

**Qualitätsmessung: L. Raphael: Probleme und Chancen der
Forschungsbewertung im Fach Geschichte**
von Lutz Raphael

Indikatoren gestützte Evaluationen von Forschungsleistung spielen auch in den Geisteswissenschaften trotz aller berechtigten Skepsis und verständlicher Widerstände bereits zum jetzigen Zeitpunkt eine große Rolle. Die folgenden Überlegungen gehen von der pragmatischen Einschätzung aus, dass der in Politik wie Gesellschaft insgesamt zu beobachtende Trend, sich auf Ratings/Rankings und andere indikatoren gestützte Verfahren vergleichender Leistungsmessung zu verlassen bzw. sie zur Basis weiterer Aushandlungsprozesse und Entscheidungen zu machen, weder aufzuhalten noch per se negativ ist. Solche Verfahren knüpfen an die Alltagspraxis wissenschaftsinterner Bewertungen - von der Benotung studentischer Leistungen, von Dissertationen bis hin zu Gutachten in Berufungsverfahren und bei Projektanträgen - an. Die möglichst fachgerechte, wissenschaftsadäquate Gestaltung dieser Prozesse sollte dementsprechend im Mittelpunkt der Kontroversen stehen.

Aus der Perspektive der betroffenen Wissenschaftler bzw. ganzer Fächer bzw. Fächergruppen geht es vor allem darum, konkret die Spielräume für eine autonome Gestaltung solcher Evaluationsprozesse zu erweitern und die zur Zeit erkennbaren Fehler bei der Übertragung sachfremder Kriterien und Bewertungsperspektiven auf die Geisteswissenschaften zu korrigieren.

Welche spezifischen Probleme muss nun ein Evaluationsverfahren im Fach Geschichte berücksichtigen? Mir scheinen mindestens sieben Punkte von besonderer Bedeutung:

1. Das Spannungsverhältnis zwischen Einzelforschung und Verbundforschung

Ganz ähnlich wie im Fach Soziologie, das Gegenstand einer Pilotstudie zum Forschungsrating seitens des Wissenschaftsrats war¹, ist die Wahl geeigneter Forschungseinheiten ausgesprochen problematisch. Im Fach Geschichte existiert nicht erst seit der finanziellen und administrativen Reduktion von Forschungschancen und -ressourcen aus der Grundausstattung von Hochschulinstituten bzw. -lehrstühlen eine Tradition koordinierter Forschung. Man denke nur an die außeruniversitären Forschungsinstitute wie das frühere Max-Planck-Institut für Geschichte in Göttingen, das Institut für Zeitgeschichte in München, aber auch viel ältere Organisationen historischer Forschung wie die Monumenta Germaniae Historica oder die Historische Kommission bei der Bayerischen Akademie der Wissenschaften. Auch die universitären Einrichtungen des Faches kennen vielfältige Formen koordinierter Forschung. Bislang bestehen in der Fachcommunity konträre Bewertungen über Sinn und Nutzen solcher Verbundforschung mehr oder weniger unverbunden nebeneinander. Die kreative Einzelleistung wird fachintern häufig in Übereinstimmung mit tiefsitzenden und gut tradierten Vorurteilen allein dem Pol individueller Forschung fernab vom Getriebe der drittmittelgetriebenen Projektwelt zugeordnet. Insbesondere DFG-Sonderforschungsbereiche oder ähnliche großformatige Forschungsverbünde müssen mit einer besonders kritischen Prüfung ihres Innovationspotentials rechnen. Demgegenüber ist daran zu erinnern, dass gerade auch projektförmige Forschung wichtige Impulse in der Geschichtswissenschaft gegeben hat. Das Bild einer durch die Abhängigkeit von Drittmitteln intellektuell verarmenden mainstream-Historiographie ist ebenso einseitig wie empirisch angefechtbar. Eine nüchterne Bilanz der größeren Verbundforschung für

¹Eine Übersicht zum Forschungsrating Soziologie findet sich auf folgender Seite des Wissenschaftsrates: <<http://www.wissenschaftsrat.de/pilot-sozio.html>> (12.05.2009). Die „Pilotstudie Forschungsrating Soziologie. Abschlussbericht der Bewertungsgruppe“ ist ebenfalls im Internet einsehbar: <<http://www.wissenschaftsrat.de/texte/8422-08.pdf>> (12.05.2009).

die Geschichtswissenschaft steht einfach noch aus. Eigene Forschungen zur Geschichte der Annales-Historiographie haben mir gezeigt, dass neben der immer wieder gern angeführten Figur des singulären Gelehrten gerade der in organisierten Arbeitszusammenhängen forschende Historiker die Entwicklung des Faches intellektuell markiert hat. Gerade die hochschulpolitische Hebelwirkung eingeworbener Drittmittel blockiert aber eine sachgerechte Diskussion dieser unterschiedlichen Forschungsformate.

Angesichts dieser Lage wäre es von großem Nutzen, wenn Evaluationen im Fach Geschichte zum einen das Ausmaß von Forschung in den Forschungseinheiten anhand der Zahl geförderter Projekte bzw. dort abgeschlossener Dissertationen/Habilitationen ermitteln, aber andererseits auch eine präzisere Bewertung bzw. Bilanzierung größerer Projektzusammenhänge ermöglichen.

2. Die fachspezifischen Publikationsformen

Die Geschichtswissenschaft kennt nach wie vor eine Vielfalt von Publikationsformen, die per se keinerlei Qualitätsgefälle erkennen lassen. Dementsprechend eignen sie sich nicht zur Berechnung quantitativer Indikatoren. Auch im Fach Geschichte kommt keine Evaluation an der Bewertung von Texten vorbei: Fachgutachter müssen lesen, die Auswahl der entsprechenden Texte wird also entscheidend sein. Hier wäre meines Erachtens eine Kategorisierung der vielen unterschiedlichen Textformate hilfreich. Die Pluralität der Publikationsformen hat sogar mit der Etablierung fachspezifischer elektronischer Publikationen noch zugenommen, so dass für eine fach- und sachgerechte Bewertung aktueller Forschungsleistungen weder eine Beschränkung auf Aufsätze in wenigen, als „erstklassig“ klassifizierten Fachzeitschriften (nach dem Muster von A-, B-Journals etc.), noch auf Printmedien eine sichere Grundlage darstellen würde. Offensichtlich ist es aber angesichts dieser Vielfalt unabdingbar, wenigstens Mindeststandards zu definieren, was als Fachpublikation Berücksichtigung finden kann. Hier wird wohl nur das Kriterium des peer review Verfahrens für alle Periodika (print wie elektronisch) eine sichere Startgrundlage liefern

können.

Für eine differenzierte vergleichende Bewertung der Texte und auch die Darstellung der Bewertungsergebnisse wiederum kann es sehr sinnvoll sein, nach Typen von Forschungsbeiträgen zu differenzieren. Ausgangspunkt können etablierte, ältere Klassifizierungsmodelle sein. So existierte lange Zeit das auch international kanonisierte Schema: Aufsätze in Fachzeitschriften oder Monographien im Buchformat dienen der Darstellung von Forschungsergebnissen. Klassiker in der Kategorie des Forschungsbuchs sind dabei die überarbeiteten, in anerkannten Reihen publizierten Dissertationen und Habilitationen geworden. Rein quantitativ haben sie erheblich zugenommen, ihre Präsenz und Relevanz in der Fachkommunikation ist aber sehr heterogen, in erheblichem Maß wird die Aufmerksamkeit der Fachkollegen durch das Renommee der Reihen und Verlage gesteuert. Etabliert ist schließlich auch das Handbuch bzw. der Handbuchartikel als Form der Synthese und der Darstellung des Forschungsstandes. Eher randständig und vielfach nicht als Forschungsleistung anerkannt waren in diesem älteren Modell programmatische Forschungsaufsätze bzw. Beiträge zu Fragen von Theorie und Methode. Die Überzeugungskraft dieses Modells beruhte und beruht in erheblichem Maße auf einem empiristischen bzw. positivistischen Vorverständnis, dessen theoretische Grundlagen seit längerem im Fach umstritten sind. Mit der wachsenden Pluralität von Forschungsansätzen hat sich auch in dieser Hinsicht das Feld der Publikationen weiter ausdifferenziert. So ist inzwischen auch in der deutschsprachigen Geschichtswissenschaft die im Umfang und Anmerkungsapparat knappere Synthese meist in Taschenbuch-/Paperbackformat getreten, die sowohl für den Lehrbetrieb wie für die Fachkommunikation über den kleinen Kreis der Spezialisten hinaus gedacht ist. Gerade auf diesem Feld sind in den letzten zehn bis 15 Jahren die überraschendsten Neuentwicklungen zu beobachten gewesen. Dieses Syntheseformat erweist sich auch viel stärker als das klassische Handbuch als Ort der Innovation, weil dort neue Perspektiven erprobt, Geschichte(n) zuweilen neu- und umgeschrieben

werden. Deutlich an Gewicht zugenommen hat schließlich auch der Forschungsessay – sowohl als Beitrag in Fachjournalen wie auch als Buch. Gerade die wachsende Pluralisierung geschichtswissenschaftlicher Forschungsansätze in den letzten drei Jahrzehnten hat hier ihren Niederschlag gefunden. Für eine Evaluierung von Forschungsleistungen wäre eine Differenzierung in entsprechende Dimensionen (Spezialforschung/ Synthesen/Konzepte) sehr hilfreich, da dies ermöglichen würde, die ganz unterschiedlichen Profile und Stärken/Schwächen von Forschungseinheiten sichtbar zu machen. Auch die Auswahl von Publikationen könnte damit gesteuert werden. Eine schwierige Frage ist es, ob Publikationen bzw. Beiträge für Reihen, die explizit für die universitäre Lehre geschrieben werden und in ihrem Format diesem Ziel entsprechend der Reproduktion bekannten Wissens erste Priorität einräumen, als eigenständige Form der Forschungskommunikation behandelt werden sollten. Angesichts des wissenschaftspolitischen Ziels, die Qualität der Lehre zu verbessern, wäre eine solche explizit eigenständige Kategorisierung sicherlich von strategischer Bedeutung.

3. Typen der Forschungsleistung

Schließlich muss die Evaluation von Forschungsleistungen in der Geschichtswissenschaft auch der Tatsache gerecht werden, dass – relativ unabhängig von der Wahl der Publikationsformate – die einzelnen Forscher und Forschungseinrichtungen arbeitsteilig ganz unterschiedliche Aspekte dieser fachspezifischen Gesamtleistung bedienen. In der Geschichte der Geschichtswissenschaft bewährt hat sich zum Beispiel die Unterscheidung zwischen der „Erfindung“ oder Exploration eines neuen Forschungsfeldes einerseits, der Untersuchung eines bereits etablierten, aber materialiter noch „unbekannten“ Untersuchungsfeldes andererseits (letztere ähnelt dem, was Kuhn „Normalwissenschaft“ genannt hat). Bei beiden Grundtypen gibt es jedoch offensichtlich unterschiedliche Qualitätskriterien: im Fall der Normalwissenschaft gibt es offenkundig den Grenzfall der „Nach„forschung, also der bloßen Wiederholung etablierter Forschungsansätze an immer kleinteiligeren und zudem weitgehend bereits erforschten, das heißt bekannten Gegen-

ständen. Der Grenzfall ist bekanntlich dann erreicht, wenn es bereits internationale Forschungsergebnisse zu einem Problem/Sachverhalt gibt, aber nur noch nicht in der deutschen Forschung! Andererseits gehört zur zukunftssträchtigen und damit hoch zu bewertenden Erschließung neuer Forschungsgebiete neben der theoretischen Konstruktionsleistung einer neuen Fragestellung (den inzwischen zur Mode gewordenen turns) auch die angemessene Behandlung methodischer Anschlussfragen (ein nahe liegendes Exempel liefert zum Beispiel die Diskursanalyse) und die materialgesättigte Erprobung. Mit Blick auf die breit diskutierten Folgewirkungen von Evaluationen wären gerade zwei typische Gefährdungen guter historischer Forschung durch diese beiden entgegengesetzten Trends, einerseits die modische Originalitätssuche auf Kosten der aufwändigeren und nur mittelfristig erreichbaren Etablierung neuer Forschungsfelder und andererseits die innovationsfeindliche Weiterführung längst hinreichend erforschter Themen und Zusammenhänge beobachtbar. Davon wäre langfristig eine Korrektur naturwüchsiger Fehlentwicklungen zu erhoffen.

4. Internationalität der Forschung

Dieses Kriterium ist ebenso unbestritten wie ungeklärt. Faktisch wird auch auf internationaler Ebene die Geschichtswissenschaft nach wie vor vom Primat nationalgeschichtlicher Perspektiven dominiert. Dies hat bereits weitreichende Folgen für die nationalspezifische Besetzung von Forschungsfeldern und die Herausbildung von besonderen Forschungsprofilen in den einzelnen nationalen Historikerfeldern. Internationale und nationale Wahrnehmung von Forschung stehen dabei zuweilen im umgekehrten Verhältnis zueinander – ohne dass dies Rückschlüsse über die Qualität der Forschung zuließe. Die kulturelle und politische Bedeutung des Faches auf nationalstaatlicher Ebene schlägt sich nicht zuletzt darin nieder, dass sich das Fach nach wie vor der Landessprache bei der Präsentation ihrer Ergebnisse bedient. Jede Bevorzugung fremdsprachiger Publikationen, zumal in englischer Sprache würde vor allem dazu dienen, die deutschsprachige Geschichtswissenschaft als Teil eines anderen, praktisch eines angel-

sächsischen Historikerfeldes (dieses ist aber nicht identisch mit der internationalen Geschichtswissenschaft) neu zu definieren. Für einzelne Forschungsgebiete mag dies angemessen sein, eine generelle Übertragung würde aber ein aktuell weitgehend forschungsfernes Kriterium einführen und damit die Ergebnisse verzerren. Dementsprechend prekär sind nach wie vor international vergleichende Beurteilungen nach quantitativen Kriterien.

Angesichts der großen Unterschiede innerhalb des Faches wäre es aus meiner Sicht am besten, den Aspekt der Internationalität nur dort systematisch zur Geltung zu bringen, wo er etwas über die fachliche Bewertung der jeweiligen Forschung aussagt – bei der Bewertung der Forschungstexte: sie sollte sich an Qualitätsstandards orientieren, die allgemeine, das heißt internationale Anerkennung in den Zentren der Geschichtsforschung finden. Internationalität heißt in diesem Fall nicht schlicht Vernetzung mit ausländischen Fachkollegen, sondern Qualitätsmessung an internationalen Standards. Die sachgerechte Anwendung dieses Kriteriums ist also konkret im Wesentlichen an die Kompetenz und Urteilsfähigkeit der Gutachter gebunden.

5. Zeiträume der Evaluation

Nach wie vor herrscht große Unklarheit darüber, in welchen Zeiträumen Evaluationen in den Geisteswissenschaften sinnvoll sind. Offensichtlich besteht die Gefahr der Evaluationitis, also des sachfremden Rankings in viel zu kurzen Abständen. Rhythmus und Tempo von Forschung im Fach Geschichte lässt sich recht gut an der Dauer ablesen, die für sehr gut bewertete Dissertationen wie für Habilitationen üblich sind. Hier liegt der Durchschnitt deutlich über vier Jahren, rechnet man die Zeit bis zur Publikation hinzu, dauert es also in der Regel sechs bis sieben Jahre, bevor innovative Leistungen kommuniziert werden und ihre Wirkung im Fach entfalten. Erst dann sind sie verlässlich bewertbar. Trotz aller Moden, die im Fach kurzfristige Spuren hinterlassen, spricht (noch) nichts dafür, dass sich die Zeiten verkürzt hätten, die individuelle Vorhaben wie drittmittelgeförderte Projekte benötigen, um neue Erkenntnisse hervorzubringen. Technologische

Innovationen haben hier keine Beschleunigung gebracht. Deshalb sind seriöse, das heißt aber nach dem bisher Gesagten arbeitsaufwändige Evaluationen nur in längeren Zeitabständen (sieben bis zehn Jahren) sinnvoll. Sie können so etwas wie nationale Leistungsbilanzen darstellen, welche es erlauben, die laufende Praxis interner wie externer Bewertungen spezifischer Forschungsleistungen zu „eichen“ und Instituten wie einzelnen Forschern dabei helfen, eine nüchterne Analyse von Stärken und Schwächen vorzunehmen.

6. Auswahl der Gutachter

Die Qualität einer fachinternen Evaluation, die eben mehr ist als die Zusammenstellung unsicherer quantitativer Indikatoren, hängt in entscheidendem Maße von der Qualität des Fachgutachtergremiums ab. Die angemessene fachliche Vertretung der einzelnen Teildisziplinen ist eine erste Voraussetzung - hier bieten aus meiner Sicht jedoch die Tendenzen zur Aufrechterhaltung der Einheit des Faches in der Lehre sowie zur Kooperation der Teildisziplinen in Forschungsprojekten eine gute Grundlage. Kompliziert stellt sich sicherlich die Frage der Abgrenzung im Bereich der Wissenschaftsgeschichte dar, während die Medizin- und Rechtsgeschichte als - vielfach randständige - Teildisziplinen von Rechtswissenschaften und Medizin zu gelten haben und nur sehr schwer nach denselben Regeln fachlich begutachtet werden könnten wie die geschichtswissenschaftlichen Teilfächer. Ein solches Gremium wird deshalb anknüpfen können an die etablierten Repräsentationspraktiken im Historikerverband und in der DFG. Ebenso stark beachtet werden müsste jedoch auch die angemessene Vertretung der unterschiedlichen Forschungsansätze im Fach. Gerade auch in aktuellen Forschungstendenzen sich artikulierende neuere Theorien und Konzepte dürfen in einem insgesamt die Pluralität der Fachströmungen repräsentierenden Fachgremium nicht fehlen. Für diesen Aspekt wie auch andere Gesichtspunkte einer fairen Evaluation wird auch eine angemessene Berücksichtigung der unterschiedlichen Altersgruppen bzw. „Generationen“ innerhalb des Faches von Gewicht sein.

7. Transfer

Die Geschichtswissenschaft ist in vielfältiger Weise in die Geschichtskultur der Bundesrepublik eingebunden. Einzelne Historiker, aber vielfach ganze Forschungseinrichtungen übernehmen auch Aufgaben bei der Begutachtung und Gestaltung von Ausstellungen, Denkmälern, Museen, Publikationen, audiovisuellen Produktionen usw. Dieses weite Feld der Transfers von Forschungsergebnissen in Kultur und Politik ist sicherlich nicht selbst unmittelbar als „Forschung“ zu bezeichnen, steht aber in vielfach engem Zusammenhang zu ihr und gehört zweifellos auch zu den genuinen Aufgaben des Faches und sollte bei einer Evaluation der Leistungen beobachtet und gewichtet werden. Eine Vermischung mit genuinen Forschungsleistungen wäre aber aus meiner Sicht sachfremd, eine schlichte Ausklammerung dieser Aufgaben der Geschichtswissenschaft andererseits nicht angemessen. Hier wäre an das Pilotprojekt im Bereich der Soziologie anzuknüpfen, um diese „dritte“ Dimension der Wissensvermittlung und des Transfers von Forschungsergebnissen sachgerecht einzubeziehen.

~~~~

*Lutz Raphael ist seit 1996 Professor für Neuere und Neueste Geschichte an der Universität Trier. Seine Forschungsschwerpunkte sind die Geschichte der Sozialpolitik im ländlichen Raum, die Geschichte der Geschichtswissenschaften im 19. und 20. Jahrhundert und der Intellektuellen in Europa im 20. Jahrhundert. Er ist seit 2007 Mitglied im Wissenschaftsrat.*