

IV.1.4

Leistungsmessung und -steuerung

Axel Oberschelp

Abstract | Mit Etablierung des New Public Management als Steuerungsparadigma hat die Messung von Leistungen im Hochschulbereich und in der Wissenschaft einen erheblichen Bedeutungszuwachs erfahren. Seitdem ist in nahezu allen Bereichen des Hochschul- und Wissenschaftssystems die Zunahme des Einsatzes metrifizierender Verfahren zur Leistungsmessung von Personen oder Organisationen zu beobachten. Von Kritiker*innen dieser Entwicklung wurden negative Effekte befürchtet, unter anderem die Einschränkung der Autonomie der Wissenschaft, die Orientierung von Forschungsthemen und -prozessen an Evaluationskriterien sowie Qualitätseinbußen infolge vorwiegend quantitativer Bewertungsmaßstäbe. Die mittlerweile weit verbreiteten Instrumente zur Leistungssteuerung wurden vor allem daraufhin untersucht, ob ihr Einsatz zu den intendierten Effekten, oder vielmehr zu Fehlsteuerungen führt.

Stichworte | Leistungsbegriff, Quantifizierung, Steuerungsinstrumente, Messen und Bewerten, Indikatoren

Einleitung

Der Bedeutungszuwachs leistungsmessender Verfahren im Hochschul- und Wissenschaftsbereich steht im Kontext der Reformen des New Public Management (NPM). Die Übertragung des Konzepts NPM auf den Hochschulbereich ging mit einer zunehmenden Orientierung an den Ergebnissen der Leistungserstellung (Output) als neuem Steuerungsprinzip und der Einführung neuer Steuerungsinstrumente einher, denen standardisierte Leistungsmessungen und -vergleiche zugrunde liegen. Metriken ermöglichen hierbei eine Quantifizierung der relevanten Sachverhalte und die Herstellung von Vergleichbarkeit (*commensuration*). Damit löste sich die Bewertung der Leistungsfähigkeit von Hochschulen von dem traditionellen Muster, wonach Hochschulen als Expertenorganisationen ausschließlich durch Expertenurteile bewertet werden können. Die mit diesem Wandel verbundenen Versprechen waren (und sind) die Herstellung von Transparenz bezüglich der Mittelverwendung, eine Verbesserung der Informationsgrundlage für Akteure in Politik und Gesellschaft, Effizienzgewinne sowie die Schaffung von Anreizen für Leistungssteigerungen durch Wettbewerb.

Parallel zu diesen Entwicklungen fand eine Intensivierung des wissenschaftlichen Diskurses zu Verfahren der Leistungsmessung und -steuerung statt. Während zugunsten des Einsatzes quantifizierender Verfahren die oben genannten Aspekte angeführt werden, weisen Kritiker*innen auf eine Vielzahl negativer Folgen hin (exemplarisch: Münch 2017; → IV.1.5 Academic Capitalism und Exzellenz). Dabei wird einerseits auf der methodischen Ebene argumentiert, wonach die Praktiken der Leistungsmessung mit der unzulässigen Vereinfachung einer hochkomplexen Wirklichkeit einher gingen und zu verzerrenden Darstellungen hinsichtlich der Leistungsfähigkeit von Einrichtungen führten.

Andererseits werden Folgerungen auf der systemischen Ebene, insbesondere die Überlagerung des wissenschaftlichen durch den ökonomischen Wettbewerb befürchtet.

Im Folgenden erläutert dieses Kapitel zunächst die häufig verwendeten Metriken in den Bereichen Lehre, Forschung und Third Mission. Anschließend werden die verbreitetsten Instrumente zur Steuerung von Leistungen behandelt und in einem Ausblick die Perspektiven künftiger Forschungsfelder skizziert. Zu einigen der hier verhandelten Themen sind keine neueren Arbeiten erschienen, weshalb die Darstellung wo erforderlich auf älteren Arbeiten basiert.

Metriken zur Messung und Bewertung hochschulischer Leistungen

Mit der Metrifizierung oder Quantifizierung ist im Wesentlichen eine Übersetzungsleistung verbunden, bei der „Phänomene, Eigenschaften oder Beschaffenheiten eines Sachverhalts [...] in einer allgemeinen, abstrakten und universell anschlussfähigen Sprache repräsentiert werden, der der Mathematik“ (Mau 2018: 27). Die Überführung der Eigenschaften von Sachverhalten oder Objekten in numerische Werte ist kein wertneutraler Vorgang, sondern Quantifizierungen sind immer mit Wertzuschreibungen verbunden (Heintz 2010). Die Vorteile numerischer Bewertungen liegen auf der Hand: Zahlen sind einfacher und schneller zu erfassen als Bewertungen in textlicher Form und bieten bessere Möglichkeiten der Vergleichbarkeit. Aber auch die Nachteile sind offenkundig: Ursache-Wirkung-Beziehungen sind nicht erkennbar und die Beeinflussbarkeit der Zahlen durch Rahmenbedingungen oder methodische Festlegungen ist nicht mehr nachvollziehbar.

Lehre

Eine Messung von Leistungen im Bereich Lehre steht vor zahlreichen Herausforderungen, die vor allem darin begründet sind, dass mehrere Akteure (Lehrende und Lernende) in die Leistungserstellung involviert sind. Zudem sind eine Vielzahl hochschulinterner und -externer Rahmenbedingungen von Relevanz, beispielsweise die Heterogenität von Fächern und Studienprogrammen sowie der Studierendenschaft. Zudem sind dynamische Entwicklungen in der Hochschulbildung zu berücksichtigen, wie eine zunehmende Wettbewerbsorientierung, Digitalisierung und Internationalisierung, sich ändernde Governance-Strukturen und demografische Rahmenbedingungen (Carlucci et al. 2019).

Ein Forschungsfeld befasst sich mit Fragen der *Qualitätsbestimmung* von Lehrleistungen und unterschiedlichen Qualitätsdimensionen (Schneijderberg/Steinhardt 2019; Pohlenz 2018). Dabei spielen divergierende, akteursbezogene Vorstellungen von Leistung und Qualität eine wichtige Rolle (Ullrich 2021).

Untersuchungen zu den Gütekriterien, und insbesondere zur Validität von Instrumenten zur Leistungsmessung in der Hochschullehre stellen ein weiteres zentrales Forschungsfeld dar. Hierbei stehen Ausgestaltung, Funktionen und Wirkungen *evaluativer Verfahren* im Mittelpunkt (Pohlenz 2018; Großmann/Wolbring 2016; Krempkow 2007). Diese Verfahren bemessen – zumeist auf der empirischen Grundlage von Studierendenbefragungen – die Studierbarkeit von Studienprogrammen (Seyfried et al. 2021), Ausstattungsmerkmale und Lehrformen, Kompetenzerwerb sowie Beschäftigungsbe-

fähigung und -adäquanz. Von herausgehobener Bedeutung ist zudem die Bestimmung von Studienerfolg und Studienabbruch, wobei deren Konzeptionalisierung als multidimensionale Konstrukte eine Voraussetzung darstellt (Seyfried et al. 2021; → II.3.3 Studienabbruchforschung).

Schließlich sind die von evaluativen Verfahren ausgehenden Wirkungen und Effekte Gegenstände der Forschung. Hier wurde festgestellt, dass leistungsmessende Verfahren zu einer Stärkung interner Hierarchien, zu mehr Detailsteuerung und zu einer Zunahme des Wettbewerbs führen (Seyfried/Pohlentz 2021). Andere Befunde zeigen hingegen, dass in der Lehre nach wie vor eine weitreichende akademische Autonomie besteht (Wilkesmann/Lauer 2020).

Nach der Einschätzung von Großmann und Wolbring (2016: 5) ist die Forschung über Evaluationen für die „Bewertung von Studium und Lehre von zentraler Bedeutung“. Kritik an der Genauigkeit und Verlässlichkeit der Messungen sollte deshalb ernst genommen und zur Entwicklung neuer bzw. modifizierter Verfahren genutzt werden. Diese könnten unter anderem experimentell ausgerichtete Wirkungsanalysen oder bessere Möglichkeiten zur Beobachtung individueller Lernprozesse beinhalten (Pohlentz 2018; → II.2.5 Lehrevaluation und Qualitätsentwicklung).

Forschung

Die verbreitetsten Metriken zu Messung von Forschungsleistungen sind Indikatoren zu eingeworbenen Drittmitteln (Input) sowie zu Publikationen (Output) und deren Wirkungen in Form von Zitationen in der *scientific community*. In beiden Fällen gehen mit den Urteilen von Peers qualitative Aspekte in die Leistungsgrößen ein: bei Publikationen durch das Peer Review der Fachzeitschriften, bei Drittmitteln im Rahmen der Antragsbegutachtung (→ III.1.6 Peer Review).

Im Kontext einer zunehmend drittmittelbasierten Forschung (Winterhager 2015: 33) hat die Bedeutung von *Drittmittelindikatoren* im Rahmen von Forschungsevaluationen zugenommen. Derartige Messungen auf der Grundlage gut verfügbarer Daten sind bspw. fester Bestandteil von Rankings und Verfahren der leistungsorientierten Mittelvergabe. Die Höhe eingeworbener Drittmittel stellt kein Maß für den Forschungserfolg, sondern zunächst lediglich eine Quantifizierung des Forschungsinputs dar, und es ist zumindest fraglich, ob Informationen zur Höhe eingeworbener Drittmittel geeignete Stellvertreterdaten für Forschungserfolg sind. Gleichwohl gibt es gute Gründe für eine Verwendung: Sofern ein Peer Review-Prozess Bestandteil des Antragsverfahrens war, enthalten Drittmittelindikatoren indirekt die Qualitätsurteile der Peers. Während für den Einsatz von Drittmittelindikatoren die Herstellung von Transparenz und (vermeintlich) leistungssteigernde Effekte sprechen, halten Kritiker*innen Fehlanreize entgegen, die zur Vereinheitlichung und Standardisierung von Forschungsprozessen führten (Münch 2006: 440). Die Auswirkungen des steigenden Drittmittelwettbewerbs auf Forschungsaktivitäten sind bislang jedoch kaum empirisch untersucht (Winterhager 2015).

Der *Bibliometrie* kommt eine hohe Praxisrelevanz in der Forschungssteuerung zu, indem häufig eine Kopplung von Messergebnissen an die Mittelallokation erfolgt: Bibliometrische Indikatoren sind Bestandteil von leistungsorientierten Mittelvergaben, von leistungsabhängigen Komponenten der Professor*innenbesoldung, von Evaluationen wissenschaftlicher Einrichtungen sowie internationalen Universitätsrankings und finden zunehmend im Kontext von Berufungsverfahren Verwendung.

Gegenstände bibliometrischer Analysen sind die Messung des Umfangs des Publikationsgeschehens sowie eine Bestimmung der hiermit verbundenen Wirkungen im Rahmen von Zitationsanalysen. Bibliometrische Forschung befasst sich traditionell intensiv mit ihren Methoden der Messung, der Validität von Indikatoren und den Wirkungen des Indikatoreneinsatzes. Beispielsweise wurde die Aussagekraft von Zitationen und deren Funktion in wissenschaftlichen Texten intensiv untersucht. Kritiker*innen des Indikatoreneinsatzes verweisen auf fachspezifische Zitationskulturen, auf die Größe und Zentralität von Publikationsmärkten sowie auf Netzwerk-Zitationen, die zu einer Verzerrung der Ergebnisse und zur Begünstigung einzelner Forschungsdesigns und Publikationsstrategien führen können (→ III.1.4 Fachkommunikation). Ein weiteres wichtiges Forschungsfeld sind intendierte und nicht-intendierte Effekte des Einsatzes bibliometrischer Indikatoren.

Die *Ausbildung des wissenschaftlichen Nachwuchses* sowie die *Bereitstellung von Forschungsinfrastruktur* stellen Bereiche der Leistungserstellung dar, in denen nicht unmittelbar neue Erkenntnisse generiert werden, die jedoch „entscheidend für die mittel- und langfristige Gewährung relevanter Forschungsergebnisse“ und für den „dauerhaften Bestand der wissenschaftlichen Forschung“ sind (Schmoch 2009: 28). Insbesondere die Bereitstellung von Forschungsinfrastruktur (z. B. Herausgeber-schaft von Zeitschriften, Bereitstellung von Forschungsdaten) wird derzeit jedoch eher selten als Messgröße verwendet (Koenig 2023: 93).

Third Mission

Der Begriff „Third Mission“ hat sich zur Beschreibung von Leistungen etabliert, die nicht (ausschließlich) den Kernaufgaben Lehre und Forschung zugerechnet werden können, und schließt an die Verwendung älterer, auf den (Technologie-)Transfer fokussierter Konzepte an. Ein wichtiges Thema der Forschung stellt die Entwicklung geeigneter *Abgrenzungskriterien* von „Third Mission“ gegenüber anderen Leistungsbereichen dar. Die Auffassung, dass „Third Mission“-Aktivitäten zwar erstens eine Kopplung zu Lehr- und Forschungsaktivitäten aufweisen sollten, jedoch nicht zu den Kernaufgaben von Lehre und Forschung zählen, und dass sie zweitens durch Interaktionen von Hochschulen mit der Gesellschaft charakterisiert seien, ist hierbei weit verbreitet (Henke et al. 2017: 66). Daraus ergeben sich drei zentrale inhaltliche Dimensionen: Technologie und Wissenstransfer (1; → III.2.2 Wissens- und Technologietransfer), Weiterbildung (2; → II.3.6. Wissenschaftliche Weiterbildung) sowie gesellschaftliches Engagement (3; → III.2.4 Soziale Innovationen für Transformationen des Wissens).

Eine Leistungsmessung im Bereich „Third Mission“ ist in vielfacher Hinsicht herausfordernd und deshalb in diesem Bereich weniger verbreitet als in den Kernbereichen Lehre und Forschung. Als praktikabel haben sich vor allem Kennzahlen in den Bereichen Patentierung, Ausgründungen und Kooperationsforschung erwiesen. Aktivitäten im Bereich Third Mission sind stark von den jeweiligen Wissensfeldern geprägt. So sind der forschungsbezogene Technologietransfer sowie Kooperationen zwischen Hochschulen und Unternehmen der Wirtschaft vor allem in den Natur- und Technikwissenschaften verbreitet (→ III.2.3 Hochschulen im regionalen Kontext). Da für die Geistes- und Sozialwissenschaften ein forschungsbezogener Wissenstransfer charakteristisch ist (Schneijderberg/Götze 2020: 31f.), begünstigen die derzeit gebräuchlichen Metriken zur Leistungsmessung besonders die MINT-Fächer.

Instrumente der Leistungssteuerung

In diesem Abschnitt werden Instrumente besprochen, bei denen der Einsatz von Indikatoren bzw. Kennzahlen erfolgt, um leistungssteigernde Effekte zu erzielen. Teilweise ist hierfür eine Kopplung der Leistungsmessung mit der Zuweisung monetärer Mittel implementiert. Ein zentrales Forschungsfeld zu allen im Folgenden behandelten Instrumenten sind intendierte als auch nicht-intendierte Wirkungen infolge der Messung (Reaktivität).

Leistungsorientierte Mittelvergabe

Verfahren der leistungsorientierten Mittelvergabe (kurz: LOM) haben seit Mitte der 1990er Jahre an deutschen Hochschulen weite Verbreitung gefunden. Das grundlegende Prinzip besteht in einer Kopplung der Leistungsmessung an die Hochschulfinanzierung. Indem die zu verteilenden Mittel nach quantitativen Leistungsparametern auf die Wettbewerber (z. B. die Hochschulen eines Bundeslandes) aufgeteilt werden, sollen Anreize zur Leistungssteigerung geschaffen werden (Jaeger 2009). LOM-Verfahren werden für die Landeshochschulsteuerung sowie für hochschulinterne Steuerungszwecke genutzt.

Wichtige Themen der Forschung sind die *Verbreitung von LOM-Verfahren* sowie deren inhaltliche Ausgestaltung. Dieses Untersuchungsfeld befindet sich im steten Wandel und ist auf eine Analyse teils schwer zugänglicher Dokumente der Landesministerien bzw. der Hochschulen angewiesen. Auf der Ebene Land-Hochschule, die im Fokus der Forschung steht, ist der Verbreitungsgrad rückläufig: Während die Untersuchung von König et al. (2012) eine aktive LOM noch für 15 der 16 Bundesländer feststellen konnte, war dies einige Jahre später nur für neun Bundesländer der Fall (vgl. Bogumil et al. 2015). Dies ist noch immer der aktuelle Stand zum Verbreitungsgrad auf dieser Ebene.¹ Verfahren, welche hochschulintern die Mittelverteilung zwischen den Fakultäten regeln, waren im Jahre 2014 an 84 % der Universitäten im Einsatz (ebd.). Weniger verbreitet sind LOM-Verfahren an HAW. Zum Einsatz kommen leicht zu messende und auf den Bereich Lehre fokussierte Kennzahlen, die vor allem im hochschulinternen Kontext um zusätzliche Messgrößen ergänzt werden.

Die vorliegenden Untersuchungen kommen zu dem Ergebnis, dass es auf allen Einsatzebenen, u. a. in Folge von begrenzten Verteilungswirkungen sowie Budget-Limitierungen, zu eher geringen Umverteilungseffekten kommt. Demzufolge resultieren aus dem Einsatz von LOM-Verfahren eher schwache *Anreizwirkungen* (Gläser/Laudel 2019; Bogumil et al. 2013). Durch die Kopplung von Leistungsmessung und Mittelallokation erweist sich die LOM in besonderer Weise als anfällig für Fehlsteuerungen und nicht-intendierte Effekte. Bewertungen des Instruments berücksichtigen zudem hohe Transaktionskosten, die aus dem Aufwand für Implementierung und operativen Betrieb resultieren sowie den Beitrag der LOM zur Transparenz der Mittelallokation.

1 Aktuelle Arbeiten hierzu liegen nicht vor. Die Einschätzung wird durch nicht-veröffentlichte eigene Recherchen gestützt.

Leistungsbezogene Besoldung

Die leistungsbezogene Besoldung des professoralen Personals im Rahmen der sog. W-Besoldung wurde im Zuge der Besoldungsreform des Jahres 2005 eingeführt. Indem die Reform eine Erhöhung der Grundgehälter durch variable Leistungsbezüge ermöglicht, sollen individuelle Leistungsanreize gesetzt werden.

Jochheim (2014) kommt in ihrer Untersuchung zu dem Ergebnis, dass ca. 80 % der W-besoldeten Professor*innen Leistungsbezüge erhalten. Neuere Untersuchungen zum *Verbreitungsgrad* liegen nicht vor. Daneben ist die konkrete Ausgestaltung der Vereinbarungen ein zentraler Forschungsgegenstand (Taferthofer 2017). Zu der Frage, ob die mit dem Instrumenteneinsatz intendierten *Effekte* tatsächlich realisiert werden, stellen mehrere Arbeiten übereinstimmend fest, dass von der leistungsbezogenen Besoldung lediglich geringe leistungssteigernde Anreize ausgehen (Jochheim 2014; Bogumil et al. 2013). Daneben sind die Akzeptanz des Instruments auf Seiten der Hochschulangehörigen sowie die mit seinem Einsatz verbundenen, hohen *Transaktionskosten* wichtige Themen der Forschung (Weißborn 2019). Einzelne neuere Untersuchungen bringen bundesland- und geschlechtsspezifische Aspekte in die Diskussion um die leistungsbezogene Besoldung ein (Kortendiek et al. 2021: 319).

Ziel- und Leistungsvereinbarungen

Ziel- und Leistungsvereinbarungen (ZLV) sind ein Instrument der partizipativen Führung und Organisationsentwicklung, bei dem Leistungsziele im Rahmen von Verhandlungen vereinbart werden. Mit der Kopplung finanzieller Zuweisungen an eine Zielerreichung sollen Anreize für Leistungssteigerungen gesetzt werden. ZLV enthalten in hohem Umfang auch quantitativ nicht operationalisierte Ziele, bspw. in Form von textlichen Zielabsprachen und Absichtserklärungen (Jaeger 2009). In den vergangenen Jahren sind nur wenige Studien zu ZLV erschienen sind, so dass sich der Kenntnisstand seit Beginn der 2010er Jahre kaum erweitert hat.

Im Hochschulbereich werden ZLV seit Ende der 1990er Jahre auf verschiedenen Ebenen eingesetzt. Viele Untersuchungen liegen zu bilateralen Vereinbarungen zwischen Bundesländern und ihren Hochschulen vor. Im Jahr 2014 hatten ca. 90 % der Universitäten Vereinbarungen mit dem jeweiligen Land abgeschlossen. Ebenfalls weit verbreitet sind hochschulinterne Vereinbarungen zwischen Hochschulleitung und Fakultäten. Bogumil et al. (2015) stellen für das Jahr 2014 fest, dass auf dieser Ebene an 68 % der Universitäten ZLV verwendet werden. Einer neueren Untersuchung zufolge, die auf 58 % *Verbreitungsgrad* kommt, war der Einsatz seitdem leicht rückläufig (Niggemann 2020).

Ein Schwerpunkt der Forschung ist die inhaltliche *Ausgestaltung* der Vereinbarungen. Neuere Untersuchungen zeigen hierbei beträchtliche Unterschiede hinsichtlich der Operationalisierung von Zielen und einer Kopplung mit Indikatoren (In der Smitten 2018). Für die Ebene Land-Hochschule wurden vor allem die Regelungsbereiche Lehre, Forschung, Organisationsentwicklung, Human Resources sowie Diversität und Weiterbildung als relevant identifiziert (In der Smitten/Jaeger 2012). Eine Kopplung von Zielerreichung und Mittelzuweisung (Budgetierungsfunktion) ist vor allem bei hochschulinternen ZLV gebräuchlich.

Bezüglich der intendierten Anreizwirkung von ZLV hat die Forschung überwiegend schwache bzw. unklare verhaltenssteuernde *Wirkungen* festgestellt. Allerdings variieren die Einschätzungen der beteiligten Akteure zur Wirksamkeit hochschulinterner ZLV erheblich. Zudem liegen Probleme bei der Zielüberprüfung sowie der Sanktionierung bei Nichterreichen der Ziele vor (Nickel 2007). In der Praxis überwiegt so häufig eine Legitimationsfunktion die erwünschte Anreizfunktion. Trotz hoher Transaktionskosten, die aus zeitaufwändigen Aushandlungsprozessen resultieren, ist eine breite *Akzeptanz* von ZLV zu konstatieren (Weißborn 2019). Diese resultiert nach Einschätzung von Jaeger (2009: 64) u. a. aus dem Umstand, dass ZLV an „Kommunikationsabläufe und Rollenverteilungen“ im Rahmen der akademischen Selbstverwaltung anknüpfen und vielfältige Partizipationsmöglichkeiten bieten.

Rankings

Seit Mitte der 1990er Jahre kommen in Deutschland vermehrt Hochschulrankings zum Einsatz. Bei einem Ranking handelt es sich um die Einordnung von Objekten in eine Rangliste zum Zwecke der Bewertung (Ruppert 2017). Anfangs sollten Rankings vor allem wissenschaftsexternen Adressaten (Politik, Unternehmen der Privatwirtschaft, Studierenden, Eltern) Informationen zum Leistungsvermögen von Hochschulen zur Verfügung stellen. Seitdem werden Rankings verstärkt für Leistungsvergleiche im nationalen und internationalen Rahmen sowie für Zwecke der Hochschulsteuerung genutzt (Espeland/Sauder 2016; Hazelnor 2016).

Neben Ansätzen zur Entwicklung einer soziologischen Theorie von Rankings (Brankovic et al. 2019) sind zwei zentrale Forschungsfelder von zentraler Bedeutung, die als Methoden-Diskurs und als Effekte-Diskurs bezeichnet werden:

Der *Methoden-Diskurs* hat die Relevanz und Konstruktion der von Rankings verwendeten Indikatoren zum Gegenstand. Klassische Rankings verdichten die Ergebnisse zu einer Zahl, bspw. zu einem Index für die gesamte Hochschule. Dem Nutzen der damit einhergehenden Komplexitätsreduktion stehen methodische Probleme gegenüber: Die Reduktion komplexer Organisationen auf eine einzige Zahl lässt die Ergebnisse von Rankings aus Sicht vieler Forschender zweifelhaft erscheinen. Ein weiterer Kritikpunkt hebt auf die Überzeichnung von Leistungsunterschieden ab, so würden geringfügige Differenzen bei den Kriterien durch die Rangplätze zu vermeintlichen Leistungsunterschieden überhöht.

Fragen nach dem Einfluss von Rankings auf die Messung und Bewertung hochschulischer Leistungen, nach den Möglichkeiten zur Beeinflussung von Ranking-Ergebnissen, nach der Nutzung von Rankings und nach einer Beeinflussung politischer Akteure lassen sich unter dem Begriff *Effekte-Diskurs* subsummieren. Arbeiten zu internationalen Universitätsrankings kommen zu dem Ergebnis, dass sich deren Einsatz formativ auf einzelne Hochschulen und auf das Hochschulsystem insgesamt auswirke (Hazelnor 2016). Die Folgen seien eine vertikale Differenzierung von Hochschulen und eine Forcierung wettbewerblichen Denkens auf allen institutionellen Ebenen. Auf organisationaler Ebene seien eine Beeinflussung interner Strukturen, Prozesse und Arbeitsaktivitäten von Hochschulen festzustellen, auf systemischer Ebene die Behinderung innovativer Forschungsvorhaben infolge der Orientierung an Ranking-Indikatoren.

Akkreditierungsverfahren

Akkreditierungsverfahren haben die Zertifizierung der Qualität von Studium und Lehre zum Ziel und werden von hochschulexternen Agenturen durchgeführt. Waren zunächst bereits laufende Studiengänge Gegenstand der Bewertung (Programmakkreditierung), finden mittlerweile verstärkt Bewertungen der Qualitätsmanagementsysteme von Hochschulen (Systemakkreditierung) statt.

Akkreditierungsverfahren sind weniger auf die Messung von Leistungen als vielmehr auf die Leistungssteuerung und die Sicherstellung von Mindeststandards der Lehr- und Studienorganisation gerichtet (Schneijderberg/Steinhardt 2019). Ein zentrales Thema der Forschung sind die Implikationen der Einführung von Akkreditierungsverfahren auf die *Governance von Hochschulen* und die damit verbundene Frage, inwieweit Mechanismen der externen Außensteuerung an Bedeutung zunehmen (Kaufmann 2012). Die Frage, ob Akkreditierungen einen Eingriff in die akademische Selbstverwaltung darstellen, oder ob sie durch ihren evaluativen Charakter in hohem Maße dem akademischen Selbstverständnis entsprechen (Steinhardt et al. 2018), ist dabei noch nicht eindeutig beantwortet.

Ausblick

Die Transformation von Hochschulen zu Akteuren in einem zunehmend global organisierten Wettbewerb wird sich in Zukunft vermutlich fortsetzen, womit die Erhebung, Auswertung und Nutzung von Daten für Entscheidungsprozesse weiter an Bedeutung gewinnt. Hinzu kommen neue Herausforderungen und Chancen im Zusammenhang mit der fortschreitenden Digitalisierung. Für die Hochschulforschung bedeutet dies, dass sich neue Themenfelder und Fragestellungen eröffnen, bspw. im Zusammenhang mit dem Einsatz von Künstlicher Intelligenz (KI), die sowohl als Untersuchungsgegenstand, der Prozesse und Praktiken in Forschung und Lehre verändert, als auch als Methodik der Hochschulforschung selbst immer stärker in den Fokus rückt.

Da in den vergangenen Jahren wenig neuere Studien zu einigen der hier besprochenen Themen erschienen sind, steht zum einen eine Überprüfung und Aktualisierung älterer Befunde aus. Zudem ist für etliche der hier skizzierten Forschungsfragen weiterer Klärungsbedarf zu konstatieren: Im Bereich der *Metriken* sind dies die Bestimmung der Aussagekraft von Indikatoren zur Drittmittelinwerbung sowie von Zitationsindices. Offene Fragen bestehen bei der Entwicklung praktikabler Messverfahren in den Bereichen Forschungsinfrastruktur und „Third Mission“. Wünschenswert wäre zudem eine stärkere empirische Fundierung des bislang eher theoretisch geführten Diskurses zu Auswirkungen des zunehmenden Drittmittel-Wettbewerbs auf Forschungsaktivitäten. Auf der Ebene des *Instrumenteneinsatzes* könnten die Auswirkungen der leistungsbezogenen Besoldung auf das Berufungsgeschehen näher bestimmt werden. Mit Blick auf ZLV sind sowohl der Einsatz als auch die Wirkungen quantitativer Messverfahren bislang noch unzureichend erforscht.

Möglicherweise kommt neueren theoretischen Konzepten zu Prozessen des Messens und Bewertens für die Bearbeitung dieser Forschungsfragen künftig eine größere Bedeutung zu. Als „Soziologie der Quantifizierung“ werden verschiedene Diskurse zusammengefasst, welche die „zunehmenden Vermessungs-, Vergleichs- und Evaluationspraktiken in der Gesellschaft“ zum Gegenstand haben (Ruppert 2017: 27, 189f.; Heintz 2010). Im Fokus steht eine Analyse von Bewertungskonstellationen, die es

ermöglicht, die Konsequenzen von Bewertungspraktiken für die wissenschaftliche Wissensproduktion besser zu erfassen. Inwieweit diese neuen Forschungszugänge unsere Kenntnisse über Verfahren und Effekte einer Leistungsmessung im Hochschulbereich erweitern können und welcher Methodeneinsatz hierbei erfolgversprechend ist, sind Fragen, die sich der Hochschulforschung für die Zukunft stellen.

Literaturempfehlungen

- Grande, Edgar/Jansen, Dorothea/Jarren, Otfried/Rip, Arie/Schimank, Uwe/Weingart, Peter (Hg.) (2013): Neue Governance der Wissenschaft. Reorganisation – externe Anforderungen – Medialisierung. Bielefeld: transcript. *In thematisch vielfältigen Beiträgen wird eine kompakte und instruktive Übersicht über die Reformen des New Public Management im Hochschulbereich geboten.*
- Heintz, Bettina (2010): Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. In: Zeitschrift für Soziologie 39, 162–181. *Der Aufsatz behandelt Vergleiche als grundlegende Sozialform und analysiert sie aus einer kommunikationstheoretischen Perspektive.*
- Krücken, Georg (2017): Die Transformation von Universitäten in Wettbewerbsakteure. In: Beiträge zur Hochschulforschung 39, 10–29. *Der Beitrag stellt die gegenwärtige Transformation von Universitäten in Wettbewerbsakteure dar und verortet diesen Prozess im Kontext allgemeiner Entwicklungen der Hochschulgovernance.*

Literaturverzeichnis

- Bogumil, Jörg/Burgi, Martin/Heinze, Rolf G./Gerber, Sascha/Gräf, Ilse-Dore/Jochheim, Linda/Schickentanz, Maren (2013): Zwischen Selbstverwaltungs- und Managementmodell. Umsetzungsstand und Bewertungen der neuen Steuerungsinstrumente in deutschen Universitäten. In: Grande, Edgar/Jansen, Dorothea/Jarren, Otfried et al. (Hg.), Neue Governance der Wissenschaft. Reorganisation – externe Anforderungen – Medialisierung. Bielefeld: transcript, 49–72.
- Bogumil, Jörg/Jochheim, Linda/Gerber, Sascha (2015): Universitäten zwischen Detail- und Kontextsteuerung. Wirkungen von Zielvereinbarungen und Finanzierungsformeln im Zeitvergleich. In: Bungarten, Pia/John-Ohnesorg, Marei (Hg.), Hochschulgovernance in Deutschland. Berlin: Friedrich-Ebert-Stiftung, 55–77.
- Brankovic, Jelena/Ringel, Leopold/Werron, Tobias (2019): Theorizing University Rankings. A Comparative Research Perspective. Working Paper SFB 1288; 2. Bielefeld: Universität Bielefeld, SFB 1288. DOI.org/10.4119/unibi/2939561 (05.07.2022).
- Carlucci, Daniela/Renna, Paolo/Izzo, Carmen/Schiuma, Giovanni (2019): Assessing teaching performance in higher education: a framework for continuous improvement. In: Management Decision 57(2), 461–479.
- Espeland, Wendy Nelson/Sauder, Michael (2016): Engines of Anxiety: Academic Rankings, Reputation, and Accountability. New York: Russel Sage Foundation.
- Gläser, Jochen/Laudel, Grit (2019): Die Unterkomplexität hochschulpolitischer Innovationen. In: Schubert, Cornelius/Schulz-Schaeffer, Ingo (Hg.), Berliner Schlüssel zur Techniksoziologie. Wiesbaden: Springer VS, 179–204.
- Großmann, Daniel/Wolbring, Tobias (Hg.) (2016): Evaluation von Studium und Lehre. Grundlagen, methodische Herausforderungen und Lösungsansätze. Wiesbaden: Springer VS.
- Hazekorn, Ellen (2016): Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence, 2. Aufl. Basingstoke: Palgrave Macmillan UK.

- Heintz, Bettina (2010): Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. In: *Zeitschrift für Soziologie* 39(3), 162–181.
- Henke, Justus/Pasternack, Peer/Schmid, Sarah (2017): *Mission, die dritte. Die Vielfalt jenseits hochschulischer Forschung und Lehre: Konzept und Kommunikation der Third Mission*. Berlin: Berliner Wissenschafts-Verlag.
- In der Smitten, Susanne (2018): *Evaluierende Einordnung der sächsischen Hochschulsteuerung im Ländervergleich mit einer besonderen Schwerpunktsetzung auf Zielvereinbarungen zwischen Ministerium und Hochschulen*. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Abschlussbericht. Hannover.
- In der Smitten, Susanne/Jaeger, Michael (2012): *Ziel- und Leistungsvereinbarungen als Instrument der Hochschulfinanzierung. Ausgestaltung und Anwendung*. Forum Hochschule, 16. Hannover: Hochschul-Informations-System GmbH.
- Jaeger, Michael (2009): *Steuerung durch Anreizsysteme an Hochschulen. Wie wirken formelgebundene Mittelverteilung und Zielvereinbarungen?* In: Bogumil, Jörg/Heinze, Rolf G. (Hg.), *Neue Steuerung von Hochschulen. Eine Zwischenbilanz*. Berlin: Edition Sigma, 45–66.
- Jochheim, Linda (2014): *Leistungsförderung durch Leistungszulagen? Effekte der W-Besoldung untersucht am Beispiel deutscher Universitätsprofessoren*. Dissertation Universität Bochum.
- Kaufmann, Benedict (2012): *Akkreditierung als Mikropolitik. Zur Wirkung neuer Steuerungsinstrumente an deutschen Hochschulen*. Zugl. Dissertation Universität Bremen, 2011. Wiesbaden: Springer VS.
- Koenig, Matthias (2023): *Zur Reform der Forschungsbewertung: Initiativen und Perspektiven*. In: *Beiträge zur Hochschulforschung* 45(2), 92–98.
- König, Karsten/Koglin, Gesa/Preische, Jens/Quaißer, Gunter (2012): *Transfer steuern – Eine Analyse wissenschaftspolitischer Instrumente in sechzehn Bundesländern*. HoF-Arbeitsbericht 3. Halle-Wittenberg.
- Kortendiek, Beate/Mense, Lisa/Beaufäys, Sandra/Bünnig, Jenny/Hendrix, Ulla/Herrmann, Jeremia/Mauer, Heike/Niegel, Jennifer (2021): *Gender Pay Gap und Geschlechter(un)gleichheit an Hochschulen*. Wiesbaden/Heidelberg: Springer VS.
- Krempkow, René (2007): *Leistungsbewertung, Leistungsanreize und die Qualität der Hochschullehre. Konzepte, Kriterien und ihre Akzeptanz*. Bielefeld: UVW – UniversitätsVerlagWeblar.
- Mau, Steffen (2018): *Das metrische Wir. Über die Quantifizierung des Sozialen*. Bonn: Bundeszentrale für politische Bildung.
- Münch, Richard (2006): *Drittmittel und Publikationen. Forschung zwischen Normalwissenschaft und Innovation*. In: *Soziologie* 35(4), 440–461.
- Münch, Richard (2017): *Die Macht der Zahlen in der Evaluation wissenschaftlicher Forschung. Eine soziologische Erklärung*. In: Hamann, Julian/Maeße, Jens/Gengnagel, Vincent et al. (Hg.), *Macht in Wissenschaft und Gesellschaft. Diskurs- und feldanalytische Perspektiven*. Wiesbaden: Springer VS, 215–228.
- Nickel, Sigrun (2007): *Partizipatives Management von Universitäten. Zielvereinbarungen, Leitungsstrukturen, staatliche Steuerung (Universität und Gesellschaft – Schriftenreihe zur Universitätsentwicklung, Bd. 5)*. München/Mering: Rainer Hampp.
- Niggemann, Felix (2020): *Interne LOM und ZLV als Instrumente der Universitätsleitung*. In: *Qualität in der Wissenschaft* 14(4), 94–98.
- Pohlentz, Philipp (2018): *Evaluation von Studium und Lehre an Hochschulen in Deutschland*. In: *Qualität in der Wissenschaft* 12(1), 10–14.
- Ruppert, Alfred (2017): *Vermessung und Quantifizierung im Hochschulsektor. Eine empirische Untersuchung über Effekte von Hochschulrankings*. Zugl. Dissertation Universität St. Gallen, 2016. Wiesbaden: Springer VS.
- Schmoch, Ulrich (2009): *Geeignete Ansätze zur Messung wissenschaftlicher Leistung*. In: *Beiträge zur Hochschulforschung* 31(1), 26–41.

- Schneijderberg, Christian/Götze, Nicolai (2020): Organisierte, metrifizierte und exzellente Wissenschaftler*innen. Veränderungen der Arbeits- und Beschäftigungsbedingungen an Fachhochschulen und Universitäten von 1992 über 2007 bis 2018. INCHER Working Paper Nr. 13. Kassel. DOI.org/10.5281/zenodo.3949756.
- Schneijderberg, Christian/Steinhardt, Isabel (2019): Accreditation of X Qualities Instead of Quality X: A Normative Analysis of Criteria of the German Higher Education Accreditation Regime. In: Higher Education Policy 32(1), 5–28. DOI.org/10.1057/s41307-018-0111-y.
- Seyfried, Markus/Hollenberg, Stefan/Baumgardt, Guido (2021): Konzeptionelle Betrachtungen zu divergierenden Konstruktionen des Studienerfolgs. In: Zeitschrift für Hochschulentwicklung 16(4), 25–44.
- Seyfried, Markus/Pohlenz, Philipp (2021): Institutionalisiertes Misstrauen und Leistungserbringung. In: Bohndick, Carla/Bülow-Schramm, Margret/Paul, Daria et al. (Hg.), Hochschullehre im Spannungsfeld zwischen individueller und institutioneller Verantwortung. Wiesbaden: Springer VS, 51–61.
- Steinhardt, Isabel/Schneijderberg, Christian/Krücken, Georg/Baumann, Janosch (2018): Externe und interne Qualitätssicherung von Studium und Lehre durch Akkreditierungs- und Evaluationsverfahren. INCHER Working Paper Nr. 9. Kassel. DOI.org/10.13140/RG.2.2.23858.71363.
- Tafertshofer, Lorenz (2017): Effekte hochschulinterner indikatorenbasierter Steuerungs- und Anreizverfahren auf das Forschungsverhalten von Professorinnen und Professoren an sozialwissenschaftlichen Fachbereichen. Dissertation. Tübingen: Universität Tübingen. DOI.org/10.15496/publikation-20970.
- Ullrich, Carsten G. (2021): Die Wahrnehmung und Deutung von Leistung und Leistungsprinzip bei Studierenden. In: Beiträge zur Hochschulforschung 43(3), 94–106.
- Weißborn, Leonie (2019): Eine qualitative Untersuchung der Transaktionskosten in Zielvereinbarungsprozessen zwischen Bundesländern und Universitäten. LCSS Working Papers, 2. Hannover: Institutionelles Repositorium der Leibniz Universität Hannover. DOI.org/10.15488/4438.
- Wilkesmann, Uwe/Lauer, Sabine (2020): The influence of teaching motivation and New Public Management on academic teaching. In: Studies in Higher Education 45(2), 434–451.
- Winterhager, Nicolas (2015): Drittmittelwettbewerb im universitären Forschungssektor. Zugl. Dissertation Universität Kassel, 2014. Wiesbaden: Springer VS.

