

## 8. Forschungs- und hochschulpolitisches Werkstattgespräch

### „Gemessen, vermessen – und zu leicht befunden? Über Sinn und Zweck von quantitativen Indikatoren für die Hochschulbewertung“

Schloss Herrenhausen, 07. bis 08. Oktober 2016

Bericht: Dr. Vera Szöllösi-Brenig

-----

Die „Vermessung der Wissenschaft“ ist voll im Gange. Nach der Humboldt'schen Universitätskonzeption, der Ordinarien- und später der Gruppenuniversität begann vor rund 20 Jahren der Einzug des New Public Management (NPM) in die deutsche Hochschulpolitik. Doch nicht nur die Wissenschaftlerinnen und Wissenschaftler werden ständig evaluiert und auf ihre Effizienz überprüft, sondern auch die Hochschulen selbst stehen – und zwar international – in dauerndem Konkurrenzdruck: Rankings, Ratings, breit gefächerte Indikatorensysteme und eine immer weiter um sich greifende Tendenz, qualitative Bewertungen durch quantitative Angaben zu ersetzen, bestimmen das Bild. Doch welche Messzahlen können welche Aussage machen und welche Folgen hat es, wenn sich hochschulpolitisches Handeln – einzig oder primär – an diesen Indikatoren ausrichtet? Wie frei, autonom und kreativ kann die vermessene Universität noch sein? Diese Fragen standen im Zentrum des 8. forschungs- und hochschulpolitischen Werkstattgesprächs der VolkswagenStiftung in Herrenhausen. In seinem Themenaufriß listete Wilhelm Krull als Gastgeber der Veranstaltung gleich eine ganze Reihe von „unintended consequences“ dieser Entwicklung auf und fragte nachdenklich, ob es sich bei diesen Indikatoren eigentlich um mehr als ein „Aktivitätsmaß“ handle mit stark selbstzweckhaften Zügen. „Wo ist der für die Wissenschaft nötige Freiraum für vorbeugendes Nachdenken und forschendes Fragen?“ Gerade mit Blick auf die bevorstehende dritte Runde der Exzellenzinitiative stellen sich diese Fragen mit neuer Dringlichkeit.

In ihrer Keynote „Gemessen – Vermessen. Von Helden und Halunken, Königen und Knechten“<sup>1</sup> identifizierte MARGIT OSTERLOH (Zürich/Basel) vier denkbare Governancetypen der Wissenschaft, je nachdem, ob eine Ausrichtung an der Öffentlichkeit oder an der Scientific Community geschieht. Als Typ (1) bezeichnete sie die „Steuerung durch Experten“. Er entspreche Polanyis Forderung „The soil of academic science must be exterritorial“ und stelle eine Art Oligarchie der Peers dar. Dieser Steuerungstyp habe aber drei Voraussetzungen: Gemeinwohlorientierung, verlässliche Prognosen und verlässliche Peer Reviews, die, wie vielfache Studien<sup>2</sup> gezeigt hätten, in der Realität nur bedingt erfüllt würden. Wenn Vielfalt und Dezentralität gesichert seien, so Osterloh, sei diese Steuerungsform jedoch unverzichtbar. Bei Typ (2) „Steuerung durch Indikatoren“ hingegen sei es die Öffentlichkeit, auf die das Wissenschaftssystem durch Rechenschaftspflicht über die Arbeitsergebnisse ausgerichtet sei – Stichwort „New Public Management“. Kompensiere

---

<sup>1</sup> Nach Julian Le Grand, *Motivation, Agency, and Public Policy: Of Knights and Knaves, Pawns and Queens*, 2004.

<sup>2</sup> Z.B. Tetlock, Philip E., *Expert Political Judgment: How Good Is It? How Can We Know?* 2005; Böhmer, Susan/ Hornbostel, Stefan/Meuser, Michael, *Postdocs in Deutschland: Evaluation des Emmy Noether-Programms*, 2008.

dieser Steuerungstyp auch die Nachteile von der Oligarchie der Peers (1), so gebe es auch hier strukturelle Voraussetzungen, die nicht erfüllt würden oder auch per se nicht erfüllbar seien: Verlässliche Peer Reviews ohne Zitationskartelle und ohne Gender- und Institutionenbias<sup>3</sup>; korrekte Aggregationsprozesse, valide Impactfaktoren, etc. Hinzu komme das Phänomen, dass – wie Frost & Brockmann 2015 gezeigt haben<sup>4</sup> – jeder Indikator über die Zeit seinen Aussagewert verliere. Das Performance-Paradox sei bei diesem Steuerungstyp besonders groß: Neben der Verschiebung der Motivation (von intrinsisch zu extrinsisch), Gaming (z.B. Slicing der Publikationsergebnisse etc.)<sup>5</sup> und „HARKING-Effekten“<sup>6</sup> auf der Ebene des einzelnen Wissenschaftlers nannte Osterloh als negative Folge auf institutioneller Ebene einen nicht überwindbaren Lock in-Effekt wie die Umverteilung von Ressourcen an „Leuchttürme“. Typ (3) „Steuerung durch Rankings“ gebe Politikern eine leichte Orientierungshilfe, unterminiere jedoch den nach Merton notwendigen „organisierten Skeptizismus der Wissenschaft“<sup>7</sup> und verstärke die Lock in-Effekte. Deswegen bezeichnete ihn Osterloh auch als „Governance by Truthiness“, denn hier würden einmal gemachte Feststellungen zu intuitiven Wahrheiten, die nicht mehr hinterfragt würden und schließlich zu falschen Ergebnissen führten<sup>8</sup> – von Absurditäten wie der Einkaufstour von Wissenschaftler-ViPs durch die King Saud Universität ganz zu schweigen. Vor diesem Hintergrund machte sich Osterloh für Typ (4) „Steuerung durch Demarchie“ stark – ein partiell eingesetzter kontrollierter Zufall helfe gegen Biases, begünstige Diversität und gebe kreativen Außenseitern eine Chance<sup>9</sup>. Dieser Steuerungstyp setze allerdings eine qualitative Vorselektion und eine quantitative Begrenzung, z.B. auf 30% der Mittel, voraus. Als konkretes historisches Beispiel nannte Osterloh die Professorenwahl an der Universität Basel im 18. Jahrhundert – drei Kandidaten wurden ausgewählt, dann aber entschied das Los.

Nach diesem historischen Beispiel führte der nächste Vortrag in die Gegenwart: Staatssekretär THOMAS GRÜNEWALD (Düsseldorf) berichtete über den Steuerungsansatz in Nordrhein-Westfalen. Steuerung im Wissenschaftsbereich sei für ihn ein zweistufiges System: Auf der Basis von Indikatoren müsse eine qualitative Bewertung erfolgen, die zur Entscheidung führe. Grünewald strukturierte seinen Vortrag in vier Fragen. Allen voran: „Wie frei ist die vermessene Universität“? Die Leistungsorientierte Mittelvergabe (LOM) beispielsweise sei in NRW 1999 eingeführt worden und biete reiche Anschauung für die natürlichen Grenzen eines quantitativen Modells in der Lebenswirklichkeit von Hochschulen. Je differenzierter die Hochschulen geworden seien, je unterschiedlicher die Profile ausgestaltet seien, je verschiedener die Hochschulen nach Größe und Spezifik geworden seien, desto weniger taue ein wettbewerbliches Umverteilungssystem, das der Diversität der Hochschulen nur unzureichend Rechnung tragen könne. Folge sei, dass die Parameter und ihre Gewichtung fortwährend angepasst würden und dass die Akteure in den Hochschulen nur unzureichend den Zusammenhang zwischen ihrer individuellen Leistung – oder auch der Leistung ihrer Organisationseinheit – und dem Abschneiden ihrer Hochschule in der leistungsorientierten Mittelverteilung nachvollziehen könnten. Die zweite Frage nach

---

<sup>3</sup> Z.B. Bornmann, Lutz, et al, „A persistent problem. Traditional gender roles hold back female scientists“, in: *EMBO Reports* 8 (11), 2007.

<sup>4</sup> Frost, Jetta/Brockmann, Julia, „When quality is equated with quantitative productivity – Scholars caught in a performance paradox“, in: *Zeitschrift für Erziehungswissenschaft* 2015.

<sup>5</sup> Osterloh, Margit/Frey, Bruno S., „Ranking Games“, in: *Evaluation Review* 1 (8), 2014.

<sup>6</sup> HARKING = Hypothesizing After Results are Known.

<sup>7</sup> Merton, Robert K, *Social Theory and Social Structure. Toward the codification of theory and research*, 1949.

<sup>8</sup> Vgl. Ioannidis, John P. A., „Why Most Published Research Findings Are False“, in: *PLoS Med* 2(8), 2005.

<sup>9</sup> Vgl. Osterloh, Margit, „Würfelt Gott? Würfelt die Wissenschaft?“, in: *Beiträge zur Hochschulforschung* (im Druck).

den gesellschaftlichen Anforderungen, denen eine Universität heute gerecht werden müsse, wollte Grünewald nicht im Widerspruch zur Freiheit der Forschung sehen. Auf das Ganze der öffentlichen Mittel für Forschung und Entwicklung in den Hochschulen Deutschlands gesehen, würden allenfalls 10 % der Mittel programmgebunden und somit auf definierte Themenbereiche hin festgelegt werden, wohingegen der weitaus größte Teil der Forschungsförderung und zumal die Mittel der DFG – wie auch die des European Research Council – programmungebunden vergeben werden. Dritte Frage: Wie werde Steuerung konkret in NRW umgesetzt? Im Landeshochschulentwicklungsplan (LHEP), den das Land erstmals im September 2016 vorgelegt habe und der zum 1. Januar 2017 in Kraft trete<sup>10</sup>, würden neun strategische Handlungsfelder für die nächsten fünf Jahre formuliert, zu denen sich jede Universität verhalten solle – aber es gehe mitnichten um die mechanische Umsetzung von Vorgaben. Vierte und letzte Frage: Wie könne sich die einzelne Universität nun auf die künftigen Aufgaben vorbereiten? Es müsse unter anderem darum gehen, weniger hochspezialisierte Bachelorstudiengänge zu entwickeln, diese müssten eher grundlegend gedacht werden. Relevanz müsse hier ein wichtiges Stichwort sein. Die Aufgabe der Wissenschaftspolitik sah Grünewald wiederum darin, der Wissenschaft im Interesse ihres Beitrags zur Zukunftssicherung ihre privilegierte Position in der Gesellschaft zu sichern. Bei der Grundfinanzierung der Universitäten jedoch müsse nachgesteuert werden. Die jetzige Schieflage sei entstanden, als die Politik die Verpflichtung, 3 % des Bruttoinlandsprodukts in Forschung und Entwicklung zu investieren, einhalten wollte. Dies sei aus guten Gründen über die Erhöhung der öffentlichen Ausgaben für Forschung und Entwicklung erfolgt; diese wiederum würden als Drittmittel in wissenschaftsgeleiteten Wettbewerben auf der Basis von Peer Reviewing allokiert. Im Zuge des Anstiegens der Drittmittel seien die infrastrukturellen Grundmittel, so Grünewald in seinem Fazit, in zu geringem Maße angepasst worden.

Die Position des Staatssekretärs wurde von zwei Rektoren von Hochschulen in NRW kommentiert. AXEL SCHÖLMERICH (Bochum) betonte zunächst, dass „wir in Deutschland immer noch auf einer Insel der Seligen leben“ – die staatlichen Universitäten von Kalifornien beispielsweise würden künftig nur mehr zu 12 % öffentlich finanziert. Danach jedoch zählte Schölmerich eine Reihe von Gegenpositionen zu Grünewald auf: Er forderte den Abschied vom New Public Management im Bereich der Wissenschaft und sprach sich insbesondere gegen das von der NRW-Landespolitik formulierte Ziel aus, die Zahl der Universitätsabsolventen zu erhöhen – dies sei kein geeigneter Weg, die soziale Ungleichheit zu verringern. Stattdessen sollte man anhand der Erwerbsbiographien von Absolventen den langfristigen Erfolg eines Studiums untersuchen. Die eingeführte LOM wiederum sei für ihn „kein Kabarett“, sondern bitterer Ernst, schließlich verliere die Universität Bochum jährlich den der Kappungsgrenze von 2,1 Mio. EUR entsprechenden Betrag. ERNST SCHMACHTENBERG (RWTH Aachen) kommentierte den Vortrag des Staatssekretärs aus der Perspektive einer sehr großen Universität und zugleich aus der Perspektive der Ingenieurskultur, von der er geprägt sei: Forschung sei für ihn nicht zweckfrei, sondern zur Lösung von Problemen gedacht. Aus diesem Grunde sei die RWTH in Deutschland die Universität mit der meisten Zusammenarbeit mit der Wirtschaft. Ihm gehe es eigentlich nicht um Transfer als eine dritte Mission der Universität, sondern um konkretes Tun als Teil der Wissenschaft. Warum beispielsweise habe jede medizinische Fakultät eine Klinik, aber den Technikfakultäten seien keine eigenen Fabriken zugeordnet? Die Politik forderte Schmachtenberg auf, klarer zu sagen, was sie wolle – zum Beispiel im LHEP. Als Rektor

---

<sup>10</sup> <https://www.land.nrw/de/pressemitteilung/hochschulen-und-land-gestalten-gemeinsam-zukunft-und-entwicklung-des>.

stehe er doch vor derselben Frage wie die Politik, wenn er die Mittel unter den Fakultäten aufteilen müsse. Gleichzeitig sprach er sich gegen die im LHEP formulierte Zielvorgabe aus, die Aufnahmekapazitäten zwischen Fachhochschulen und Universitäten im Verhältnis von 40 zu 60 festzusetzen, ohne dabei auch die Finanzierung zu betrachten: Da die Hochschulen nach den Bachelorzahlen bezahlt würden, werde diese Vorgabe automatisch zu einer Reduzierung der Mittel für die Universitäten führen.

In der folgenden Diskussion ging es zum einen um die Bedeutung von Indikatoren für die Steuerung des Systems. Der Wert von Zahlen wurde vor allem darin gesehen, Komplexität so zu reduzieren, dass Entscheidungen – sowohl innerhalb der Universität als auch seitens der Wissenschaftspolitik – politisch legitimiert werden können. Sämtliche Entscheidungen dürften aber das soziale System nicht in Frage stellen. Beispielsweise seien die Haushalte der Universitäten in den 1970er Jahren de facto festgeschrieben worden, was jede LOM in ihrer Wirkung per se reduziere. Dennoch hätten alle Hochschulen in NRW nicht mehr das alte Gepräge, und die Zusammenlegung von Duisburg-Essen könne regelrecht als Erfolgsbeispiel verbucht werden. Einigkeit wurde unter den Anwesenden darin erzielt, dass Indikatoren keinen operativen Zweck haben könnten, um das Finetuning an einer Hochschule durchzuführen. An der ETH Zürich, an deren Exzellenz kein Zweifel bestehe, gebe es beispielsweise ein elaboriertes Evaluationssystem, jedoch werde an der Universität nichts von außen gesteuert. Dafür gebe es dort eine klare Eingangssteuerung: Durch eine Maturaquote von nur 20 % – im Verhältnis zu rund 50 % in Deutschland – sei die Qualität der Studierenden gesichert. In diesem Punkt wurden die Unterschiede in den Positionen von Wissenschaftspolitik auf der einen Seite und Universitäten auf der anderen Seite deutlich, ebenso bei der Bewertung der von NRW geplanten Output-Steuerung durch die Zahl von Absolventen. Die hier zu erwartenden „unintended consequences“ – möglichst viele Absolventen auf Kosten der Qualität – wurden jedoch von den Rektoren und Präsidenten nicht klar benannt; es wurde nur darauf hingewiesen, dass die Wissenschaft ein „Ort des systematischen Marktversagens“ sei, da Wissenschaft öffentliche Güter erzeuge, durch Serendipitätseffekte von Unsicherheit geprägt sei und schließlich „Vertrauensgüter“ produziere, deren Güte und Validität sich oft erst nach langer Zeit herausstelle. Dieser Zeitfaktor wurde in der Diskussion auch bei sämtlichen Steuerungsversuchen als kritisch festgestellt: Der Erfolg einer Steuerung stelle sich im Wissenschaftsbereich erst nach 10 bis 15 Jahren heraus. Funktionierende Märkte gebe es im Wissenschaftsbereich ausschließlich bei der Berufung neuer Kolleg(inn)en; diesen müssten hinreichend Mittel gegeben werden, damit sie nicht gleich Anträge schreiben müssten. Unterschiedliche Akzente wurden von Politik und Wissenschaft am Schluss der Diskussion gesetzt: Während ein Rektor die Abschaffung der Kapazitätsverordnung durch die Politik forderte, machte der Staatssekretär deutlich, dass Hochschulen in Deutschland Instrumente der regionalen Strukturpolitik seien.

Am zweiten Veranstaltungstag wurden Innensichten aus verschiedenen Hochschulen im Umgang mit Kennzahlen vorgestellt. KATHARINA KRAUSE, Präsidentin in Marburg, stellte ihren Ausführungen die These von Foucault voran, dass Kennzahlen immer Machttechnologien seien. Sie erweckten den Anschein von Objektivität, folgten aber expliziten und impliziten Setzungen. Neben der Politik würden Zahlen u.a. auch von der DFG und von den Medien gefordert. Universitätsintern würden Zahlen beispielsweise bei der Budgetierung, der Studiengangsplanung, dem Flächenmanagement und der Zulagen der W-Besoldung genutzt. Die politischen Vorgaben hätten zur überkapazitären Aufnahme von Studierenden geführt. De facto gebe es nur bei 130%iger Auslastung 100% Geld. Durch die Gleichbehandlung aller in der Basisfinanzierung gebe es für die Methodik – trotz der bekannten Probleme – eine hohe Akzeptanz, auch weil dadurch sowohl inneruniversitärer

Streit vermieden werde als auch keine diskretionäre Zuweisung durch das Präsidium erfolge. Die erfolgreiche Einwerbung von Drittmitteln honoriere die Hochschulleitung in Marburg im Übrigen durch Weiterleitung eines großen Anteils an Overheads an die erfolgreichen Wissenschaftler(innen), um so die Projektleiter(innen) handlungsfähig zu machen und ihnen ein inneruniversitäres Antragsverfahren für die notwendige Zusatzausstattung zu ersparen.

SASCHA SPOUN (Lüneburg) stellte folgende Grundsatzüberlegung in den Mittelpunkt: Die Forderung nach permanenter Evaluation der Wissenschaft sei kontraproduktiv, denn sie schränke die Freiheit ein, lasse zu wenig Zeit, um langwierige und langfristige Erfolge zu erzielen und sei ein Ausdruck von fehlendem Vertrauen – und ohne Vertrauen gebe es seiner Überzeugung nach keine Leistungsfähigkeit. Derzeit fehlten vor allem Mechanismen, mit denen sich die Wissenschaft der Politik verständlich machen und für Vertrauen werben könne. Für ihn als Präsident einer jungen, kleinen Universität außerhalb der Top 10 hätten manche Kennzahlen nur bedingten Aussagewert, denn es gehe nicht darum, diesen nachzulaufen, sondern um ein eigenes Profil der Universität. Er frage sich, welche „anderen“ Indikatoren für seine Universität relevant seien: beispielsweise die Anzahl von Studienstiftern unter den Studierenden, die Anzahl erfolgreicher Ausgründungen oder die Bewerberzahlen von Bildungsausländern. Spoun mutig: „Wir gehen in einen nicht-messbaren Experimentalmodus und leisten so einen Beitrag zu einer vielfältigen Hochschullandschaft.“

WOLFGANG-UWE FRIEDRICH (Hildesheim) zeichnete in seinem Vortrag die Entwicklung der Hochschulfinanzierung in Niedersachsen seit Ende der 1990er Jahre nach – und damit die Entwicklung der Indikatoren für die Finanzierung der Hochschulen in ganz Deutschland. Nach der Entwicklung des New Public Management-Konzepts in den 1990er Jahren<sup>11</sup> habe Thomas Oppermann (SPD) als Wissenschaftsminister 2002 das „Hochschulreformgesetz“ beschlossen, das das Stiftungsmodell für Hochschulen entwickelte und den Globalhaushalt und neue Leitungsstrukturen an Universitäten einführte: Unter dem Schlagwort „Mehr Autonomie für die Hochschulen“ sei das Ziel die output-orientierte Steuerung durch die Formelfinanzierung gewesen. Sein Nachfolger Lutz Stratmann (CDU) habe diese Entwicklung 2005 in einem "Hochschuloptimierungskonzept" weitergeführt, das zu Einsparungen in Höhe von 50 Mio. EUR – gegenfinanziert durch 90 Mio. EUR Studiengebühren – geführt habe. Das von der Politik behauptete Ziel, den Wettbewerb zu fördern, habe bedingten Erfolg erzielt; die Verringerung der Detailsteuerung sei nicht erreicht worden, jedoch der Aufbau von mehr Daten und Kennzahlen und damit ein Mehr an Transparenz durchaus: „Das hat gut geklappt“. Die von Niedersachsen zwischen MWK und Landeshochschulkonferenz entwickelte „Formel“ sei in ähnlicher Form in allen deutschen Ländern übernommen worden wie beispielsweise die Einteilung in drei Formelfächergruppen und die Mittelverteilung 48 % Lehre (Auslastung = 28%; Absolventen nach RZ = 68%; Bildungsausländer = 2%; outgoing students = 2%), 48 % Forschung (Drittmittel = 74%; Promotionen = 24%; AvH Stipendiaten und Preisträger = 2%) und 4 % Gleichstellung (weibl. Wiss.Pers. = 20%; Neuber. Prof. = 40%; Promotionen/w = 20%; Absolventinnen = 20%). Auch wenn die Formel insgesamt zu einer Leistungssteigerung geführt hat, so Friedrich in seinem Fazit, sei das Qualitätsproblem z.B. bei Promotionen dadurch nicht zu lösen.

Über den Profilbildungsprozess an der Universität Leipzig berichtete MATTHIAS SCHWARZ, Prorektor für Forschung und Nachwuchsförderung. Top down und Bottom up-Steuerung

---

<sup>11</sup> Budäus, Dietrich/ Conrad, Peter./ Schreyögg, Georg (Hrsg.), *New Public Management*, Berlin/New York, 1998. Wissenschaftlicher Beirat (Hrsg.), *Ein neues Verfahren der staatlichen Mittelvergabe an die Universitäten in Niedersachsen: Anforderungen, Prinzipien und Umsetzungsvorschläge*, 1998.

habe man in Leipzig als eine Art „Gegenstrommodell“ begriffen. 24 Vorschläge für Profillinien seien 2013 von der Wissenschaft gemacht worden, aus denen mit Hilfe externer Expertise neun Profillinien für die Universität beschlossen wurden. Der Steuerungsprozess habe eine positive Dynamik ausgelöst und den stärksten Effekt vielleicht in den Geisteswissenschaften gehabt. Demgegenüber würden individuelle Zielvereinbarungen inzwischen nur auf freiwilliger Basis abgeschlossen. LOM wiederum könne nicht genutzt werden, „da wir zu wenig zu verteilen haben“. Von Overhead-Geldern würde ein Drittel an den Wissenschaftler weitergereicht, zwei Drittel verblieben bei der Universitätsleitung. Einen gewissen finanziellen und strategischen Spielraum verschafften Gelder aus dem Europäischen Fonds für regionale Entwicklung. Die beste Steuerungswirkung hätten aus Sicht der Universität Berufungen. So sei man im Prozess, sich eng mit den drei in Leipzig ansässigen Max-Planck-Instituten zu vernetzen und habe dabei auch schon zu dem Instrument außerordentlicher Berufungen gegriffen. Und zusammen mit der Leipziger Musikhochschule sei man im Aufbau eines neuen, aus sechs Professuren bestehenden Instituts in der Musikwissenschaft. Schwarz‘ Fazit: „Die besten Steuerungsmöglichkeiten sind die sanften“.

Die Frage, welche Funktion Zahlen bei der Steuerung einer Hochschule haben, bestimmte auch die Diskussion zu diesem Themenkreis. Deren Begrenztheit war Konsens unter den anwesenden Rektor(inn)en und Präsident(inn)en. Zum einen war man sich bewusst, dass allein durch die Erhebung von Zahlen die Wirklichkeit schon verändert wird. Die Wissenschaft sei ein adaptives System und die Berücksichtigung der Indikatoren – das zeige beispielhaft das Slicing – führe automatisch zu einer Veränderung wissenschaftlichen Handelns. Hinzu komme, dass Indikatoren eher zur Bilanzierung der Vergangenheit geeignet seien als zur Erlangung sicherer Prognosen über die Zukunft. Kritisch wurde auch angemerkt, dass durch den Einsatz von Zahlen nur die Verteilung zugewiesener Mittel innerhalb der Universität gesteuert werden könne; absolute Leistungssteigerung einer Universität führe jedoch nicht zur Zuweisung höherer Mittel durch die Politik. Außerdem drohe die Ausrichtung an Kennzahlen die notwendige Reaktion auf neue Herausforderungen, auch auf Disruptionen, regelrecht zu verhindern. Als Beispiel wurde die Digitalisierung der Lehre genannt, die eigentlich neue Antworten von den Hochschulen wie beispielsweise kooperative Lehrkonzepte verlange. Im Forschungsbereich wiederum könne gerade die Profilbildung zur Fortschreibung des Status quo führen. Jede Universität müsse sich darum bemühen, Denkräume einzurichten – genannt wurden hier Entwicklungspläne für Fakultäten ebenso wie externe Beiräte –, damit sich nicht immer das Gleiche perpetuiere. Für kleine Fächer wiederum müssten andere Berechnungen zugrunde gelegt werden – ein Fach wie die Keltologie beispielsweise könne nicht nach der Anzahl der Studierenden vergütet werden, hier müsse auch die Bedeutung des Fachs für das gesamte deutsche Wissenschaftssystem eine Rolle spielen. Die eigene Rolle als Rektor und Präsident einer Hochschule beschrieben die Anwesenden als Mischung von „Kümmerer, Kommunikator und Impresario“.

BIRGITTA WOLFF (Frankfurt am Main) unterschied zwischen der leistungsorientierten Mittelzuweisung durch A) die Politik und B) innerhalb der Universität. A) Die Universität erhalte 57 % Grundfinanzierung vom Land. Insofern sei insbesondere die Möglichkeit, den 40-seitigen Kennzahlenbericht im Parlamentsausschuss selbst vorstellen zu dürfen, ein wichtiges Kommunikationsinstrument zwischen Politik und Universität in Hessen. Weitere externe Adressaten seien der Hochschulrat, die Stiftungskuratoren und die Freunde und Förderer der Universität. B) Innerhalb der Universität werde ausschließlich das Sachmittelbudget als leistungsorientierte Mittelzuweisung über die Fachbereiche berechnet. Parameter sind Drittmittelannahmen, Promotionen, Habilitationen, rechnerische

Studierende, Absolventen und die Frauenförderung, nicht aber Publikationen. Eine Ausnahme unter den 16 Fachbereichen stellt der Fachbereich Medizin dar, der 16 % der Mittel auf der Grundlage bibliometrischer Daten an Kliniken und Institute verteilt. Die individuelle Leistung einer Professorin bzw. eines Professors wird im Rahmen einer gesamtuniversitären Evaluation der W-Besoldung nach dem „Frankfurter Modell der W-Besoldung“ bewertet: Grundlage ist ein Selbstbericht, der von einer sogenannten „Elder-Statesmen-Kommission“ nach bewährten Kriterien, wie insbesondere Leistungen in Forschung und Lehre, ausgewertet und innerhalb der Fachcluster verglichen wird. Die Kommission spricht dem Präsidium eine Empfehlung aus, ob die festgelegten evaluationsabhängigen Leistungsbezüge für 3 Jahre weitergewährt werden oder bei einer Platzierung in der Spitzengruppe (Top 10 %) – bis zur nächsten Evaluation – verdoppelt werden.

Die nächsten beiden Impulsvorträge kamen aus medizinischen Hochschulen. CHRISTOPHER BAUM (MHH Hannover) kritisierte sowohl die Auswirkung von LOM als universitätsinternes Belohnungssystem wie auch die Diagnosis Related Groups (DRG), das Klassifikationssystem zur Bezahlung von medizinischen Leistungen in den Krankenhäusern. Beide Systeme führten zur Beschleunigung, zum Verlust der Kommunikation und zur Reduktion produktiver Auseinandersetzung mit Komplexität und Risiko. DRG bedeute die Bezahlung nach Fallpauschalen und stelle damit die Quantität („volume of care“) ins Zentrum, nicht die Qualität („value of care“). Die Folge sei eine systematische Umorientierung der Universitätsmedizin. Auch LOM fokussiere auf quantitative Parameter und nicht auf die Qualität. Nachhaltigkeit werde vernachlässigt, Mut zum Risiko nicht belohnt, Nachwuchsförderung stehe nicht im Zentrum. Baums Fazit: „Ich sehe die immanente Werteverchiebung kritisch“. Es müsse wieder mehr Zeit für die optimale Trias von Kommunikation, Reflexion und Experiment eingeräumt werden, die Grundausstattung habe daher eine hohe Bedeutung. Da die Notwendigkeit ökonomischen Handelns jedoch unabdingbar ist, müsse eine akademisch-ökonomische Kodominanz im Sinne einer gesunden Balance angestrebt werden.

In diese Kritik stimmte ULRICH DIRNAGL (Charité Berlin) ein und fokussierte sich vor allem auf die innermedizinischen Prozesse. Die Indikatorik, eigentlich als Steuerungsgröße gedacht, werde in der Universitätsmedizin mehr und mehr zum Selbstzweck. Dies führe dazu, dass die Indikatoren steigen, aber die Qualität falle<sup>12</sup>. Er zitierte: „Selection for high output leads to poorer methods and increasingly high false discovery rates“<sup>13</sup>. An der Charité würden etwa 10 Mio. EUR auf die individuellen Forscher(innen) verteilt, und zwar zur Hälfte auf Drittmittelinwerbung und zur anderen Hälfte für Publikationsleistungen via Journal Impact Factor. Dirnagl will dies verändern. Zum einen, indem die Mittel für Publikationen wiederum aufgeteilt würden, um die relative Zitationsrate des NIH zu berücksichtigen. Zum anderen durch die Einführung eines Sondertopfes für „Open Science“ in Höhe von (nur) 500.000 EUR. Die bisherigen Indikatoren im Bewerberportal wolle er durch neue Angaben ergänzen wie Open Data bei Publikationen, die Anzahl replizierter Studien oder auch Angaben zu Nullresultaten. Dirnagl wies darauf hin, dass es an den Institutionen sei, das System top down zu verbessern, der einzelne Mediziner hingegen habe sich in dem System eingerichtet. Softwareprogramme wie PIPredictor – „Predict your probability to become a

---

<sup>12</sup> Hinweis auf Glasziou, Paul/ Chalmers, Iain, „Avoidable waste in the production and reporting of research evidence“, in: *Lancet*, 2009 Jul 4;374 (9683); Baker, Monya „Is there a reproducibility crisis in science?“ in: *Nature* 25 May 2016/28 July 2016; Ionnidis, AA.

<sup>13</sup> Paul E. Smaldino, Richard McElreath „The natural selection of bad science“, Royal Society Open Science, 21 September 2016. DOI: 10.1098/rsos.160384.

Principal Investigator“ – führten vor Augen, dass es eine Fixierung auf die Publikation in High Impact Journalen in diesem System gebe, ohne Berücksichtigung des Inhalts, der Werthaltigkeit und der Robustheit der Befunde.

In der Diskussion waren sich die Teilnehmer(inn)en des Workshops einig, dass die Situation in der Medizin ein Sonderfall sei. Die zuständige Kommission beim Wissenschaftsrat bestehe beispielsweise ausschließlich aus Medizinerinnen, dabei sei der Blick von außen auf das Fach eigentlich dringend nötig. Eine wichtige Rolle könne hier auch die Forschungsförderung spielen, indem bei jedem Antrag abgefragt werde, wie die Daten zustande gekommen sind und was zu ihrer Reproduzierbarkeit getan worden ist. Dass die DFG in der Psychologie, in der sich vergleichbare Probleme auftun, Replikationsstudien nun endlich anerkenne und dass in Journals Bereiche für Replikationsstudien eingeräumt würden, seien wichtige Schritte in die richtige Richtung.

In einer Podiumsdiskussion zum Abschluss des forschungs- und hochschulpolitischen Gesprächs wurden zunächst die Unterschiede zwischen den einzelnen Bundesländern deutlich: In Mecklenburg-Vorpommern beispielsweise (ELEONORE WEBER, Greifswald) wurde LOM wieder abgeschafft, da es zu wenig Mittel zu verteilen gibt. In Thüringen (WALTER ROSENTHAL, Jena) habe man Ziel- und Leistungsvereinbarungen zugunsten strategischer Gespräche wieder abgeschafft. Nur 10 % würden hier nach fünf Indikatoren vergeben, was durchaus positive Wirkung habe. In Niedersachsen (MAY-BRITT KALLENRODE, Osnabrück) wiederum gebe es letztlich ein Benchmarkingsystem, mit dem eine Verknüpfung zwischen Leistung und Haushaltsmitteln hergestellt wird. Die „Benchmark“ ermittelt sich prozentual am Anteil der durch eine Universität eingebrachten formelrelevanten Haushaltsmittel an den insgesamt durch alle Hochschulen eingebrachten Haushaltsmitteln in einer Formelfächergruppe. Insgesamt waren sich die anwesenden Rektor(inn)en und Präsident(inn)en jedoch einig, dass die Hochschulen nun neu mit der Politik sprechen müssten. „Wir haben aus dem System rausgeholt, was man rausholen kann und die Steuerungsmöglichkeiten sind jetzt ausgereizt. Wir müssen Maßnahmen gegen die ‚Übersteuerungseffekte‘ vorbereiten.“ Dabei wurde der Wissenschaft ins Stammbuch geschrieben, dass sie bei den Gesprächen mit der Politik nicht Unterfinanzierung ins Feld führen dürfe, da Bildung und Wissenschaft schon jetzt in jedem Landeshaushalt die zweitgrößte Haushaltsposition sei. Außerdem sprächen die an den Hochschulen gebildeten Rücklagen der angeblichen Finanznot Hohn. Jetzt aber gehe es darum, Transparenz zu schaffen bzw. aufrechtzuerhalten durch die Entwicklung von qualitativen Parametern. „Wir müssen“, so das Fazit der Veranstaltung, „Qualität an die erste Stelle setzen“.