus weitere elf FER korrigiert)
- (3) Nicht-deutsche Publikationen
- (4) Drittmittelfinanzierte Mitarbeiter

Hinzugefügt werden folgende Indikatoren:

- (5) Reputation 1: Herausgeberschaften, Beiräte, DFG-Fachgutachter mit Skalenwerten (Rep.)
- (6) Reputation 2: Wert der drei besten FE einer FER (FQ1, zugeordneter Skalenwert auf Skala von 0 bis 100, FQ2)

Es werden vier Werte eines neuen mittleren Skalenwerts angegeben:

- a) mittlerer Skalenwert aus (1) bis (5), Perzentilwerte für PR-Artikel für erste 15 plus weitere elf FER korrigiert nach PR 1
- b) mittlerer Skalenwert aus (1) bis (5), Perzentilwerte für PR-Artikel durchgehend nicht korrigiert (PR 2)
- c) mittlerer Skalenwert aus (2) bis (6), Perzentilwerte für PR-Artikel für erste 15 plus weitere elf FER korrigiert nach PR 1
- d) mittlerer Skalenwert aus (1) bis (6), Perzentilwerte für PR-Artikel für erste 15 plus weitere elf FER korrigiert nach PR 1

Außerdem kann als Kriterium dienen, wie häufig der Perzentilwert 90+ für 5 (exzellent) bzw. 70+ für 4 (sehr gut) erreicht wird. Die Häufigkeit wird in den Spalten 70+ und 90+ angegeben.

Wie schon die FQ lässt sich auch der Impact nicht nach einer einzigen Maßzahl ermitteln. Jeder einzelne Indikator ist einseitig. Um nur annäherungsweise der Komplexität der realen Verhältnisse gerecht zu werden, müssen mehrere Indikatoren nebeneinander gestellt werden. Tabelle 7 enthält dementsprechend die mittleren Skalenwerte a) bis d) zum Impact. Dabei liegt der Wert a) der Reihung der FER in Spalte 1 zugrunde. Die Werte b) bis d) können zur ersten Relativierung genutzt werden. Zusätzlich helfen die folgenden Werte bei der Interpretation und Relativierung der Reihung: die Häufigkeit des Überschreitens der Perzentilwerte 70 bzw. 90, der korrigierte und der nicht korrigierte Wert für PR-Artikel, der Reputationswert, die FQ der besten drei FE einer FR, der Skalenwert der FQ, der mittlere Perzentilwert der Impactmessung im Forschungsrating und die im Forschungsrating vergebene Note für Impact (Tab. 5).

Tabelle 5: Impact

FER	70+	90+	PR1	PR2	Rep.	FQ 1		FQ 2	Mittlerer Skal.				Mittl. Perz. FR	Note FR
									a	b	c	d		
Bielefeld	5	5	(100)	100	100	12	444	80	98	99	95	96	99	5
Bamberg	5	4	(98)	87	100	14	554	100	93	91	97	95	89	4
Göttingen	4	3	(78)	81	60	11	443	70	83	84	79	81	90	5

FER	70+	90+	PR1	PR2	Rep.	FQ 1		FQ 2	Mittlerer Skal.				Mittl. Perz. FR	Note FR
									a	b	c	d		
HU Berlin	5	1	(77)	95	60	11	443	70	81	85	75	79	92	5
Konstanz	4	1	(75)	76	80	11	533	70	80	71	80	78	80	4
Halle-Wittenberg	2	2	(60)	71	40	9	333	50	73	75	63	69	81	4
Frankfurt / Main	3	2	(60)	79	30	10	433	60	72	76	65	70	85	4
LMU München	3	2	(93)	93	60	13	544	90	71	71	72	74	77	5
MPIfG	3	2	(94)	84	60	(15)	5	100	69	67	74	74	72	4
FU Berlin	4	0	(80)	89	40	11	533	70	69	71	67	70	81	4
Mannheim	2	2	(99)	96	30	(15)	5	100	69	68	77	74	78	5
Jena	3	0	(75)	82	30	11	443	70	69	71	66	69	81	4
Hamburg	4	0	(80)	92	10	(9)	33	50	69	72	62	66	88	4
U Bremen	2	0	(76)	71	50	11	443	70	65	64	67	66	68	4
Bochum	2	0	(60)	60	30	10	433	60	64	64	60	63	70	3
WZB	1	0		66	60	(10)	43	60	64	64	63	63	65	3
Duisburg-Essen	2	0		62	40	9	333	50	62	62	57	60	66	3
SOEP	3	1	(97)	87	20	(15)	5	100	61	59	63	67	72	4
U Köln	2	0	(76)	76	30	11	443	70	61	61	60	63	70	3
Kassel	3	1	(30)	54	70	10	433	60	59	74	64	67	71	3
Erfurt	1	0	(60)	66	50	(11)	53	70	56	54	63	58	59	3
Dortmund	2	0		35	50	9	333	50	57	57	52	56	54	3
TU Berlin	0	0	(40)	40	50	10	433	60	52	52	50	53	50	3
TU Chemnitz	0	0		54	40	10	433	60	52	52	53	53	55	3
Düsseldorf	1	0	(60)	92	40	(12)	44	80	51	57	57	56	68	4
Freiburg				51	40	(9)	33	50	47	41	43	48	49	3
Marburg				58	30	9	333	50	47	45	45	47	52	3
Augsburg			(55)	62	0	9	333	50	46	46	42	45	58	3
Gießen			(45)	46	20	8	332	40	46	46	45	45	51	3
JU Bremen	2	0	(78)	74	10	12	444	80	44	42	52	50	56	3
TU Darmstadt				40	30	(8)	32	40	43	47	43	42	45	3
TU Dresden				35	10	9	432	50	42	40	41	43	47	3
Osnabrück				3	30	8	332	40	41	35	42	41	43	1
Wuppertal				46	30	(9)	3	50	40	38	41	42	44	3
Trier				49	10	10	433	60	36	34	38	40	44	3
Erlangen-Nürnberg				8	60	(9)	3	50	36	36	40	38	25	3

FER	70+	90+	PR1	PR2	Rep.	FQ 1		FQ 2	Mittlerer Skal.				Mittl. Perz. FR	Note FR
									a	b	c	d		
Siegen			(45)	44	0	11	443	70	34	34	40	40	42	3
Heidelberg				54	30	11	443	70	32	32	40	38	37	3
Rostock				40	20	(8)	42	40	32	32	30	33	36	3
Magdeburg				11	30	7	322	30	32	30	32	31	28	2
Stuttgart				18	30	(9)	33	50	31	31	38	34	29	2
Tübingen				30	50	8	431	40	31	27	36	33	27	2
Hagen				25	50	10	433	60	30	30	33	35	25	2
DSPH Köln				3	20	(9)	33	50	30	30	37	34	33	2
Hannover				18	30	(9)	3	50	24	24	32	29	22	2
Mainz				18	40	9	432	50	24	22	29	29	20	2
Aachen				32	30	(9)	3	50	21	21	29	26	21	2
Eichstätt				27	20	8	431	40	21	19	21	23	21	2
Potsdam				8	0	(7)	32	30	19	19	24	21	21	1
Bayreuth				5	30	7	322	30	18	18	19	20	13	1
Kiel				18	0	(9)	33	50	17	17	25	23	21	2
UBW München				11	10	(10)	43	60	17	15	25	24	21	2
Frankfurt / Oder			(30)	27	0	(12)	4	80	16	16	30	27	22	2
Bonn				18	0	(9)	3	50	12	12	21	18	15	2
UBW Hamburg				35	0	8	422	40	10	10	16	15	16	1
Leipzig				68	30	10	442	60	43	45	47	46	51	n.b.
Münster				18	10	(9)	3	50	13	11	23	20	15	n.b.

Legende: **70+** = Häufigkeit der Überschreitung des Perzentilwerts 70, Indikatoren 1-5; **90+** = Häufigkeit der Überschreitung des Perzentilwerts 90, Indikatoren 1-5; **PR1** = Korrigierter Skalenwert auf Skala 0-100 für PR-Artikel; **PR2** = nicht korrigierter Perzentilwert der PR-Artikel im Forschungsrating; **Rep.** = Skalenwert auf Skala von 0-100 für Reputation, ermittelt aus Herausgeberschaften und Beiratsmitgliedschaften in Fachzeitschriften sowie DFG-Fachgutachter- und DFG-Ausschusspositionen; **FQ1** = Einzelnoten für Forschungsqualität und deren Summe der jeweils besten drei FE einer FER, in Klammern geschätzte Werte, wenn keine Unterteilung in mehrere FE vorliegt; **FQ2** = zugeordneter Skalenwert auf Skala von 0-100 der in FQ1 ermittelten Forschungsqualität; **Mittlerer Skal.** = neu errechneter mittlerer Skalenwert für Impact auf Skala von 0 bis 100; **Mittleres Perz. FR** = im Forschungsrating ermittelter mittlerer Perzentilwert für Impact; **Note FR** = Note im Forschungsrating.

Die in Tabelle 5 wiedergegebenen Ergebnisse zeigen, dass man mit verschiedenen Indikatoren zu unterschiedlichen Bewertungen gelangen kann und kleine Veränderungen zu veränderten Einstufungen führen. Außerdem können qualitative Einschätzungen zahlenmäßige Ergebnisse relativieren und eine andere Einstufung als die Zahlen vorgeben, rechtfertigen. Dabei ist es allerdings auch möglich, dass bestimmten Indikatoren über verschiedene Fälle hinweg nicht immer genau das gleiche Gewicht gegeben wird. Auch das besondere Profil einer FER kann Anlass für die geringere Gewichtung eines Indikators (z.B. Drittmittel) sein.

Ganz unverkennbar ist der ermittelte Impact einer FER hoch mit deren Größe korreliert. Mit der Größe wachsen die Chancen, mit Drittmiteinnahmen, Publikationsoutput und Ämtern in der wissenschaftlichen Gemeinschaft präsent zu sein und einen entsprechend großen Einfluss auszuüben. Die in der Spalte „Rep.“ wiedergegebenen Reputationswerte nach Herausgeberschaften, Beiratsmitgliedschaften und DFG-Tätigkeiten drängen den Größeneffekt etwas in den Hintergrund, wie sich unmittelbar an den Diskrepanzen zur Spalte „Mittleres Perz. FR“ erkennen lässt.

4 Kriterium III: Effizienz

Die Überlegungen zum Kriterium Impact lassen sich auch auf das Kriterium Effizienz übertragen. Hier fällt auf, dass alle vier im Forschungsrating mit der Note 5 als exzellent eingestuften FER sehr niedrige Perzentilwerte bei den drittmittelfinanzierten Mitarbeitern erreichten (25 %, 16 %, 5 %, 14 %). Ihre relativ hohen Perzentilwerte bei den Publikationen ergeben sich demnach schon aus der niedrigen Zahl von angerechneten drittmittelfinanzierten Mitarbeitern. Es wurden nämlich diese Werte und nicht die Werte pro grundfinanziertem Personal zugrunde gelegt. Dagegen weisen fast alle FER in der Notenstufe 4 (sehr gut) höhere Perzentilwerte bei diesem Indikator auf. Es wurde demnach sehr niedrige Drittmittelaktivität mit hohen Effizienzwerten belohnt. Das verweist darauf, dass sich hohe Drittmittelsummen negativ auf die Publikationsproduktivität auswirken können (Münch 2007). Es gab drei Publikationsindikatoren, aber nur einen Drittmittelindikator. Offensichtlich wurde bei der Einzelfallprüfung der an sich besonders gewichtete Indikator „Drittmittel-Mitarbeiter“, von der vorgegebenen Maßgabe abweichend, als nachrangig eingestuft. Es fällt außerdem auf, dass einige sehr große und auch drittmittelstarke FER nicht mehr als mittlere oder sogar noch schwächere Effizienzwerte erreichten, dagegen einige kleine FER sehr gute oder sogar exzellente Werte. Hier zeigen sich die Nachteile der Überinvestition von Forschungsmitteln an sehr großen Standorten.

Es gab zwei Effizienzmaße: ein Effizienzmaß A, bezogen auf das gesamte Personal einschließlich Drittmittel-Mitarbeiter, und ein Effizienzmaß B, allein bezogen auf das grundfinanzierte Personal. Die Unterschiede im Ergebnis waren erheblich. Man könnte deshalb für das Effizienzmaß einen mittleren Perzentilwert aus A und B errechnen. Auf der Basis der unter „Impact“ erfolgten Auszählung der PR-Aufsätze wurden die Effizienzwerte A und B für diesen Indikator bei den FER des oberen Drittels neu berechnet (daneben stehen zugeordnete Skalenwerte auf einer Skala von 0 bis 100). Alte Werte ergeben sich aus den unkorrigierten PR-Artikel-Daten, neue Werte aus den korrigierten (gewichtet nach Kern- und Spezialzeitschriften).

Die besonders starke Korrektur beim SOEP ist darauf zurückzuführen, dass dem alten Wert eine rechnerisch falsche doppelte Begünstigung durch den Faktor 0,5 für Professoren und 0,75 für Mitarbeiter zugrunde lag. Das SOEP wurde mit der Begründung von Serviceleistungen wie die Universitätseinrichtungen behandelt. Bei den beiden anderen außeruniversitären FER, dem Wissenschaftszentrum für Sozialforschung Berlin (WZB) und dem Max Planck-Institut für Gesellschaftsforschung in Köln (MPIfG) wurden die Professoren mit 0,9 und die Mitarbeiter mit 1,0 berechnet (Tab. 6).

Tabelle 6: Effizienzwerte

	Effizienz							
	A				B			
	Alt	Perz.	Neu	Skal.	alt	Perz.	Neu	Skal.
Düsseldorf	13,35	98	13,07	100	10,15	98	9,94	100
IU Bremen	10,52	95	10,2	96	9,85	96	9,54	98

	Effizienz							
	A				B			
	Alt	Perz.	Neu	Skal.	alt	Perz.	Neu	Skal.
SOEP	19,37	100	10,77	98	14	100	7,78	96
Mannheim	8,75	93	7,2	93	6,89	93	5,76	93
Erfurt	7,1	89	6,12	88	6,44	91	5,55	91
Bamberg	5,02	75	4,82	83	5,23	89	5,01	90
Frankfurt / Oder	5,45	82	5,45	86	5	85	5	89
LMU München	10,87	96	6,48	90	7,38	95	4,41	86
Augsburg	5,62	84	4,41	80	5,09	86	4	82
MPIfG	7,56	91	6,14	89	4,85	81	3,94	81
Jena	5,32	81	3,96	75	5,21	88	3,88	79
U Bremen	4,38	72	4,24	72	3,26	61	3,16	61
Konstanz	3,6	54	2,65	40	4,25	77	3,13	60
Gießen	3,71	56	3,28	50	3,4	63	3	58
Göttingen	3,37	49	3,2	49	2,91	51	2,76	51
HU Berlin	5,31	79	2,97	45	4,21	75	2,36	40
TU Berlin	3,14	44	2,93	44	2,51	39	2,34	39
Kassel	3,17	46	1,66	15	2,71	46	1,42	17

Ergebnisse

Im Folgenden wurden für das obere Drittel die mittleren Skalenwerte für die Effizienzwerte A und B neu berechnet. Außerdem wurden die Spalte zu den Perzentilwerten der DM-Mitarbeiter und die Note des Forschungsratings übernommen. Auch bei den Effizienzwerten gilt: Die Welt sieht mit unterschiedlichen Indikatoren betrachtet im Detail und in Einzelfällen jeweils anders aus. Dabei wird auch hier nicht der Anspruch auf eine „wahre“ Rangfolge erhoben. Es soll nur deutlich werden, dass auch Effizienz nicht einfach da ist, sondern durch den Kriterienkatalog und den Messvorgang konstruiert wird (Tab. 7).

Tabelle 7: Effizienz

FER	Effiz. A		Effiz. B		Mittel A / B		DM-Mit.	Note FR
	alt	neu	alt	neu	alt	neu		
Bamberg	82	84	90	90	86	87	95	4
Jena	79	78	86	84	83	81	93	4
SOEP	81	81	81	80	81	80	25	5
Erfurt	79	79	80	80	80	80	68	4
JU Bremen	74	74	77	78	76	76	16	5
Düsseldorf	76	76	74	75	75	75	28	4
LMU München	74	73	72	70	73	72	5	5
Göttingen	72	72	71	71	72	72	93	4
Konstanz	66	62	82	77	74	70	100	3
Augsburg	69	68	71	70	70	69	86	4
Mannheim	69	69	68	68	69	69	51	4

FER	Effiz. A		Effiz. B		Mittel A / B		DM-Mit.	Note FR
	alt	neu	alt	neu	alt	neu		
Gießen	68	67	72	71	70	69	70	3
U Bremen	61	61	77	77	69	69	44	4
Kassel	73	65	73	68	73	67	61	3
Siegen	65	65	69	69	67	67	75	3
HU Berlin	74	66	71	63	73	65	54	4
MPIfG	70	69	58	58	64	64	14	5
Frankfurt / Oder	62	63	64	64	63	64	30	4
Bochum	62		65		64		74	3
Halle-Wittenberg	60		67		64		98	3
TU Berlin	63	63	61	61	62	62	47	3
Freiburg	57		60		59		53	3
Osnabrück	53		64		59		77	2
Dresden	58		57		58		60	3
Bielefeld	59		55		57		56	3
Chemnitz	57		53		55		64	3
FU Berlin	59		48		54		30	3
(Leipzig)	56		52		54		18	n.b.
Tübingen	51		55		53		79	3
Bonn	48		53		51		42	3
Heidelberg	50		45		48		9	3
Wuppertal	46		50		48		20	3
Marburg	44		47		46		12	3
Eichstätt	47		42		45		4	3
Köln	49		41		45		37	3
Frankfurt / Main	39		45		42		37	2
Dortmund	37		44		41		88	2
Stuttgart	42		40		41		72	3
Darmstadt	40		40		40		67	3
DSPH Köln	38		42		40		91	2
Magdeburg	39		40		40		70	2
Rostock	37		40		39		58	3
Hamburg	34		39		37		33	3
WZB	40		32		36		49	3
Aachen	37		33		35		32	3
Erlangen	35		33		34		39	3
Duisburg-Essen	34		32		33		40	3
UBW München	31		35		33		64	2
Kiel	29		27		28		26	2
Potsdam	28		27		28		89	1
Hagen	27		24		26		7	2
Bayreuth	24		21		23		26	2

FER	Effiz. A		Effiz. B		Mittel A / B		DM-Mit.	Note FR
	alt	neu	alt	neu	alt	neu		
UBW Hamburg	22		18		20		2	1
Hannover	19		19		19		53	1
Trier	20		15		18		21	2
Münster	16		17		17		35	n.b.
Mainz	18		14		16		20	2

Auch die Effizienzmessung hat ihre Schattenseite. Sie fördert die Strategie der Publikation kleinster Einheiten aus größeren Datenmengen, die Publikation in immer größeren Autoren-gemeinschaften, um mit geringstmöglichem Aufwand die größtmögliche Präsenz zu erreichen. In den Naturwissenschaften ist das längst Realität, in den vernaturwissenschaftlichen Grenz-gebieten der Humanwissenschaften wie der Psychologie breitet sich diese Strategie massiv aus. Zehn Autoren produzieren fünfseitige Aufsätze, in denen die Ergebnisse eines Experi-ments berichtet werden. Von den Siegern des Effizienzwetzbewerbs zu lernen, heißt, genau solche Strategien zu verfolgen, bei denen die Publikation kleinstmöglicher Einheiten mit der größtmöglichen Autorenzahl zum Selbstzweck wird, völlig ungeachtet der Frage, welche Er-kenntnisse auf diese Weise entstehen. Von dieser Schattenseite aus betrachtet, wäre es für die Kreativität der Forscher und die Wissensrevolution schädlich, wenn alle den Strategien der Tabellenführer in der Effizienztabelle nacheifern würden. Die Erzeugung solcher Tabellen durch Forschungsrating hat jedoch genau diesen Effekt. Sie übt einen normativen Druck auf die Forscher aus, es so zu machen wie die Sieger. Die Rhetorik des „Benchmarking“, dass von den „Besten der Besten“ gelernt werden soll, um besser zu werden, setzt genau diesen Maßstab.

5 Kriterium IV: Nachwuchsförderung

Hier bestanden große Schwierigkeiten, das Datenmaterial zu bewerten und in eine Rangord-nung zu bringen. In die Bewertung einbezogen wurden die folgenden Daten:

- Promotionen absolut
- Promotionen pro Professor
- finanzierte Doktoranden (stärker gewichtet)
- Promotionsprogramme (stärker gewichtet)
- angenommene Erstrufe (stärker gewichtet)

Analog zur Darstellung von FQ, Impact und Effizienz sollen auch bei der Nachwuchsförderung mehrere Maßzahlen sichtbar nebeneinander gestellt werden. Aus den im Forschungsrating entscheidenden Perzentilwerten für die absolute Zahl von Promotionen, die Promotionen pro Professor, die finanzierten Doktoranden und die Erstrufe sowie einem Skalenwert für Promo-tionsprogramme wird ein Mittelwert errechnet. Außerdem wird die Note im Forschungsrating genannt. Für die Promotionsprogramme werden die folgenden Skalenwerte vergeben:

Promotionsprogramme:

1 DFG + 1 anderes externes Programm	100
1 DFG	95
1 anderes externes	70
1 internes	60
keines	40

Die höchste Bewertung für ein DFG-Graduiertenkolleg ergibt sich aus dem besonders stren-gen Prüfungsverfahren und der inzwischen sehr hohen Hürde. Der Aufwand ist sehr hoch. Der

Reihung der FER in Spalte 1 liegt der Mittelwert aus den fünf genannten Indikatoren zugrunde. Die einzelnen Indikatoren sowie die Note im Forschungsrating können zur Interpretation und Relativierung der in Spalte 1 vorgenommenen Reihung herangezogen werden (Tab. 8).

Tabelle 8: Nachwuchsförderung

FER	Mittlerer Skalenwert	Promotionen		Finanzierte Doktoranden	Erstrufe	Promotionsprogramme	Note FR
		abs.	pro Prof.				
Bielefeld	92	98	60	100	100	100	5
Göttingen	86	84	69	89	89	100	4
HU Berlin	85	88	61	83	92	100	5
Bamberg	81	73	50	88	97	95	4
Bochum	80	93	86	68	82	70	4
FU Berlin	78	76	56	98	89	70	4
U Bremen	77	68	69	95	82	70	5
LMU München	73	79	88	64	92	40	3
Dortmund	71	68	36	86	71	95	4
WZB	71	96	100	18	82	60	3
Aachen	70	51	91	75	95	40	4
Duisburg-Essen	69	82	39	80	51	95	3
Frankfurt / Main	69	91	36	93	19	95	4
MPIfG	69	47	96	46	97	60	5
Halle-Wittenberg	67	76	42	96	51	70	3
Gießen	67	61	80	27	71	95	3
Darmstadt	65	68	69	75	19	95	3
Osnabrück	65	86	95	54	19	70	2
Köln	63	82	56	64	71	40	2
Marburg	62	47	19	91	51	100	3
U Hamburg	60	95	33	80	51	40	3
Freiburg	56	89	89	11	51	40	3
Jena	56	68	69	83	19	40	3
Magdeburg	56	56	83	54	19	70	3
Dresden	55	43	42	41	51	100	3
Mannheim	54	47	28	75	82	40	4
TU Berlin	53	61	80	64	19	40	2
Trier	51	53	30	60	71	40	3
Heidelberg	50	56	77	4	71	40	3
Tübingen	48	35	69	46	51	40	3
Augsburg	47	56	83	36	19	40	2
Erlangen-Nürnberg	47	9	8	54	71	95	3
TU Chemnitz	46	43	42	32	71	40	3
Frankfurt / Oder	44	20	56	32	19	95	3

FER	Mittlerer Skalen- wert	Promotionen		Finan- zierte Dokto- randen	Erstrufe	Promo- tionspro- gramme	Note FR
		abs.	pro Prof.				
Düsseldorf	43	27	32	64	51	40	3
SOEP	42	20	13	54	82	40	3
Eichstätt	41	35	69	11	51	40	2
Kiel	41	35	93	18	19	40	2
Mainz	41	35	69	41	19	40	2
Stuttgart	41	15	46	54	51	40	3
Hannover	40	63	53	27	19	40	2
DSPH Köln	40	15	46	46	51	40	2
Bonn	39	35	69	4	19	70	2
Kassel	39	40	23	18	19	95	2
Konstanz	39	73	50	11	19	40	2
Wuppertal	38	24	11	36	51	70	2
Erfurt	36	15	13	71	19	60	3
Rostock	31	30	18	46	19	40	2
UBw München	25	15	26	27	19	40	2
Siegen	24	5	5	18	51	40	1
FU Hagen	23	27	25	4	19	40	1
Potsdam	22	9	8	4	51	40	1
Bayreuth	21	9	21	18	19	40	2
UBw Hamburg	17	3	3	18	19	40	1

Es soll hier nicht behauptet werden, dass die Spalte zum mittleren Skalenwert die „wahre“ Rangfolge enthält. Es wird erkennbar, dass man in Einzelfällen zu anderen Bewertungen kommt, wenn den einzelnen Indikatoren nicht das gleiche Gewicht gegeben wird. Was dabei ausbleibt, ist die qualitative Beurteilung nach zusätzlichen Kriterien, die in die Notengebung des Forschungsratings eingegangen sind. Das ist im informed peer review möglich. So kann z.B. die stärkere Gewichtung der Promotionen pro Professor, der Erstrufe oder der Publikation der Dissertationen in angesehenen Verlagen die eine oder andere Abweichung erklären.

Die Tabelle zeigt unmissverständlich, dass Nachwuchsförderung im großen Umfang an den größeren FER prämiert wurde, ein ganz überwiegender Größeneffekt wie beim Impact. Das Ergebnis ist nahezu deckungsgleich mit dem Ergebnis der Impactmessung. Lediglich fünf der ersten 20 Standorte in der Nachwuchsförderung sind nicht unter den ersten 20 Standorten der Impact-Tabelle zu finden. Die Schattenseite dieser Bewertungsstrategie ist die Förderung von „Promotionsfabriken“. Die Wissenschaft hat allerdings keinen Nutzen von dieser Konzentration der Promotionen auf Großstandorte, wenn dadurch die Vielfalt der Forschungsansätze und damit die Offenheit der Wissensrevolution unter das kritische Niveau gesenkt wird.

6 Kriterium V: Transfer

Transferleistungen konnten nicht angemessen berücksichtigt werden, schon wegen der unzureichenden Daten und ihrer schwierigen Quantifizierung in der Praxis. Die wegen der Schwierigkeiten vollzogene Reduktion auf eine Dreierskala bedeutet zwangsläufig eine Abwertung dieses Kriteriums und widerspricht der Intention des mehrdimensionalen Forschungsratings

des Wissenschaftsrates. Dieser Abwertung entsprach auch ein geringes Interesse der öffentlichen Kommunikation an dem Kriterium. Sie bestraft diejenigen FER, die in diese Funktion investieren, und bevorteilt jene, die das unterlassen. Wissenstransfer ist aber genau das, was die Soziologie praktisch relevant macht und was ihren Einfluss auf die Praxis sichert und in den verschiedenen Funktionsbereichen der Gesellschaft (Politik, Verwaltung, Recht, Dritter Sektor, Wirtschaft) Anerkennung bringt. Wenn man mit solchen Leistungen aber keine Prämien gewinnen kann, dann heißt das zwangsläufig, nur noch in den sehr engen Bereich des Social Science Citation Index der Professionellen Soziologie zu investieren. Die Konsequenz wäre die Verarmung der Soziologie zu einem *l'art pour l'art* des methodisch ausgefeiltesten Aufsatzes, bei dem es in aller Regel um nichts unmittelbar praktisch Verwertbares geht.

Die Bewertungsgruppe hat sich bei den anderen Kriterien leichter getan, Abstufungen in einer Fünferskala vorzunehmen, weil die Perzentilwerte eine Handhabe gegeben haben, deren Problematik aber bei genauerem Hinsehen offenbar wird. Dagegen lag beim Transfer nur ein einziger quantitativer Indikator vor: die Drittmittel von Ministerien, Unternehmen und Verbänden. Darüber hinaus existierte nur die Aufzählung entsprechender Aktivitäten. Es wäre aber im Interesse des Faches, hier zu investieren. Man könnte z.B. die erhobenen Transfer-Drittmittel zu einem harten Kriterium machen, das die Basislinie definiert, von der durch die qualitative Auflistung von Transferleistungen nur noch begrenzte Auf- und Abstufungen von maximal einer Notenstufe möglich sind. Ein weiterer harter Indikator wäre die Existenz von An-Instituten, die man als Einheiten gleich oder differenziert nach Personalstärke als Pluspunkt zu den Transfer-Drittmitteln hinzurechnen könnte (Tab. 9).

Tabelle 9: Auftragsforschung

Uni	Perz.	Note FR	Uni	Perz.	Note FR
Bielefeld	100	4	Siegen	44	2
Halle-Wittenberg	98	4	Stuttgart	46	4
Göttingen	96	4	HU Berlin	42	4
Dortmund	95	4	Düsseldorf	40	3
Jena	93	4	Mainz	39	4
TU Darmstadt	91	3	FU Berlin	37	3
Osnabrück	89	4	Aachen	35	2
Freiburg	88	4	Mannheim	33	2
Bochum	86	4	Rostock	32	3
Trier	84	4	Heidelberg	30	2
Bamberg	82	4	Chemnitz	28	3
Duisburg-Essen	81	4	Wuppertal	26	3
Hamburg	79	3	Hagen	25	2
WZB	77	3	Leipzig	23	n.b.
SOEP	75	4	MPiFG	21	3
Bremen	74	3	Frankfurt / Oder	19	2
DSPH Köln	70	4	Marburg	18	3
TU Dresden	68	3	Gießen	16	3
Potsdam	67	2	JU Bremen	14	3
Köln	65	3	TU Berlin	12	2
Kiel	63	3	Bayreuth	6	2
Erfurt	61	2	Eichstätt	6	2

Uni	Perz.	Note FR	Uni	Perz.	Note FR
Münster	60	n.b.	UBw Hamburg	6	2
Bonn	58	3	Konstanz	6	2
Tübingen	56	3	LMU München	6	2
Erlangen-Nürnberg	54	3	UBw München	6	2
Frankfurt / Main	53	3			
Kassel	51	3			
Hannover	49	3			
Augsburg	47	3			

Weil hier nur ein Indikator zugrunde gelegt werden kann, darf die Tabelle auf keinen Fall als komplettes Rating für Transfer interpretiert werden. Hier war es nicht anders möglich, als durch die qualitative Bewertung von Transferleistungen zu einem endgültigen Urteil zu gelangen, was sich aber als sehr schwierig darstellte.

7 Kriterium VI: Wissensvermittlung und Wissensverbreitung in die Öffentlichkeit

Auch die Wissensvermittlung und die Wissensverbreitung in die Öffentlichkeit konnten nur sehr eingeschränkt beurteilt werden. Infolgedessen wurde nur eine Dreierskala verwendet. Die Herabstufung dieser Dimension auf eine Dreierskala setzt jedoch ein Signal, das zum Abbau der entsprechenden Aktivitäten beiträgt. Dazu gehören insbesondere Zeitdiagnosen in Büchern sowie Kommentare zum Zeitgeschehen in Presse, Rundfunk und Fernsehen. Wenn das nicht gleichwertig zu den anderen Kriterien honoriert wird, dann nimmt die Soziologie weniger Teil an der Beeinflussung der gesellschaftlichen Entwicklung durch gesellschaftliche Selbstbeschreibung. Die Soziologie überlässt diese Aufgabe dann ganz dem Journalismus, der Politik, der Wirtschaft und den Ökonomen. Hier ist jedoch der Ort der soziologischen Aufklärung. Aber wie könnte diese Intention umgesetzt werden? Es liegt für dieses Kriterium überhaupt keine Kennziffer vor. So wie Aufsätze in Zeitschriften oder Sammelbänden, Herausgeber-schaften und Monographien ohne Unterschiede nach Länge in ihrer Zahl erfasst wurden, hätte man auch hier nach Kategorien zählen können:

1. Zeitdiagnostische Monographien (evtl. Resonanz in Presse und Rundfunk)
2. Aufsätze in allgemeinen Publikumszeitschriften wie *Merkur* und *Internationale Politik*
3. Beratende Veröffentlichungen
4. Interviews
5. Radiobeiträge
6. Pressebeiträge überregional
7. Pressebeiträge regional

Man könnte wie bei den Publikationen Perzentilwerte ausrechnen, zumal sich die FER die Mühe gemacht haben, die Daten selbst zu liefern. Allerdings haben einige FER diese Daten beim Kriterium V statt beim Kriterium VI eingegeben. Sie müssten umsortiert werden.

Wir befinden uns hier im Feld der Öffentlichen Soziologie (Burawoy 2005; Bude 2005). Dass deren Leistungen nicht gleichwertig zu den Leistungen der Professionellen Soziologie dargestellt werden konnten, ist ein Manko des Forschungsratings. Die Folge wird eine Vermehrung von PR-Artikel-Produzenten sein, die Soziologie ohne öffentliche Resonanz betreiben, während die Spezies von Soziologen, die sich an die Öffentlichkeit wenden, aussterben wird. Das ist die Schattenseite der unzureichenden Würdigung der Verbreitung soziologischen Wissens in der Öffentlichkeit.

8 Forschungsrating: ein vermintes Feld

Wie leicht zu erkennen ist, begibt sich jedes Forschungsrating in ein hochgradig vermintes Feld (vgl. Hirschauer 2004). Überall lauert die Gefahr einer Realitätskonstruktion mit kontraproduktiven Konsequenzen für den Erkenntnisfortschritt (Frey 2006; Münch 2007). Es muss gesehen werden, dass die hier durchgeführten Analysen immer noch das Spiel des Ratings mitmachen, das man als solches mit guten Gründen in Frage stellen kann. Die Bewertung der FE und FER erfolgt nach einem Kriterienkatalog, der keine selbstverständliche Geltung beanspruchen kann. Seine Anwendung ist blind für eine Vielzahl relevanter Forschungsleistungen, die durch das Rating unsichtbar gemacht werden. Diesem Dilemma kann sich auch der hier vorgelegte Beitrag nicht entziehen. Ebenso wenig entgeht er dem Dilemma, die Grenzen des Vermessens der Wissenschaft durch eigenes Vermessen zu bestimmen, die Konstruktion der soziologischen Realität durch das Forschungsrating durch eine weitere Konstruktion zu verdoppeln. Allerdings kann nur auf diese Weise ein über das bloße Klagen hinausgehender Beitrag zur Selbstaufklärung der Soziologie über den sozial konstruierten Charakter ihrer Forschungsrealität geleistet werden. Es wird das Bewusstsein dafür geschärft, dass Gewinner und Verlierer dieses Spiels nach Regeln erzeugt werden, die auch anders aussehen könnten. Dazu gehört auch die Schärfung des Bewusstseins dafür, dass die Soziologie der große Verlierer wäre, wenn alle den Gewinnern des Forschungsratings nacheifern würden. Sie würde auf das im Rating prämierte Wissen verarmen.

Ein Forschungsrating – wie multidimensional es auch ausgerichtet sein mag – ist ein bedeutsamer Schritt von der dezentralen marktförmigen Zuweisung von Reputation zu einer zentralen und hierarchischen. Damit verliert die Zuweisung von Reputation an Offenheit für Wandel und Innovation. Es entstehen verfestigte Statushierarchien von Forschungseinrichtungen (Münch 2008). Jedes Forschungsrating ist hochgradig selektiv und hat in dieser Selektivität weit reichende Konsequenzen. Nur leichte Veränderungen in der Art und Zahl der Kriterien und in ihrer Operationalisierung durch Indikatoren führen zu erheblichen Verschiebungen in der Bewertung und Rangordnung von Forschungseinheiten bzw. Forschungseinrichtungen. Die selektierten Indikatoren spiegeln nicht direkt die alltägliche Praxis der Anerkennung von Forschungsleistungen wider. Es wird in Ränge eingestuft, was vorher in eigener Würde nebeneinander existierte und Anerkennung gefunden hat. Das Spektrum des Anerkannten wird auf das immer auch anders mögliche Maß der angewandten Indikatoren reduziert. All das hat eine schließende Wirkung auf die Wissensrevolution. Wir bewegen uns ein erhebliches Stück weg von der Feyerabend-Welt des *anything goes* (Feyerabend 1976) und hin zur Thomas von Aquin-Welt der Scholastik. Umso mehr bedarf es einer frühzeitigen und regelmäßigen Evaluation der Evaluation, um die de facto zur Geltung gekommenen Regeln zu erkennen und über die Folgen ihrer Anwendung reflektieren und debattieren zu können (vgl. Weingart 2001). Dazu sollte die hier vorgelegte Sekundäranalyse einen Beitrag leisten.

Es bleibt allerdings das Problem, dass die Veröffentlichung der Sekundäranalyse selbst zur Stratifizierung des soziologischen Feldes und damit zur Beseitigung der idealen Sprechsituation beiträgt. Da aber die Ergebnisse des Forschungsratings nun einmal in der Welt sind, ist es besser, über ihr Zustandekommen zu sprechen, als darüber zu schweigen, weil sich die betroffenen FER nur dann überhaupt darauf einstellen und auch darüber debattieren können. Nur auf diesem Wege kann Evaluation als Lernprozess institutionalisiert werden.

Bei allen Gefahren, die hinter dem Forschungsrating lauern, muss anerkannt werden, dass es für die Forschungspolitik wichtige Ergebnisse hervorgebracht hat. Dazu sind insbesondere zu zählen: (1) die breite Streuung sehr guter bis exzellenter FE über die Standorte, (2) die Kluft zwischen der hohen Bewertung von PR-Artikeln und englischsprachigen Publikationen und der noch wenig darauf eingestellten Publikationspraxis der deutschen Soziologie, (3) die Schwierigkeiten, die Vielfalt von Forschungsleistungen und die Kreativität von Forschung in

einem Ratingverfahren angemessen zu würdigen, (4) die kontraproduktiven Effekte einer auf Drittmiteleinwerbung fixierten Forschungspolitik, (5) die unzureichende Förderung der selbstständigen Publikationstätigkeit von Nachwuchswissenschaftlern im Rahmen der herrschenden oligarchischen Strukturen, die sich in der geringen Präsenz der Nachwuchswissenschaftler in den Publikationslisten geäußert hat. Es gibt demnach Reflexionsbedarf auf beiden Seiten: auf der Seite des Forschungsratings selbst wie auch auf der Seite der „gerateten“ Realität soziologischer Praxis.

Literaturverzeichnis

- Bude, H. (2005): Auf der Suche nach einer öffentlichen Soziologie. Ein Kommentar zu Michael Burawoy, in: *Soziale Welt* 56, S. 375-380.
- Burawoy, M. (2005): For Public Sociology, in: *Soziale Welt* 56, S. 347-374.
- Feyerabend, P. K. (1976): Wider den Methodenzwang. Skizze einer anarchistischen Erkenntnistheorie, Frankfurt / Main.
- Frey, B.S. (2006): Evaluitis – Eine neue Krankheit. Working Paper No. 293. Institute for Empirical Research in Economics, University of Zurich.
- Gerhards, J. / G. G. Wagner (2008): Es geht um die Qualität. Was ist das Besondere an dem Forschungs-rating des Wissenschaftsrats und warum sollte es Schule machen?, in: *Forschung & Lehre* 15, S. 532-533.
- Hirschauer, S. (2004): Peer Review auf dem Prüfstand. Zum Soziologiedefizit der Wissenschaftsevaluation, in: *Zeitschrift für Soziologie* 33 (1), S. 62-83.
- Münch, R. (2007): Die akademische Elite, Frankfurt / Main.
- Münch, R. (2008): Stratifikation durch Evaluation. Mechanismen der Konstruktion von Statushierarchien in der Forschung, in: *Zeitschrift für Soziologie* 37, S. 60-80.
- Neidhardt, F. (2008): Das Forschungsrating des Wissenschaftsrats. Einige Erfahrungen und Befunde, in: *Soziologie* 37 (4), S. 421-432.
- Power, M. (1997): *The Audit Society: Rituals of Verification*, Oxford.
- Weingart, P. (2001): Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft, Weilerswist.
- Wissenschaftsrat (2007): Forschungsleistungen deutscher Universitäten und außeruniversitärer Einrichtungen in der Chemie, Köln.
- Wissenschaftsrat (2008 a): Forschungsleistungen deutscher Universitäten und außeruniversitärer Einrichtungen in der Soziologie, Köln.
- Wissenschaftsrat (2008 b): Ergebnisse der Pilotstudie Forschungsrating im Fach Soziologie: tabellarische Übersichten vom 6.8.2008, abgerufen am 16.12.2008 unter http://www.wissenschaftsrat.de/texte/pilot_uebersicht_sozio.pdf.

Prof. Dr. Richard Münch
Otto-Friedrich-Universität Bamberg
Lehrstuhl für Soziologie II
Lichtenhaidestr. 11
96045 Bamberg
richard.muench@uni-bamberg.de