

Prof. Dr. Christian Simon

Qualitätssicherung, Leistungsmessung und Evaluation

Die im Titel genannten Themen sollen im Sinn eines Inputs in dieser Reihenfolge kurz angesprochen werden.

Dabei gehen wir nicht von der Suche nach spezifischen Aspekten für die Geschichtswissenschaft aus, sondern von Ergebnissen von Debatten, die meist von Naturwissenschaftlern geführt worden sind. Man wird sehen, dass in diesem Fall keine Vergewaltigung einer „weichen“ Wissenschaft durch „harte“ Wissenschaften erfolgt, sondern dass ein Lernprozess durchaus von letzteren angestoßen worden ist, den auch erstere rezipieren können.

Dies impliziert einen Bruch mit der geisteswissenschaftlichen Gewohnheit, die ich die nachholende Modernisierung nennen möchte: weg vom schlechten Image der individuellen Forschung ohne „Projekte“ und „Ziele“, hin zu Programm und Projekt und Vernetzung; weg vom Buch hin zum Journal Article, weg von der Verankerung in einer bestimmten Sprache und deren Ausdrucksmöglichkeiten hin zum Englischen, weg von der eigenen Suche nach Sinn und Inhalt hin zum Beitrag zu einer universellen Theorie und Thematik, sich einfügen in einem Stream, mag er auch ein Mainstream sein, mit gültigen Paradigmen und Begriffen, die bloss noch anzuwenden sind.

Das Anliegen, das im Wissenschaftsrat, dem ich nur noch wenige Wochen dienen werde, an den Naturwissenschaften immer deutlicher aufgeschienen ist, ist die Sicherung der Möglichkeit und der Freiheit, originär Neues zu suchen und zu finden. Ohne dauernd dem Zwang zu unterliegen, im Projektrhythmus vorgeformte Eier mit vorgegebenen Inhalten legen zu müssen, um im System zu bleiben oder in es hineinzugelangen.

Der Wissenschaftsrat steht mit seinem Warnruf, dass Substanz verlorengelange durch falsche Anreize, die sich aus falschen Übertragungen betriebswirtschaftlicher Konzepte ergeben, nicht allein. Peter Weingart, in den 1980er Jahren selbst ein Promotor der angewandten Bibliometrie, hat in den letzten 10 Jahren wiederholt darauf hingewiesen, was geschieht, wenn Forschende ihr Verhalten so ausrichten, dass sie bei der gegenwärtigen Art der Leistungsmessung und –bewertung optimal abschneiden. Viel radikaler argumentiert das Academic Manifesto aus Amsterdam, das in der von Weingart geleiteten, hochrangigen allgemeinen Wissenschaftszeitschrift *Minerva* im Frühjahr 2015 publiziert wurde. Alarmiert sind auch Investoren, die in Unternehmen Geld stecken, die auf wissenschaftlichen Ergebnissen aufbauen. Der *Economist* hat vor einiger Zeit eine harte Kritik an der schlechten Qualität wissenschaftlicher Publikationen in Molekularbiologie veröffentlicht. Die Industrie klagt über mangelnde Reproduzierbarkeit von Forschungsergebnissen, und diese Probleme werden dadurch verursacht, dass die Forschenden für die Zahl und Frequenz der Veröffentlichungen belohnt werden, nicht für die Solidität ihrer Erkenntnisse. Und wer die Resultate anderer kritisch nachprüft, kann sich die damit verbrachten Stunden und die entsprechenden Aufsätze nicht als originäre Forschung gut schreiben lassen, erleidet somit Nachteile im Wettbewerb und wird vom System bestraft.

So sieht es in den Naturwissenschaften aus, und als Ursachen werden die Praktiken der

Leistungsmessung, Evaluation und Qualitätssicherung angesehen. Sollen wir also als Historiker dazu übergehen, uns an diesen Praktiken zu orientieren?

Zur Qualitätssicherung

Für die Forschung sollte dies eigentlich gar kein Thema sein. Wissenschaft wäre nicht Wissenschaft, wenn deren Resultate nicht dauern der manchmal unfairen Kritik der lieben Kolleginnen und Kollegen ausgesetzt wäre. Ich spreche nicht vom „peer reviewed journal“, einem Ideal der nachholenden Modernisierung in den Geisteswissenschaften. Und zwar deshalb nicht, weil es ein Mythos ist oder geworden ist in den Naturwissenschaften. Das Peer Review System ist dort am Zusammenbrechen oder doch unter starkem Beschuss, weil es gekoppelt ist mit der Zeitschriftenkrise, d.h. einer Hyperkommerzialisierung der wichtigeren Journals, des Zugangs zu diesen für die Autorinnen und Autoren und der Distribution, mit der den öffentlichen Bibliotheken fast das Licht ausgeblasen worden ist. Wer publiziert, bezahlt, und wer mehr bezahlt, profitiert von einem „fast track“ der Peer Review, die einen rascheren Entscheid bewirkt.

Die wichtigste Massnahme der Qualitätssicherung der Forschung ist in fast allen Fächern: genügend Zeit geben zum Lesen, zum Nachdenken, zum Recherchieren und zum Schreiben, zum Diskutieren und redigieren.

In der Lehre steckt wohl das eigentliche Problem der Qualitätssicherung. So lange allen klar war, was Wissenschaft sein soll, und so lange die Verhältnisse überschaubar und elitär waren, war Qualitätslehre „Meisterlehre“. Das ist nicht Nostalgie: Sie ist es heute noch in vielen Fächern. Ich meine damit: Wer von einem Nobelpreisträger oder ähnlichem Eliteforscherpersönlichkeit in den inneren Kreis der Schülerinnen und Schüler zugelassen worden ist, hat qualitätsvolle, individualisierte, forschungsbezogene Lehre – ein optimaler Start für eine Karriere in der Wissenschaft.

Nur eben: wenige Studierende studieren, um der Wissenschaft näher zu kommen. Man müsste also wissen, was das Ziel universitären Unterrichts in einem Fach ist. Für Geschichte, die kein „Beruf“ ist, ist diese Frage besonders schwer zu beantworten. So lange wir nicht einfach und klar (!) kommunizieren können, was das Ziel des Geschichtsunterrichts an der Hochschule ist, können wir auch keine Qualität bestimmen und keine solche sichern. Hochschuldidaktik und Kurrikulumsentwicklung bringen sicher Fortschritte gegenüber den früher (heute noch?) üblichen langweiligen und irgendwie zweckfreien Vorlesungen und Seminarritualen. Kompetenzkataloge in Ehren, aber was konnte ich machen mit Studierenden (bis Herbst 2015), die weder gezielt ihre „soft skills“ im Argumentieren, Strukturieren und Kommunizieren erweitern wollten (dafür sind wir nicht an der Uni!) noch eine Ausbildung zu Forschung haben wollten (kann mir nicht vorstellen, monatelang zu forschen!).

Zur Leistungsmessung

Leistung aber ist ein noch viel tückischerer Begriff. Wenn ich nochmals auf die Überlegungen des Wissenschaftsrats in den letzten Jahren zurücklenken darf, so ist die Frage nach der Leistung irgendwie geprägt von einer Aussensicht (gespiegelt in unserer Innensicht), dass die da in den Universitäten doch auch endlich sagen sollten, was sie eigentlich leisten. Oder kurz: Tun die den überhaupt etwas? Es ist das Misstrauen derer, die ihr Leben fristen mit einer

undankbaren Bütz und denken, diese Gelehrten sollten eigentlich auch etwas leisten für das viele Geld. Ehrliche Bütz misst sich an den Produkten, die hergestellt werden, deren Zahl und Nützlichkeit, Verkäuflichkeit. Das war sicher richtig in der klassischen Welt der Industrie, sieht aber auch dort inzwischen etwas anders aus: „Industrie 4.0“. Nun sind die Gelehrten, sind als wir vor gut 20 Jahren auf die Idee verfallen, eben auch unsere Leistung in dieser Weise darzustellen und sie in Messgrößen erfassbar zu machen.

Das hatte schon seine Richtigkeit. So kennen wir alle Kollegen, die wir nicht nennen wollen, die das Privileg, als UniversitätslehrerInnen nicht in der industriellen Weise einem Vorarbeiter, Meister, Abteilungsleiter etc. unterstellt zu sein, ausnützen und so gut wie nichts tun. Das kann aber nicht der Grund sein, ganze Hochschulen auf ein System der Leistungsmessung auszurichten, das die Anzahl Publikationen, die eingeworbenen Drittmittel, die Zitationen und die erhaltenen Ehrungen zu Messlatten macht, nach denen sich Belohnung und Strafe ausrichten. Nochmals zum Wissenschaftsrat und wohlgermerkt zu den Naturwissenschaften. Die Folgen sind innerwissenschaftlich negativ. Statt Erkenntnis wird Aktivität belohnt. Das führt zu Betrieb statt zu Einsicht. Und der Nachwuchs wird entsprechend selektioniert. Ressourcen werden investiert in Leistungserfassungs- und Kontrollsysteme. Und gegen aussen wird wissenschaftliche Tätigkeit mit Zahlen repräsentiert und legitimiert. Daran ist die Politik, die uns das Geld gibt, inzwischen gewöhnt, geradezu darauf konditioniert. Ohne Leistungsausweis keine Gelder für Wissenschaft. Gibt es dazu Alternativen? Im Moment gibt es eine um sich greifende Kritik, sie kommt von erfolgreichen Forschenden und Lehrenden, und sie richtet sich gegen die Leitung und gegen die Administration der Hochschulen. Das erwähnte Academic Manifesto sieht denn auch gerade darin einen Feind der Wissenschaft. Das deutsche Hochschulforschungsprojekt KORFU ist zu ähnlichen Schlüssen gekommen und empfiehlt, die Entscheidungsgewalt über die einzelnen Hochschulen von der Ebene der Schulleitung herunterzunehmen auf die Ebene der Fakultäten, um sie näher zur Wissenschaft zu bringen und um die nötige Vielfalt wieder möglich zu machen, die die Wissenschaft braucht.

„Post NPM“ steht an der Wand geschrieben, aber was die Zeichen für uns genau heissen, wissen wir noch nicht.

Zur Evaluation

Evaluation – ich liebe sie, habe sie 20 Jahre lang geübt, ich möchte sie nicht missen. Aber woher kommt sie und was meint sie wirklich? 20einhalb Jahre zuvor hatte ich Gelegenheit, zu lernen, was damals zu Evaluation gelernt werden konnte. Dabei erfuhr ich, dass ein Ursprung der Evaluation in der US-amerikanischen Sozialarbeit zu suchen sei. Programme zur Verbesserung der gesellschaftlichen Lage sollten darauf hin untersucht werden, ob sie eine Wirkung hatten und ob diese Wirkung in die intendierte Richtung zeigte. Dabei war Evaluation als eine Begleitung der Projekte gedacht, die einen Lernprozess erleichterte und auf beiden Seiten, den Evaluatoren, den Evaluierten und auch auf der dritten, der Zielgruppen der Massnahmen, Einsichten beförderte, formativ wirkte. Zugleich erfuhr ich von Schweizer Politikwissenschaftlern, dass in ihrem Feld Evaluation gewissermassen nachholend zu einer Profession gemacht werden sollte, die die Wirkungen von staatlichen Massnahmen messend bestimmen sollte und damit zur Rechtfertigung rationalen Verhaltens im Politischen beizutragen hatte. Das war ein Ziel eines Nationalen Forschungsprogramms „Wirksamkeit

staatlicher Massnahmen“ der 1990er Jahre, aus dem unter anderem die Schweizerische Evaluationsgesellschaft SEVAL hervorgegangen ist.

Wissenschaftsevaluation war von beiden Ideenwelten geprägt, aber mit keiner identisch. Mit der Politikevaluation hatte sie gemein, dass festgestellt werden sollte, ob das Geld der Steuerzahler sinnvoll eingesetzt wurde bei der Finanzierung von Programmen oder Institutionen. Mit der sozialen Tradition der Evaluation teilte sie den Anspruch, Lernprozesse anzustossen und Dinge besser zu machen, die Akteure zu ermächtigen, ihre Sache selbst und richtig in die Hand zu nehmen.

Schliesslich führten Bund und Kantone die Pflicht für die von ihnen finanzierten Hochschulen ein, sich den periodischen Überprüfungen ihrer Qualitätssicherungssysteme durch das Organ für Akkreditierung und Qualitätssicherung zu unterwerfen. Diese Überprüfungen fanden in Gestalt von Evaluationen statt. Unter anderer Bezeichnung wird das weitergehen unter dem HFKG.

Dass so etwas gelegentlich gut und nötig sein kann, glaube ich noch heute. Was ich aber sehe in den Hochschulen unter Evaluation, ist etwas ganz anderes. Es ist periodisch, nicht durch das Auftreten eines besonderen Problems motiviert, das nicht anders angegangen werden kann. Es geschieht in einem Klima des Misstrauens und reproduziert Herrschaftsverhältnisse zwischen oben und unten, draussen und drinnen, Habenden und von diesen Abhängigen. Es stimuliert Wohlverhalten mit dem unbestimmten Versprechen, dass dieses dereinst belohnt werden könnte. Und es wird von der Furcht geleitet, schlecht evaluierte würden Ressourcen verlieren, und nicht Evaluierte gälten draussen, bei den Geldgebern der Wissenschaft, als inexistent. Evaluation wird zu einem Instrument, das uns sichtbar machen soll.

An der Tür eines Kollegen hing lange Zeit ein Ausriss aus einem Comic. Die Situation war die Hölle. Der Oberteufel sagt, wir machen wieder einmal Evaluation. Er dreht den Gashahn auf, die Flammen lodern, die Kreaturen in der Hölle schreien auf, und der Oberteufel sagt befriedigt, es sind noch alle da.

Anders gesagt: Evaluation ist ein mächtiges Instrument, um in besonderen Lagen Probleme zu erkennen und Problemlösungen zu finden. Es ist kein Routinevorgang und dient nicht zum Controlling und nicht zur Qualitätssicherung. Dazu dienen Reviews und Assessments, Audits und wie die Verfahren alle heissen, die in der Industrie geläufig sind. Diese sind verbunden mit der Idee der Optimierung von Prozessen, die als Routinen beschrieben werden können. Wissenschaft ist das, wenn sie gut ist, gerade nicht.

Qualität, Leistung, sichtbar Machen des Guten, das an Hochschulen geschieht, meint etwas sehr Wichtiges. Wenn aber Mittel und Medien gewählt werden, die nicht das Wesentliche der Wissenschaft erfassen und vermitteln, sondern das Unwesentliche, die Routine, das einfach Beschreibbare, dann gehen die Ressourcen den falschen Weg und dann belohnen wir einen Schein von Wissenschaft, den Betrieb und die Betriebsamkeit. Es geht aber um Einsichten, Erkenntnisse, deren Ermöglichung und Weitergabe, das Wecken des Sinnes für das, was an der Grenze des bisher Begriffenen liegt. Hohe Erwartungen, hohe Werte, die wir nicht jeden Tag genügen und auch nicht immer alle von uns. Dort wo das geschehen kann, sollten wir nicht durch falsch gewählte Massnahmen erschweren oder verhindern, sondern durch Freiheiten fördern.

Das ist ganz schwierig zu vermitteln, aber sicher nicht in der inzwischen gewohnten Art der Kommunikation über Wissenschaft in Zahlen und Evaluationsberichten und schematisierten Empfehlungen. Wenn es geschieht, kann es aber miterlebt werden und im Dialog verständlich gemacht werden. Im Forschungsgesetz heisst es, der Bund „betreibe“ die ETH. Das sagt man von einem Produktionsunternehmen. Aktionäre betreiben eine Fabrik. Eine Hochschule wird geschaffen, wird gefördert mit dem Vertrauen, dass in ihr etwas geschieht, was sonst nur selten geschieht. Dass es Enttäuschungen gibt, ist menschlich, menschlich ist auch der Stolz auf die eigene Hochschule, wenn es gelegentlich funktioniert. Wir sind hier in der SGG, Geschichte hat ganz selten einen Marktwert wie die Kenntnis eines molekularen Vorgangs, der von der Industrie technisch verwertet werden und mit dem man Geld verdienen kann. Gemeinsam ist aber, dass falsche Anreize zu einem Betrieb führen, der gerade das verpasst, was von den Investoren erwartet wird, nämlich solides, überprüfbares Wissen zu generieren, das im Kern neue Erkenntnis, neue Einsicht ist. Entscheidend ist deshalb, dass nicht nur Resultate im Sinne von Fakten, die den Erwartungen entsprechen, „produziert“ und deren Produktion belohnt werden, sondern die interne Kritik. Erkenntnis ist immer vorläufig, und das sollte verhindern, dass wir ihr Warencharakter zusprechen.

Was tun

Es gibt den pragmatischen Weg: Der Hochschulleitung und damit der Trägerschaft und der Politik Zahlen und Prozeduren anzubieten. Wir haben dazu 20 Jahre lang geraten, wir müssen es jetzt auch tun. Unter dieser Prämisse ist Linderung der Last, die das mit sich bringt, angesagt, und Abwendung von Schaden, den falsche Kriterien und damit falsche Anreize anrichten. Man sollte sich bewusst sein, dass dieser Schaden in den Naturwissenschaften

bereits Tatsache ist und dass dort der Komplex aus Leistungsnachweis, Qualitätssicherung, Sichtbarmachung und Belohnung in der Krise ist. Man sollte in unseren Wissenschaften daraus lernen.

Ein Lernversuch ist der Rückgriff auf die Schaffung empirischer Daten. Das Projekt der Gruppe um Hans-Dieter Daniel zielt in diese Richtung: Durch Befragungen zu ermitteln, wie GeisteswissenschaftlerInnen selbst die Leistung und Qualität in ihren Fächern adäquat beurteilen (möchten). Der pragmatische Weg erhält dadurch eine Basis. Er setzt aber immer schon voraus, dass der erwähnte Komplex unausweichlich sei, und die empirische Studie erstreckt sich nicht auf die sogenannten unerwarteten oder perversen Effekte, die die Krise der Naturwissenschaften prägen. Hier geht es um Projekte, also letztlich, wenn ich das richtig verstehe, z.B. um die Begutachtung von Anträgen beim Nationalfonds, allenfalls um die Begutachtung von Manuskripten, die bei Zeitschriften eintreffen. Jedenfalls gewinne ich diese Vorstellung aus der Delphi-Umfragerunde, die mich Ende 2015 erreicht hat.

Oder die radikale Alternative: Kritik und Ablehnung des Komplexes, Aufbrechen der Systemlogik, die die perversen Effekte produziert. Was aber soll an dessen Stelle treten: hier sind die Antworten bisher eher schlagwortartig. Autonomie den Fakultäten oder den WissenschaftlerInnen statt der Hochschule als Unternehmen zuzugestehen. Vertrauen als Grundsatz im Verhältnis zwischen Geldgebern und Wissenschaft statt Misstrauen und Kontrolle. Nützlichkeit der Hochschulen nicht für die private Wirtschaft argumentieren, sondern für die Gesellschaft. Das Academic Manifesto hat dies gefordert, und so wurde in den Niederlanden, woher der Text stammt, eine Volksbefragung durchgeführt: Was erwarten Bürgerinnen und Bürger von der Wissenschaft. Daraus resultierte eine Liste mit Forschungsthemen, die nun teilweise staatlich gefördert werden. Diese Liste hat bedenkliche Aspekte, die ich hier nicht diskutieren kann. Das Grundproblem ist, dass Wissenschaft Autonomie, besser Freiheit braucht, um ihre eigentlichen Aufgaben gut erfüllen zu können. Die Legitimierung von Wissenschaft geschieht im öffentlichen Raum jedoch über deren Nutzen, also utilitaristisch. Und Wissenschaft selbst ist für die Öffentlichkeit eine Black Box.

Ich habe also Respekt vor den PragmatikerInnen. Denn sie möchten der Wissenschaft helfen, sich in einer NPM-Welt zurechtzufinden und ihre öffentlichen Ressourcen zu verteidigen. Kurzfristig können sie viel bewirken: Ende der routinemässigen Pseudo-Evaluation; Beurteilungen von Inhalten statt von Formen und Zahlen und Indikatoren als „Proxis“,

Entlastung des akademischen Personals von überbordendem Reporting, Beurteilungen über längere Zeiträume, damit die Wissenschaft den langen Atem bekommt, den sie braucht, um Resultate zu erlangen, den Disziplinen angemessene, spezifische Beurteilungskriterien etc. Aber langfristig ist vielleicht der Pragmatismus ein Teil eines Problems: Er kann nicht wirklich mit dem Risiko umgehen, dass unsere Wissenschaften in ähnlicher Weise umgestaltet werden die die Naturwissenschaften und damit in die Krise geraten, denn er hebt das System nicht aus den Angeln, das dort als geschlossener Regelkreis die Krise immer weiter vertieft. Es ist der alte Gegensatz zwischen Reformern und Revolutionären.