

Risiken und Nebenwirkungen der Leistungsmessung von Professoren: Warum wir sie brauchen und warum wir dennoch aufpassen müssen

Kaum ein Thema hat die Wirtschaftspresse in den letzten Monaten mehr beschäftigt als die Frage nach einer leistungsgerechten Vergütung von Managern. Während die Forderung schnell ausgesprochen ist, ist die Frage nach der adäquaten Definition und Messung von »Leistung« alles andere als trivial: Wann ist die Leistung eines Managers als hoch zu bewerten, wann ist sie niedrig? Die Unternehmensperformance, wie auch immer man sie im Einzelnen misst, mag als Indikator für die Management-Leistung zwar nahe liegen, ist aber offensichtlich auch von anderen, exogenen Faktoren beeinflusst. Natürlich kann man versuchen, diese exogenen Faktoren bei der Bemessung der Vergütung in irgendeiner Form zu berücksichtigen (etwa indem man relative Leistungsindikatoren heranzieht und die Unternehmensperformance bzw. deren Entwicklung im Zeitablauf im Vergleich zum Markt oder zur Branche betrachtet), aber damit wird man wohl regelmäßig nur einen Teil des Problems erfassen: Interessant wäre letztlich die Beantwortung der kontrafaktischen Frage, wie sich denn das betreffende Unternehmen entwickelt hätte, wenn es anders geführt worden wäre.

Auch für die Personengruppe der Professoren ist – nicht zuletzt im Rahmen der Professorenbesoldungsreform – wiederholt diskutiert worden, wie für diese denn adäquate Leistungsanreize gesetzt werden könnten. Während einerseits grundsätzlich in Frage gestellt wird, ob Professoren denn überhaupt über Leistungsanreize gesteuert werden können und sie nicht vielmehr intrinsisch motiviert sind (eine Motivation, die möglicherweise durch extrinsische Anreize sogar noch verdrängt werden könnte), stellt sich andererseits die Frage, wie denn ein sinnvolles System der Leistungsmessung für Professoren aussehen könnte – wenn man denn die Annahme positiver Anreizeffekte leistungsorientierter Vergütung für die Gruppe der Professoren nicht *a priori* ausschließen will. Wie wäre demnach die »Leistung« eines Professors zu definieren und wie sollte sie gemessen werden? Zunächst ist dabei zu beachten, dass Professoren ihre Leistung, grob vereinfacht, in drei verschiedenen Teilbereichen erbringen: in der Forschung, in der Lehre und in der

Selbstverwaltung. Mit anderen Worten, der Arbeitsplatz eines Professors (ähnlich wie der des Managers) ist durch das gekennzeichnet, was man gemeinhin als »Multi-Tasking« bezeichnet. Für Multi-Tasking-Arbeitsplätze ist jedoch bekannt, dass man mit der Einführung von leistungsorientierten Vergütungsbestandteilen extrem vorsichtig sein muss, da sie letztlich nur dann gelingen kann, wenn adäquate Anreize für *jede* der zu leistenden Aufgaben gesetzt werden. Versieht man hingegen nur eine Aufgabe mit Leistungsanreizen, dann wird ein solchermaßen angereizter Hochschullehrer (so weit jedenfalls die Theorie) nur noch in dieser einen Anstrengungsdimension etwas leisten – es sei denn, er ist intrinsisch motiviert (siehe oben), sich auch in den anderen, nicht angereizten Dimensionen zu engagieren, und/oder die Aufgaben sind insofern miteinander verknüpft, dass sich eine Mehrleistung in einer Aufgabe auch auf die Leistung in der anderen vorteilhaft auswirkt, das heißt, dass es so etwas wie Komplementaritäten zwischen den Aufgaben gibt. Aber wie könnten adäquate Leistungsmaße für die drei Leistungsbereiche von Professoren aussehen?

Beginnen wir mit der *Forschung*, denn hier scheint es noch die klarsten Leistungsindikatoren zu geben: Man kann Publikationen auszählen und dabei Seitenzahlen und die Zahl der Koautoren berücksichtigen. Man kann darüber hinaus versuchen, die Qualität der Beiträge über Zitationen zu erfassen oder über die unterschiedliche Gewichtung von Publikationsorganen. Für letztere benötigt man allgemein akzeptierte Rankings, wobei wir in der Betriebswirtschaftslehre – trotz des soeben veröffentlichten Handelsblatt-Journal-Rankings – immer noch weit davon entfernt davon zu sein scheinen, uns auf eine allgemein akzeptierte Liste einigen zu können; ein »deutscher SSCI« zur ergänzenden Erfassung von Impact-Faktoren der für die deutschen Betriebswirte immer und zu Recht noch wichtigen deutschsprachigen Zeitschriften wäre wohl die – zugegebenermaßen wenig realistische – Lösung. Ein anderer und aktuell immer bedeutsamer werdender Weg zur Erfassung der Forschungsleistung von Professoren besteht darin, nach der Einwerbung von Drittmitteln und dabei insbesondere nach der von hoch reputierlichen Drittmitteln wie etwa der DFG zu fragen. Während Drittmittel ganz offensichtlich ein Input-Faktor für die Forschung sind, ist die Frage, ob sie auch als Output-Maß zurückliegender Forschungsleistungen heran-

gezogen werden können, weniger eindeutig zu beantworten. Unzweifelhaft erhöht zwar eine hohe Forschungsaktivität eines Wissenschaftlers die Chance auf die Bewilligung weiterer Projekte, aber da die Akquirierung von Drittmitteln nicht für jeden Forscher von substantiellem oder gar existentiellm Interesse ist, stellt die Generierung von Drittmitteln letztlich kein eindeutiges Signal für Forschungsaktivitäten und -erfolg dar. Dennoch: Bei aller Problematik, die in der adäquaten Messung von Forschungsleistung liegt, lässt sich diese durch die Vielfalt der verfügbaren Indikatoren am Ende wahrscheinlich immer noch ganz gut approximieren – jedenfalls im Vergleich zu den anderen Leistungsdimensionen.

So sieht es etwa bei der Erfassung der Leistung eines Professors in der *Lehre* völlig anders aus. Idealerweise würde man wissen wollen, welchen Beitrag ein Dozent dazu leistet, dass Studierende auf ihre zukünftigen Aufgaben in Wirtschaft, Wissenschaft, Politik oder Gesellschaft vorbereitet werden, und vielleicht auch darin, dass er den Studierenden dabei hilft, überhaupt erst zu erkennen, in welchem Bereich sie die höchsten Beiträge werden leisten können. Man sieht schon an der Fragestellung, dass eine Erfassung dieser Dimension kaum zu erreichen sein dürfte. Auch eine spätere Befragung der Alumni dürfte hier kaum Aufschluss geben: Zum einen würde ein nennenswerter zeitlicher Abstand benötigt, um überhaupt die relevanten Informationen abfragen zu können, zum anderen würde aber gerade dieser zeitliche Abstand verhindern, dass ein Alumnus seine Lernerfolge auf bestimmte Dozenten oder konkrete Veranstaltungen zurückzuführen in der Lage wäre. Natürlich könnte man sich hier auch auf ein Studienprogramm/einen Studiengang beziehen anstatt auf einen einzelnen Dozenten, mit dem Argument, dass es sich bei der Vorbereitung der Studierenden auf ihre zukünftigen Aufgaben letztlich um ein Teamprodukt handelt, an dem mehrere Dozenten mitwirken. Eine solche Interpretation würde es dann auch möglich machen (jedenfalls im Prinzip), dass man die im Rahmen eines Studiengangs erbrachte Team-Lehrleistung der Dozenten dadurch erfasst, dass man den Berufserfolg der Absolventen unterschiedlicher Studiengänge an unterschiedlichen Universitäten miteinander vergleicht und nach signifikanten Unterschieden sucht (d. h. es ginge dann auch hier um die relative Performance). Natürlich wäre dabei das möglicherweise auftretende Trittbrettfahrer-Prob-

lem bei einer auf dem Teamergebnis beruhenden Vergütung der Professoren noch nicht gelöst. Des Weiteren wären unterschiedliche Ausgangsbedingungen zu kontrollieren (unterschiedlich gute Studierende, unterschiedliche regionale Arbeitsmarktlagen, unterschiedliche – nur sehr langfristig veränderbare – Reputation der Institutionen). Am Ende würde man also wieder nach einer Antwort auf eine kontrafaktische Frage suchen: Was wäre aus den Studierenden geworden, wenn sie eine andere Ausbildung genossen hätten? Die Idee, auch Lehrevaluationen der aktuellen Studierenden als Indikator für die Lehrleistung heranzuziehen, mag dabei zwar nahe liegen, erweist sich bei näherem Hinsehen aber als wenig hilfreich: Es gibt nämlich Hinweise darauf, dass die Evaluation der Lehrleistung eines Dozenten empirisch eng verknüpft mit der Beantwortung der Frage ist, ob der präsentierte Stoff als einfach oder schwierig empfunden wird. Würde man entsprechend die Ergebnisse von Lehrevaluationen zur Leistungsmessung (und womöglich -vergütung) heranziehen, wäre es für einen Hochschullehrer ein Leichtes, seine Evaluationen durch Reduktion des Anspruchsniveaus in seinen Veranstaltungen aufzubessern. Alternativen wie etwa die Vereinheitlichung des Lehrstoffes und die Einführung zentraler Prüfungen sind im Hochschulbereich – angesichts des Grundsatzes der Freiheit von Forschung und Lehre – nicht verfügbar. Mit anderen Worten, eine an der Lehrleistung der Professoren orientierte Entlohnung wird es kaum geben können, wenn man nicht Anreize zur Senkung des Anspruchsniveaus induzieren möchte. Wenn und insofern jedoch Komplementaritäten zwischen Forschung und Lehre bestehen (woran zumindest Humboldt glaubte und wofür es – jedenfalls für den Bereich von Fortgeschrittenen-Lehrveranstaltungen – auch empirische Hinweise gibt), bliebe zu hoffen, dass die einseitige Anreizung von Forschungsleistungen am Ende zumindest *auch* der Lehre zu Gute kommt.

Bliebe die Frage nach der Leistungsmessung von Hochschulprofessoren im Bereich der *Selbstverwaltung*. Nichts scheint einfacher zu realisieren. Denn schließlich kann man leicht erfassen, ob ein Kollege ein Amt im Rahmen der akademischen Selbstverwaltung innehat oder nicht. Wie stark er sich hier engagiert, ist schon wieder ungleich schwerer zu messen. Auch ist nicht klar, wie unterschiedliche Ämter zu gewichten wären. Was bedeutet etwa eine Tätigkeit als Dekanin, was eine

Tätigkeit als Studiendekan? (Dem Land Baden-Württemberg etwa ist die Übernahme des Dekansamtes in der C-Besoldung aktuell brutto 66,58 € im Monat wert; Studiendekane erhalten keine Funktionszulage; die W-Besoldung eröffnet hier allerdings mehr Spielräume.) Wie ist die Übernahme von Ämtern und Aufgaben in der Scientific Community zu gewichten, etwa die Organisation wissenschaftlicher Tagungen, der Vorsitz in wissenschaftlichen Fachverbänden, die Erstellung von Fachgutachten oder die Herausgabe wissenschaftlicher Fachzeitschriften? Hier wird – typischerweise ehrenamtlich – ganz erheblich in Öffentliche Güter investiert, die der Scientific Community zugute kommen. Kann man *bemessen*, was diese Tätigkeiten im Einzelnen wert sind? Ist die Scientific Community, sind Berufungskommissionen bereit, diese Investitionen zu honorieren? Im Moment scheint dies nicht der Fall zu sein. Auch hier ist zwar nicht auszuschließen, dass es Komplementaritäten zwischen Forschungsleistungen und Aufgaben im Rahmen der Scientific Community gibt (etwa wenn es um die Organisation von Tagungen, die Erstellung von Gutachten oder die Herausgabe von Fachzeitschriften geht; für Aufgaben im Fakultätsmanagement dürfte dies weniger gelten). Wenn es aber – anders als im Bereich der universitären Lehre – keine Verpflichtung für die Übernahme von Aufgaben in der akademischen Selbstverwaltung und der Scientific Community gibt, dann dürften sich hier ganz stark auch die substitutiven Beziehungen zwischen den Aufgabenbereichen auswirken: Zeit steht eben nur einmal zur Verfügung.

Es wird derzeit gezählt, gewichtet und gerankt. Die einzige Leistungsdimension, die dabei aktuell erfasst wird, sind die Publikationen in ausgewählten und bei weitem nicht allen für unsere Profession relevanten Fachzeitschriften sowie die Einwerbung von Drittmitteln. Nachdem bekanntlich allein schon das Sammeln und Veröffentlichen von Kennzahlen Druck auf die Betroffenen ausübt, ohne dass es dabei überhaupt zu Gehaltskonsequenzen kommen muss, sollte man sich nicht wundern, wenn es immer schwieriger wird, Kollegen für die Übernahme von Aufgaben in der Scientific Community zu gewinnen und wenn irgendwann auch die Lehrleistung (insbesondere im Bereich der grundständigen Lehre, in der Komplementaritäten zu Forschungsaktivitäten wenig stark ausgeprägt sein dürften) in ihrer Qualität abnehmen.

Dabei ist an dieser Stelle keineswegs intendiert, einem Verzicht auf die Leistungsmessung im Hochschulbereich das Wort zu reden (wenn wir es so vehement für andere fordern, sollten wir wohl mit gutem Beispiel voranzugehen bereit sein) – im Gegenteil. Aber wir sollten nicht den Fehler machen, auch noch aktiv daran mitzuwirken, dass unser Berufsstand nur noch auf eine einzige, wenngleich wichtige, Dimension reduziert wird: Das Aufgabenfeld des Hochschullehrers ist vielschichtiger als es die derzeitige Diskussion um Zeitschriftenrankings und Drittmittel-Statistiken wiedergibt.

Kerstin Pull