

Rechtliche Rahmenbedingungen von Evaluationen



Margrit Seckelmann

Evaluationen haben in den letzten Jahren als Elemente der Qualitätssicherung Eingang in die Landeshochschulgesetze gefunden. Diesen neuen Steuerungsinstrumenten wird im Zeichen gesteigerter Hochschulautonomie mit großen Erwartungen begegnet. Sie sollen nämlich u. a. als institutionelle Äquivalente als Instrumente der sogenannten „Output-Steuerung“ für eine Rücknahme der das Handeln von Wissenschaftseinrichtungen normierender Vorschriften (sogenannte „Input-Steuerung“) dienen. Im Hochschulbereich sollen überdies Qualität und gesellschaftliche Relevanz von Forschung und Lehre untersucht werden; auch darüber sollen Evaluationsverfahren Auskunft geben. Eine Gewinnung eindeutiger Qualitätszuschreibungen mithilfe von Evaluationen wird indes durch die Breite und Unschärfe des Evaluationsbegriffs und die Vielzahl der Verfahren zur Leistungsbeurteilung (extern und intern) erschwert. Daher ist eine Konkretisierung des Begriffs und der darunter zu verstehenden Verfahren und Indikatoren wie die Beleuchtung der (verfassungs-) rechtlichen Rahmenbedingungen geboten. Dieser Artikel will einen Beitrag dazu leisten.

Gliederung	Seite
1. Evaluationen als Element der Qualitätssicherung an Hochschulen	2
1.1 Evaluation – ein problematisches Unterfangen	2
1.2 Etablierung von Evaluationen an Hochschulen	2
1.3 Varianz der rechtlichen Rahmenbedingungen je nach Evaluationszweck und -form	3
2. Evaluationszwecke	4
2.1 Evaluationen als neue Legitimationsinstrumente	4
2.2 Evaluationen als Elemente neuer Steuerungsansätze	5
2.3 Evaluationen als neue Finanzsteuerungsinstrumente	5
3. Evaluationsarten und Evaluationsverfahren	6
3.1 Zum Begriff der Evaluation	6
3.2 Elementare Differenzierungen bei Evaluationen	7
3.3 Spezifische Differenzierungen	8
3.4 Typische Form der Evaluationsverfahren	9
4. Evaluationsmethoden, Evaluationsqualität und Evaluationseffekte	10
4.1 Negative Effekte	10
4.2 Eignung der Indikatoren?	10
5. Verfassungsrechtliche Rahmenbedingungen: Schutzgüter	13
5.1 Grundlagen	13
5.2 Organisationelle Schutzbereiche	14
5.3 Verfassungsrechtliche Zulässigkeit von Evaluationen: „Praktische Konkordanz“	14
6. Fazit	17

1. Evaluationen als Element der Qualitätssicherung an Hochschulen

1.1 Evaluation – ein problematisches Unterfangen

Evaluation als unscharfer Begriff

Evaluationen stellen ein bedeutendes Element der Qualitätssicherung an Hochschulen dar. Evaluationen und/oder andere Qualitätssicherungsmaßnahmen haben (in unterschiedlicher Detailschärfe) daher auch Eingang in die Landeshochschulgesetze gefunden. Stefan Hornbostel zufolge könnte man allerdings den Begriff der Evaluation „mit einigen Erfolgsaussichten für die Wahl der Unwörter des Jahres anmelden“, so sehr habe sich der Begriff in relativ kurzer Zeit zu einer „inhaltlich weitgehend unklaren catch-all-phrase“ entwickelt, die gerade durch ihre begriffliche Unschärfe hervorragend dazu geeignet sei, „sowohl fundamentale Kritik anzubringen wie eschatologische Hoffnungen und Visionen damit zu verbinden“ (Hornbostel 2010: 293). In der Tat wird die Berechtigung, aber auch die praktische Durchführung von Evaluationen als Element der Qualitätssicherung an Hochschulen derzeit intensiv diskutiert (vgl. u. a. Frey/Osterloh 2006; Knie/Simon 2008: 24; Walach 2009: 342; Hornbostel 2010: 293 sowie die Beiträge in Matthies/Simon 2008). Es ist in der Tat zu konstatieren, dass die Berechtigung von Evaluationen als Rahmenbedingungen „der Effektivität und Effizienz von Forschung und Lehre“ (Winter/Reil 2002: 7) ebenso umstritten ist wie der Evaluationsbegriff selbst, der sich kaum „trennscharf von Managementtechniken (incl. Qualitätsmanagement, Controlling, Auditing)“ abgrenzen lässt (Hornbostel 2010: 295). Unter dem Begriff der „Evaluation“ ist im aktuellen Sprachgebrauch nämlich weniger eine eindeutig fixierte Methodik oder Verfahrensweise als eine Art „Kunstlehre“ (Hornbostel 2010: 295) für die unterschiedlichsten Verfahren zu verstehen, die „zu einer Bewertung von Leistung führen“ (Müller-Böling 1996: 4), und auf deren Grundlage die eigentliche – politische – Entscheidung erfolgt.

1.2 Etablierung von Evaluationen an Hochschulen

Einführung verschiedener Bewertungsverfahren

Ausgehend von Empfehlungen des Wissenschaftsrats von 1985 zur Herstellung von Leistungstransparenz (Wissenschaftsrat 1985, dazu Knie/Simon 2010: 33) und der Bologna-Deklaration der europäischen Bildungsminister von 1999 (Röbbcke 2010: 335) haben sich Evaluationsverfahren als Bestandteil der Qualitätssicherung (zumindest im Bereich der Lehre) an Hochschulen etabliert (Kaufmann/Hornbostel 2008). Anknüpfend an Erfahrungen vor allem aus den Niederlanden sind zunächst der Verbund Norddeutscher Universitäten und die Zentrale Evaluationsagentur des Landes Niedersachsen, dann auch weitere Hochschulen, Hochschulverbünde und Agenturen dazu übergegangen, eine regelmäßige Evaluation von Studiengängen einzuführen; diese

treten zu den eingeführten Verfahren der Qualitätssicherung wie der Studierendenbefragung („Lehrevaluation“) und den Lehrberichten hinzu (Lütje 2010: 273). Teilweise werden Evaluationsverfahren in Form der Kombination von Selbstevaluation und externer Begutachtung in Akkreditierungsverfahren integriert, teilweise stehen Evaluation und Akkreditierung nebeneinander oder es kommt zu einer Art „Verdrängungswettbewerb“ von Evaluationen durch Akkreditierungen (vgl. Röbbcke 2010: 342; Lütje 2010: 273). Bei der Akkreditierung neuer Studiengänge wird auch das Qualitätssicherungssystem in die Begutachtung einbezogen. Die Kultusministerkonferenz hat am 14. Juni 2007 die Einführung sogenannter *Quality Audits* an Hochschulen beschlossen (318. Sitzung, einzusehen unter www.kmk.org). Diese können die erhobenen Daten zum Aufbau eines ‚Hochschulcontrolling‘ nutzen, in welchem Ziele und Schwerpunkte der Hochschulen definiert werden, Steuerungsgrößen und Indikatoren zur Messung der Zielerreichung festgelegt und Informationen über die Zielerreichung gewonnen werden, die ihrerseits evaluiert und bei der Aufstellung der neuen Ziele berücksichtigt werden (vgl. Vernau 2007: 2).

Verfahren der Evaluation bzw. der Qualitätssicherung finden ihre Erwähnung und zugleich ihre verfassungsrechtliche Absicherung in den Landeshochschulgesetzen (z. B. in § 5 des baden-württembergischen Landeshochschulgesetzes, Art. 10 des Bayerischen Hochschulgesetzes oder § 69 des Bremischen Hochschulgesetzes, in Berlin sind Evaluationen bzw. ist diese Form der Qualitätssicherung allerdings nur in den Bestimmungen über den Datenschutz im Hochschulgesetz erwähnt, vgl. § 6 Abs. 1 S. 1 Nr. 4 des Berliner Hochschulgesetzes [Berliner Hochschulgesetz vom 13.02.2003, GVBl. Berlin 59.2003,9, S. 82 ff., zul. geänd. durch Gesetz vom 19.03.2009 (GVBl. Berlin 65.2009,6, S. 70 ff.)]).

Gesetzliche Grundlagen

1.3 Varianz der rechtlichen Rahmenbedingungen je nach Evaluationszweck und -form

Die Vagheit des Evaluationsbegriffs macht eine rechtliche Bewertung und Umrahmung dieser Verfahren schwierig. Je nachdem, ob die erhobenen Daten veröffentlicht werden sollen, oder ob sie mit Konsequenzen finanzieller oder anderer Art verknüpft werden sollen, stellen sich unterschiedliche Anforderungen nicht nur an die rechtliche Bewertung, sondern auch schon an das Design von Evaluationen – von dem wiederum auch die rechtliche Bewertung abhängt. Von ausschlaggebender Bedeutung ist beispielsweise, ob die Evaluationen als internes Qualitätssicherungselement dienen sollen, oder ob diese etwa von einer externen Bewertungskommission im Auftrag eines Ministeriums oder einer Wissenschaftsorganisation durchgeführt werden. Aus diesem Grunde sollen nachfolgend die unterschiedlichen Evaluationszwecke und -formen vorgestellt werden, bevor dann auf die rechtlichen Rahmenbedingungen derselben einzugehen ist.

Evaluationsbewertung von Evaluationszweck abhängig

2. Evaluationszwecke

2.1 Evaluationen als neue Legitimationsinstrumente

Zusammenhang von Neuen Steuerungs- modellen und Autonomie mit Demokratieproblem

Evaluationen haben sich im Zeichen eines Neuen Steuerungsmodells von Hochschulen in den letzten Jahren etabliert (dazu u. a. Knie/Simon 2008). Sie stellen eine besondere Form der Qualitätssicherung dar, die einer gesteigerten Autonomie von Wissenschaftseinrichtungen korrespondiert. Diese Autonomie ist dadurch gekennzeichnet, dass inhaltliche Vorgaben hinsichtlich der *Art und Weise* der Mittelverwendung als Elemente direkter Steuerung ganz oder teilweise zurückgenommen wurden (beispielsweise durch die Herstellung weitgehender Finanzautonomie durch Umstellung auf Globalbudgets, teilweise auch gesteigerte Kompetenzen hinsichtlich der Gebäudebewirtschaftung, oder Zuwächse an Personalgewinnungs- oder -bewirtschaftungskompetenzen von Hochschulen, z. B. im Bereich von Berufungsverfahren), teilweise wurden auch die Genehmigungsvorbehalte im Bereich der Prüfungsordnungen zurückgenommen oder auf den Rektor bzw. Präsidenten der Hochschule delegiert (vgl. Stüber 2009: 99). Auf der anderen Seite verlangt das Demokratieprinzip, das aus den Artikeln 20 Abs. 1 und 2 GG abgeleitet wird, dass keine Bereiche im öffentlichen Sektor entstehen, die der Rückführbarkeit ihrer Entscheidungen auf den Volkswillen gänzlich entzogen sind. Dieser Volkswille wird im Parlament gebildet und durch den von diesem gewählten Regierungschef und die von ihm eingesetzten Minister umgesetzt. Nach dem Modell einer „Legitimationskette“ vom Volk zu den mit staatlichen Aufgaben betrauten Organen und Amtswaltern (vgl. BVerfGE 47, 253, 275; 52, 95, 130; 77, 1, 40; 83, 60, 72 f.; 93, 37, 66) verlangt das Bundesverfassungsgericht daher grundsätzlich ein Vorliegen ununterbrochener Zurechnungszusammenhänge der Entscheidungen nachgeordneter Bereiche an das Parlament (zu den Ausnahmen im Bereich funktionaler Selbstverwaltung vgl. BVerfGE 107, 59, neuerdings auch das Bundesverwaltungsgericht, BVerwG vom 27.06.2004, 10 B 72.04). Wenn dies nicht über die exekutivische Spitze geschieht, so muss nach dem neueren Modell alternativer Legitimationsverfahren das Parlament in den Stand gesetzt werden, über die entscheidenden Fragen, etwa bei Hochschulgesetzen und deren Novellen oder bei der Verabschiedung des Landeshaushalts zu entscheiden. Hierzu benötigt es Informationen, die ihm über Berichte und Evaluationen hinsichtlich der *Ergebnisse* der Mittelverwendung gegeben werden.

2.2 Evaluationen als Elemente neuer Steuerungsansätze

In der Literatur wird der Wechsel von der sogenannten „Input-Steuerung“ zur „Output“-Steuerung noch deutlicher akzentuiert (vgl. Schubert/Schmoch 2010: 245; Jansen 2010: 41). Gesteigerten Handlungsspielräumen von Wissenschaftseinrichtungen auf der „Input“-Seite stehen neue Darlegungspflichten über dem „Output“ der Wissenschaftseinrichtung gegenüber. Während bei der ersteren die Ordnungsmäßigkeit der Mittelverwendung im Mittelpunkt steht (durch die Vereinbarkeit mit der diese normierenden Vorschriften), sind bei der „Outputsteuerung“ die Effizienz und Effektivität die entscheidenden Kriterien unter dem Stichwort der „Ergebnistransparenz“ (Seckelmann 2009a) oder des „*value for money*“ (Jansen 2010: 42). Unter dem Gesichtspunkt des „*value for money*“ ist dem Modell nach entscheidend, dass den staatlichen Geldern, die in Wissenschaftseinrichtungen investiert werden, je nach Einrichtung möglichst qualitätsvolle Forschung und/oder Lehre gegenübersteht. Dabei versucht man, die Qualität der Forschung über im Wege von Eigenberichten erhobene quantitative Indikatoren und über Evaluationen (qualitative Bewertung – auch aufgrund der Indikatoren – durch Fachkollegen, sogenannte „*peer review*“, hierzu Neidhardt 2010: 280) zu ermitteln. Einen Eindruck von der Qualität der Lehre versucht man sich hingegen über sogenannte „*performance indicators*“ zu verschaffen, welche über Leistungskennzahlen Aspekte des sogenannten „Inputs“ (Studierendenzahlen), „Throughputs“ (Zwischenprüfungen, aber auch Studienabbrucherquote) und des „Outputs“ (Absolventenzahlen, Studiensemester, Promotionen, Publikationen) Auskunft geben können (vgl. Müller-Böling 1996: 3). Zusätzlich werden bzw. wurden qualitative Lehrberichte oder Selbstberichte der Hochschulen eingesetzt, die dann ggf. ebenfalls zusammen mit den „*performance indicators*“ qualitativ bewertet werden können (dazu Schenker-Wicki 1999).

**Evaluationen als Mittel
performanzbezogener
Output-Steuerung**

2.3 Evaluationen als neue Finanzsteuerungsinstrumente

In der Mehrheit der Bundesländer wird ein Teil der staatlichen Zuwendungen an die Hochschulen nach Leistungskriterien vergeben. Es lässt sich eine Tendenz zur Umwandlung von Teilen der (nicht-kompetitiv vergebenen) Grundfinanzierung und wettbewerblich vergebene Drittmittel ausmachen (nach Schubert/Schmoch 2010: 250). In den Zuwendungsbescheiden oder in den teilweise auch als „Hochschulverträgen“ bezeichneten Ziel- und Leistungsvereinbarungen werden die Hochschulen dazu berechtigt und verpflichtet, ihrerseits interne (leistungsbezogene) Steuerungselemente gegenüber ihren Fachbereichen einzuführen (Kracht 2006). Dem Modell nach sollen sich in den Hochschulen ähnliche Modelle der Qualitätssicherung und der

**Externe und interne
Mittelzuweisung auf
Grund von Evaluations-
ergebnissen**

„Belohnung“ forschungs- (oder lehr-) starker Fachbereiche etablieren, die ihrerseits im Rahmen der leistungsorientierten Besoldungsbestandteile (Leistungszulagen) Zielvereinbarungen mit den einzelnen Hochschullehrern treffen sollen. Hochschulen, Fachbereiche und Hochschullehrer sollen die kompetitive Erringungsmöglichkeit von Mitteln, die zu einer Grundausstattung hinzutreten, einerseits zu verbesserten Leistungen angehalten und andererseits zu einer größeren Ressourcensensibilität angeregt werden. So wird ein Teil der in den Leistungsvereinbarungen oder Hochschulverträgen festgeschriebenen Zuschüsse einbehalten und auf der Basis eines formelgebundenen Zuweisungsverfahrens zwischen den Hochschulen umverteilt. Hinsichtlich der der Berechnung zugrunde zu legenden Formeln und Parameter in den Aufgabenbereichen Lehre (wie etwa Erfolgsquote, Regelstudienzeitquote), Forschung (etwa über die Einwerbung kompetitiv vergebener Drittmittel) und Gleichstellung (etwa über die Anzahl der Berufung von Professorinnen) vgl. Jaeger/In der Smitten 2009: 2 (für Berlin), Seckelmann 2009, Schubert/Schmoch 2010: 250 sowie für die medizinischen Fakultäten und Universitätsklinika Neufeld 2009.

3. Evaluationsarten und Evaluationsverfahren

3.1 Zum Begriff der Evaluation



Definition

Unter Evaluation soll im Folgenden „die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes“ verstanden werden, wobei mögliche Gegenstände beispielsweise „Programme, Projekte, Produkte, Maßnahmen, Leistungen, Organisationen, Politik, Technologien oder Forschung“ sein können“. Systematisch meint dabei, dass die Empfehlungen nachvollziehbar sein und auf „empirisch gewonnenen qualitativen und/oder quantitativen Daten beruhen“ müssen (nach der Definition der Deutschen Gesellschaft für Evaluation e. V., DeGEval 2008: 15; Hornbostel 2010: 294).

Evaluation

Der Begriff der Evaluation ist indessen weit, da verschiedene Formen der Qualitätsbeurteilung von Forschung und Lehre (teilweise auch des wissenschaftsunterstützenden Bereichs) an Hochschulen darunter verstanden werden können. Daher soll dieser nachfolgend für die vorliegende Untersuchung präzisiert werden. Unter Evaluation (von lat. *valere*: wertschätzen) wird entweder der Vorgang (dann teilweise auch Evaluierung genannt) oder das Ergebnis einer Bewertung verstanden. Wie bereits erwähnt, lassen sich unter diesem „Sammelbegriff“ (Müller-Böling 1996: 4) oder „Oberbegriff“ (Hornbostel 2010: 295) eine Vielzahl von Verfahren und zu beurteilenden Gegenständen aus den „unterschiedlichsten Leistungszusammenhängen und Produktionsbereichen“

(Lange 2008: 12) verstehen, sofern diesen eine gewisse methodische Vorgehensweise zugrunde liegt, wie die nebenstehende Definition verdeutlicht.

Da sich die jeweiligen rechtlichen Rahmenbedingungen aber je nach Evaluationsgegenstand unterscheiden, ist eine gewisse Einengung des Begriffs geboten.

Der Begriff der „Evaluation“ soll in diesem Artikel nur auf Gegenstände bezogen werden, die einen engen Zusammenhang mit der Wissenschaft aufweisen, also von Lehre und Forschung. Wissenschaftsunterstützende Verwaltungstätigkeiten an Hochschulen und Forschungsinstituten, die teilweise auch Qualitätsbewertungsprozessen unterworfen werden (wie etwa im Verbund Norddeutscher Universitäten), sollen nachfolgend außer Betracht bleiben, da dort andere Kriterien und Verfahren der Qualitätsmessung und -bewertung zugrunde zu legen sind als bei der Beurteilung von Lehr- und Forschungsleistungen und diese Untersuchung den zur Verfügung stehenden Platz überschreiten würde.

Schon die letztgenannten Kriterien, also die der Beurteilung von Forschungs- oder Lehrleistungen, sind keinesfalls „kanonisiert“ und variieren stark zwischen den einzelnen Evaluationsgegenständen und -verfahren. Gemeinsam ist ihnen jedoch, dass für die Beurteilung von wissenschaftlichen Leistungen *wissenschaftsinterne* Kriterien der Qualität und Relevanz der Forschung bzw. Lehre wie der Effizienz oder Effektivität des dafür aufgewendeten Mitteleinsatzes ausschlaggebend sein müssen (BVerfGE 111, 333, 333; vgl. auch Knie/Simon 2008: 29; Seckelmann/Lauer/Jug 2006: 195; Hornbostel 2004: 173).

3.2 Elementare Differenzierungen bei Evaluationen

Evaluationen können unterschieden werden nach den zu beurteilenden Gegenständen, den Evaluationsobjekten oder -subjekten und den verwendeten Verfahren. Bezüglich der *Evaluationsgegenstände* sind Evaluationen von Forschung von denen der Lehre zu differenzieren. Hinsichtlich der *Evaluationsobjekte* sollen individuelle Leistungen von denen von Hochschulen, deren Fachbereichen und von außeruniversitären Forschungsinstituten unterschieden werden. Bezogen auf die *Evaluationssubjekte* lassen sich zum einen Beurteilungen durch Private (etwa in Foren wie „MeinProf.de“ oder durch das sogenannte „Spiegel-Ranking“) von solchen durch oder für Forschungseinrichtungen, die Hochschulen oder den Fachbereich selbst und schließlich durch oder für Ministerien (die Aufzählung ist nicht abschließend) trennen. Beurteilungen durch Privatpersonen sollen nachfolgend außer Betracht bleiben.

**Differenzierung der
Evaluationsansätze**

Bei den anderen Beurteilungen ist (unabhängig von der Rechtsform des Instituts oder der Universität) zwischen einer *internen* Selbstevaluation – durch den Fachbereich oder das Institut oder dessen Organe von einer *externen Evaluation* zu unterscheiden, die „von einem Organ erstellt wird, das außerhalb der Behörde liegt, die gesetzlich für das Objekt der Evaluation verantwortlich ist“ (Vedung 1999: 104).

**Interne und externe
Evaluation**

Das Beispiel Leibniz-Gemeinschaft

In der außeruniversitären Forschung ist das Beispiel der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz e. V. (Leibniz-Gemeinschaft) hervorzuheben. Diese kombiniert derartige Verfahren zeitlich: neben Evaluationen in siebenjährigen Abständen durch eine von der Leibniz-Gemeinschaft eingesetzte Bewertungskommission (externe Evaluation, hier in Form eines „*informed peer review*“) wird eine regelmäßige Zwischenbewertung durch den Wissenschaftlichen Beirat der Institute vorgenommen, die neben einer Selbstbewertung des Instituts auch in die abschließende Bewertung einfließen kann. Das Verfahren baut in Teilen auf demjenigen auf, das der Wissenschaftsrat bereits für die von Bund und Ländern gemeinschaftlich geförderten Institute der sogenannten „Blauen Liste“ entwickelt hatte. Wie dieser prüft die Leibniz-Gemeinschaft neben der wissenschaftlichen Qualität auch die überregionale Bedeutung und das allgemeine wissenschaftspolitische Interesse an der Einrichtung (zum – inzwischen geringfügig modifizierten – Verfahren vgl. Beisiegel/Hüttl/Klein/Müller 2007 bzw. Arnold/Groß 2005).

3.3 Spezifische Differenzierungen

Evaluationen lassen sich nicht nur nach den Akteuren aufteilen (interne und externe). Vielmehr lassen sich auch andere Unterscheidungen treffen.

Querschnitte und Systeme

Zunächst kann man nach dem *Evaluationsgegenstand Einzelbegutachtungen* von Hochschulen, Fachbereichen oder Instituten von sogenannten *Querschnittsevaluationen* unterscheiden, in denen bestimmte Fachgruppen evaluiert werden. Eine *Systemevaluation*, wie sie Ende der 1990er Jahre in Deutschland durchgeführt wurde, prüft demgegenüber grundsätzlich die Bewährung eines Systems, hier der Forschungsorganisationen.

Rating und Ranking

Ratings sind Einstufungen von Leitungen, die anhand von „*performance indicators*“ (dazu grundlegend Schenker-Wicki 1997) vorgenommen werden. Demgegenüber werden bei einem *Ranking* (Rangordnung) Leistungen von (Wissenschafts-)Einrichtungen etc. in eine auf- oder absteigende Reihenfolge gebracht. In der Regel baut ein Ranking auf einem Rating auf (vgl. auch R. Lange 2010).

Ein Ranking wird in der Regel (zumindest überwiegend) durch Fachkollegen („*peers*“) vorgenommen, wenn es durch oder für den Staat oder durch Wissenschaftsorganisationen vorgenommen wird (dazu Neidhardt 2010). Das hat zumindest in Hinblick auf den Staat seinen Grund auch in der Wissenschaftsfreiheitsgarantie nach Art. 5 Abs. 3 S. 1 GG, die eine gewisse „Staatsfreiheit“ der Wissenschaft bzw. eine „wertneutrale Wissenschaftspflege des Staates“ (BVerfGE 35, 79, 114) verlangt.

Der Begriff der Evaluation ist nicht auf eine Ergebniskontrolle begrenzt. Man unterscheidet *summative* und *formative Evaluationen* (vgl. u. a. Promberger/Bernhart/Niederkofler 2006: 47). Während die erstgenannten „statisch“ Ergebnisse oder Stände bewerten, tauschen sich bei den letztgenannten *formativen Evaluationen* (auch „ongoing-Evaluationen“) Evaluierende und Evaluierte über die Ergebnisse aus und suchen gemeinsam nach Verbesserungsmöglichkeiten. Man kann daher formative Evaluationen auch als eine Art „*Monitoring*“ bezeichnen. Ein solches Monitoring kann beispielsweise in Form von Lehrevaluationen geschehen.

**Summative und
formative Evaluation**

Evaluieren kann zudem nicht nur begleitend oder nachträglich („*ex post*“), sondern auch „*ex ante*“ – ein Verfahren, das in der Wissenschaft früher das übliche war, etwa bei Berufungen oder auch bei der Begutachtung von Förderanträgen durch die DFG, nunmehr auch bei der Exzellenzinitiative durch die DFG bzw. den Wissenschaftsrat. Das Monitoring kann dabei als eine „begleitende“ Evaluation angesehen werden.

**Ex-post und ex-ante
Evaluation**

3.4 Typische Form der Evaluationsverfahren

In den meisten Bundesländern hat sich bei den Hochschulen, sofern dort Fachbereiche evaluiert werden, ein zwei- bis dreistufiges Verfahren durchgesetzt: Auf eine interne Evaluation als Selbstbeschreibung und Selbstevaluation des Instituts, Fachs oder Fachbereichs, vor allem hinsichtlich quantitativer Parameter des „Forschungsausputs“ oder der Studierendenzahlen etc. folgt eine externe Evaluation durch ein nicht mit der Einrichtung verbundenes Expertengremium (*peer review* mit Vor-Ort-Begehung), die unter anderem auch aufgrund des vorgelegten Selbstberichts vorgenommen und durch eine Bewertung ergänzt wird. Diese wird mit einem schriftlichen Gutachten abgeschlossen, zu dem die betroffene Einrichtung in der Regel noch einmal Stellung nehmen kann (in der Gottfried-Wilhelm-Leibniz-Gemeinschaft e. V. nehmen ergänzend die Träger des Instituts zu aufgeworfenen Fragen, etwa hinsichtlich der Ressourcenausstattung, Stellung). Auf dieser Grundlage werden die entsprechenden ressourcenbezogenen Entscheidungen gefällt; in besonders partnerschaftlich organisierten Prozessen kann noch ein weiterer Verfahrensschritt zwischengeschaltet werden wie die sogenannte „auswertende Konferenz“ im Verbund Norddeutscher Universitäten (Kromrey 2005: 2). Im Universitätsbereich überlagert allerdings, wie unter 1.3 erwähnt, zunehmend die Akkreditierung die Evaluation, da die Hochschulen meist auf eine Akkreditierung angewiesen sind (Röbbecke 2010: 342; Lutje 2010: 273). Zugleich aber bewertet die Akkreditierung die Qualitätssicherungssysteme der Fachbereiche oder Hochschulen, so dass man auch davon sprechen kann, dass durch Akkreditierungen die Evaluationen verstetigt werden.

4. Evaluationsmethoden, Evaluationsqualität und Evaluationseffekte

Warum nun spricht Stefan Hornbostel aufgrund seiner langjährigen, intensiven Beschäftigung mit Evaluationen in seiner eingangs zitierten Bemerkung (Hornbostel 2010: 293) davon, man habe Aussicht auf Erfolg, wenn man das Wort „Evaluation“ zum „Unwort des Jahres“ vorschläge?

4.1 Negative Effekte

Fehlerhafte, nicht überzeugende Resultate

Hierzu tragen die Erfahrungen bei, die Evaluierte gemacht haben, die im Rahmen der Selbst- oder der Vorbereitung einer Fremdevaluation universitärer oder außeruniversitärer Wissenschaftseinrichtungen bibliometrische Daten liefern und/ oder aufbereiten mussten oder die im Rahmen von Lehrevaluationen (die eher eine Form der Vorbereitung einer Evaluation darstellen, vgl. Mutz/Daniel 2008: 34) lernen mussten, dass fordernde und anspruchsvolle Lehre oftmals schlechter bewertet wird als weniger ambitionierte.

Motivationsminderung

Ebenso können sich zu häufig durchgeführte Evaluationsverfahren motivationsmindernd auf die Wissenschaftler auswirken, denen dadurch weniger Zeit zum Forschen und Lehren bleibt (vgl. Frey/ Osterloh 2006). Um keine motivationsmindernden Ermüdungseffekte zu erzielen, sollten (externe) Evaluationen nur in größeren Zeitabständen durchgeführt werden.

4.2 Eignung der Indikatoren?

Zweifel an Indikatoreneignung

Es kommt hinzu, dass man auch bei den zu erhebenden Indikatoren deren quantitativer wie qualitativer Bewertung keinesfalls von „kanonisierten „Indikatoren oder Verfahren sprechen kann (vgl. u. a. Seckelmann/Lauer/Jug 2006: 195; Krawietz 2006: 23; Hornbostel 2010 sowie die Beiträge in Matthies/Simon 2008).

Die Diskussion macht sich vor allem an der Frage der Tauglichkeit der Verwendung bibliometrischer Daten, der Leistungsbeurteilung durch Fachkollegen (*peer review*) und der Aussagekraft von Drittmittelwerbungen fest. Weniger kontrovers werden demgegenüber in Hinblick auf die *Lehre/Ausbildung* die Messung von Studienverläufen bzw. akademischen Qualifikationsverläufen beurteilt, bei der Erstsemesterzahlen zu Abbrecherzahlen und Zahlen erfolgreicher Abschlüsse zueinander in Beziehung gesetzt bzw. für die späteren Qualifikationsphasen die erfolgreichen (teilweise auch gesondert die mit einer Prädikatsnote versehenen) Zahlen der Promotionen und die der Habilitationen in Relation zueinander gesetzt werden. Allerdings können

auch in diesem Bereich falsche Anreize durch zu wenig differenzierte Indikatoren gesetzt werden, z. B. durch eine „Belohnung“ möglichst vieler Promotionen mit Prädikatsnote (vgl. Kracht 2006).

Deutlich kontroverser werden die Verfahren zur Messung der Qualität und Relevanz von Forschung beurteilt. Zur Messung beider Größen werden namentlich qualitative Parameter (gewichteter) Publikationen, in der Regel auch eine Zitationsanalyse und die Zahl der eingeworbenen Drittmittel herangezogen, wobei den sogenannten „kompetitiv eingeworbenen“ Drittmitteln der EU-Forschungsförderungsprogramme, der Deutschen Forschungsgemeinschaft und einigen Stiftungen ein besonderer Stellenwert zukommt. Oftmals wird dann noch die internationale Vernetzung oder Attraktivität einer Hochschule bzw. eines Fachbereichs oder Instituts gemessen, beispielsweise anhand der sich dort aufhaltenden Stipendiaten der Alexander-von-Humboldt-Stiftung. Bei natur- oder ingenieurwissenschaftlich arbeitenden Fachbereichen oder Instituten werden die Zahlen der (erfolgreichen) Patentanträge herangezogen.

Jenseits der Messung der internationalen Attraktivität einer Einrichtung (bei der aber auch der jeweilige Standort zu beachten ist), sind diese Indikatoren keineswegs „neutral“, da ihnen eine Wertung vorge-schaltet ist.

Wertungs-Bias

Dies betrifft zum einen die *Zitationsanalyse* (ursprünglich von Garfield 1979), welche unter Zugrundelegung scientometrischer Methoden die Beachtung wissenschaftlicher Texte in der entsprechenden *community* aufgrund der Häufigkeit ihrer Zitationen misst. Dabei werden Zitationen im natur- bzw. ingenieurwissenschaftlichen und medizinischen Bereich über den *Science Citation Index* des *Web of Science* (www.isiknowledge.com) oder neuerdings auch mithilfe des sogenannten *Hirsch-Index* (dazu Näheres bei Hirsch 2005) ermittelt. Die Zitationen sozialwissenschaftlicher Texte lassen sich über den *Social Science Citation Index* (www.isiknowledge.com) ermitteln. Zitationsmessungen sind äußerst nützliche Hilfsmittel, sie können jedoch im Einzelfall ein etwas verzerrtes Bild von der wissenschaftlichen Wirklichkeit bieten: Es gibt zum einen Wissenschaften (beispielsweise die Rechtswissenschaft), die in diesen Zitationsindizes nur an ihren Randbereichen (etwa der Rechtssoziologie oder Rechtstheorie) vorkommen, nicht aber im eigentlichen (rechtsdogmatischen) „Kernbereich“. Zum anderen kann das reine Abstellen auf Zitationen falsche Anreize setzen. Kritiker rein bibliometrischer Verfahren beklagen eine Ermunterung zum „konservativen“ und wenig riskanten Forschen, das mit Fehlschlägen oder mangelnder Abdruckbarkeit verbunden sein kann, ein wissenschaftliches „Modebewusstsein“ und das Zerlegen innovativer Forschung in viele kleine Publikationen (sogenanntes „*salami slicing*“), die Bildung von Zitationskartellen, Eigenzitationen und eine Tendenz zum „Marketing“ von Forschung, die zeitlich zu Lasten der eigentlichen Forschung geht (zu allem Seckelmann/Lauer/Jug 2006:

Zitationsanalyse

199 mit weiteren Hinweisen sowie die Beiträge in Matthies/Simon 2008 und Hornbostel 2010).

Peer Review

Aber auch qualitative Beurteilungen durch die Fachkollegen („*peer review*“, dazu Neidhardt 2010) gestaltet sich nicht immer problemlos, denn diese setzt eine gewisse Distanz („mittlerer Reichweite“, vgl. Hornbostel 2010) voraus. In kleinen wissenschaftlichen „communities“ sind aber die Akteure durch zahlreiche Betreuungs- und Kooperationsverhältnisse miteinander vernetzt oder konkurrieren umgekehrt miteinander, was alles an sich begrüßenswert ist, aber die Neutralität bei der Begutachtung beeinträchtigen könnte.

Insbesondere bei der Kombination von internen und externen Evaluationen (etwa bei der Auditierung durch Beiräte) ist auf dem Hintergrund externer Evaluatoren, insbesondere von „peers“, auch davor zu warnen, dass die Kategorien zwischen „Beratung“ und „Entscheidung“ verschwimmen (dies betrifft auch die Frage der Hochschulräte deren Qualifikation als Organe der Hochschule oder als externe Beurteilungsgremien schwierig und je nach Bundesland anders zu beurteilen ist, vgl. Laqua 2004, Hütther 2009, St. Lange 2010; Seckelmann 2010.)

Drittmittel

Hinsichtlich der (kompetitiv) eingeworbenen *Drittmittel* ist auf das Problem hinzuweisen, dass Drittmittel selbst nur ein Indikator für die Relevanz von Forschungsprogrammen darstellen, nicht aber die Forschung selbst ersetzen können. Auch der Vergabe von Drittmitteln ist ein subjektives Element vorgeschaltet, diese werden zudem immer häufiger im Rahmen der EU-Forschungsförderungsprogramme, aber auch durch Stiftungen, aufgrund der Einfügbarekeit der Forschung in bestimmte *Programme* vergeben. Der Drittmittelindikator gibt daher nur über einen bestimmten *Typus* von Forschung Auskunft (dazu Seckelmann/Lauer/Jug 2006: 199 mit weiteren Hinweisen, zur Auswirkung des Drittmittelindikators auf den Forschungsausput vgl. Jansen/Wald/Franke/Schmoch/Schubert 2007). Zum anderen ist die Unterteilung zwischen „kompetitiv“ eingeworbenen und anderen Drittmitteln (etwa Projektförderungen durch Bundes- oder Landesministerien) problematisch, da die letztgenannten die *Relevanz* von Forschung, also ihre wirklichkeitsgestaltende Dimension, erfassen, ohne notwendigerweise weniger qualitativ sein zu müssen als die im Rahmen von Programmen von Forschungsförderungseinrichtungen geförderte Forschung.

In der Praxis werden daher qualitative Indikatoren und quantitative Verfahren miteinander kombiniert, um die „blinden Flecken“ der jeweiligen Methoden möglichst zu minimieren (so etwa bei Sondermann/Simon/Scholz/Hornbostel 2008). Man spricht von einer „Triangulation“, also einer Messung eines Punktes von zwei Punkten aus. Eine solche Methode stellt auch das „informed peer review“ dar, wie es beispielsweise durch den Wissenschaftsrat seinen Pilotstudien zum Forschungsrating zugrundegelegt wurde (Wissenschaftsrat 2008).

5. Verfassungsrechtliche Rahmenbedingungen: Schutzgüter

5.1 Grundlagen

Die Einführung von Evaluationsverfahren an Hochschulen (auf diese beschränken sich die nachfolgenden Ausführungen) geschieht keineswegs im „rechtsfreien Raum“. Ein Indiz für die Zulässigkeit von Evaluationsverfahren könnte es zwar grundsätzlich sein, dass Leistungsvergleiche im Bereich des Bildungswesens im Rahmen der ersten Stufe der Föderalismusreform in Art. 91b Abs. 2 4 Grundgesetz (GG) neuer Fassung und im Verwaltungsbereich (sogenanntes „Verwaltungs-Benchmarking“) gemäß dem neu eingeführten Art. 91d GG Eingang in das Grundgesetz fanden (Seckelmann 2009b). Dessen ungeachtet ist jegliche Evaluation aber aufgrund des Rechtsstaatsprinzips aus Art. 20 Abs. 3 GG an Gesetz und Recht gebunden, und gemäß Art. 1 Abs. 3 GG hat sie nach ihrer Konzeption und Durchführung die Grundrechte zu wahren.

**Rechtsstaatliche
Grenzen**

Als rechtliche Rahmenbedingungen für Evaluationen kommen für die Hochschulen und die an ihr in Forschung und Lehre Tätigen neben dem erwähnten Rechtsstaatsprinzip und Grundrechtsvorbehalt in Betracht das Demokratieprinzip aus Art. 20 Abs. 1 und 2 GG, die Wissenschaftsfreiheit gemäß Art. 5 Abs. 3 S. 1 GG in ihrer organisationsbezogenen Aussage (dazu sogleich und Seckelmann 2010), das Allgemeine Persönlichkeitsrecht der Evaluierten (Art. 2 Abs. 1 in Verbindung mit 1 Abs. 1 GG) bzw. das jenes konkretisierende Recht auf informationelle Selbstbestimmung sowie ihre Berufsausübungsfreiheit gemäß Art. 12 Abs. 1 GG, die aber im Wege der Grundrechtskonkurrenz ebenso hinter den organisationsbezogenen Aussagen der Wissenschaftsfreiheit zurücktreten dürfte wie – außer in besonders krassen Fällen – die Menschenwürde des Art. 1 Abs. 1 GG hinter das speziellere Allgemeine Persönlichkeitsrecht (das zumeist in den Bestimmungen über den Datenschutz näher konkretisiert wird). Bei den beamteten Hochschullehrern sind die hergebrachten Grundsätze des Berufsbeamtentums (Art. 33 Abs. 5 GG) zu wahren. Auch der Allgemeine Gleichheitsgrundsatz nach Art. 3 Abs. 1 GG kann im Einzelfall von Bedeutung sein. Schließlich könnte auch die Rechtsweggarantie nach Art. 19 Abs. 4 GG einschlägig sein. Bei den Evaluierenden sind hingegen die Meinungsfreiheit aus Art. 5 Abs. 1 GG sowie gewisse urheberrechtliche Aspekte zu beachten, die auf das Allgemeine Persönlichkeitsrecht nach Art. 2 Abs. 1 in Verbindung mit 1 Abs. 1 GG zurückzuführen sind. Schließlich spielt auch das Recht der Studierenden auf eine qualitätvolle Ausbildung aus Art. 12 Abs. 1 GG und deren „Studierfreiheit“ (dazu Glaser 2008) eine Rolle. Das Bundesstaatsprinzip bleibt nachfolgend außer Betracht, da dies in erster Linie die gemeinschaftlich geförderten außeruniversitären Forschungseinrichtungen betrifft.

**Elementare
verfassungsrechtliche
Schutzbereiche**

5.2 Organisationelle Schutzbereiche

Demokratieprinzip

Hinsichtlich des Demokratieprinzips gilt das unter 2.1 ausgeführte: Das Bundesverfassungsgericht verlangt in seiner ständigen Rechtsprechung (vgl. BVerfGE 47, 253, 275; 52, 95, 130; 77, 1, 40; 83, 60, 72 f.; 93, 37, 66) ein Vorliegen ununterbrochener Zurechnungszusammenhänge der Entscheidungen nachgeordneter Bereiche an das Parlament, sofern nicht ein spezieller Bereich funktionaler Selbstverwaltung vorliegt (vgl. BVerfGE 107, 59; ebenso BVerwG vom 27.06.2004, 10 B 72.04). Bei universitätsinternen Evaluationen erscheint dies unter diesem Gesichtspunkt als weniger problematisch, da diese wohl dem universitären Selbstverwaltungsbereich zuzurechnen sind. Schwieriger wird es bei externen Evaluationen, denn das Demokratie- wie das Rechtsstaatsprinzip verlangen eine klare Zuordnung von Handlungen an bestimmte Staatsgewalten. Besonders problematisch wird dies regelmäßig dann, wenn eine Beratung eine Entscheidung quasi „vorwegnimmt“. Problematisch kann eine solche Vorwegnahme auch in Hinblick auf die Rechtsweggarantie des Art. 19 Abs. 4 GG erscheinen, hier ist aber die Klage gegen die Körperschaft zu richten, die die fragliche Entscheidung aufgrund des ihr vorgelegten Evaluationsberichts trifft (schwieriger ist diese Frage bei der Akkreditierung zu beurteilen, vgl. Stüber 2009).

Evaluationsgesteuerte Mittelverteilung und Grundausrüstungs-garantie

Besonders relevant für die rechtliche Beurteilung von Evaluationen sind die organisationsbezogenen Aussagen der Wissenschaftsfreiheit. Das Bundesverfassungsgericht hat in ständiger Rechtsprechung auch eine zur Wahrnehmung der grundrechtlich garantierten Forschungs- und Lehrfreiheit der Hochschullehrer angemessene Grundausrüstung von der positiven Grundrechtsaussage der Wissenschaftsfreiheit nach Art. 5 Abs. 3 S. 1 GG für umfasst erklärt (BVerfGE 35, 79, 113; 90, 1, 11f.; 111, 333, 353; ähnlich das Bundesverwaltungsgericht, BVerwGE 102, 304, 307; vgl. auch Seckelmann 2009a, Seckelmann 2010).

5.3 Verfassungsrechtliche Zulässigkeit von Evaluationen: „Praktische Konkordanz“

Grundlagen

Das Bundesverfassungsgericht hat im Jahre 2004 die Verfassungskonformität eines Hochschulgesetzes festgestellt, dessen § 7 die Durchführung sogenannter Lehrevaluationen vorsah (anders wohl Hufen 1995; gegen die Verfassungskonformität von Evaluationen v. a. Mußnug 1992, zu den Folgerungen aus der Entscheidung für eine Neubestimmung der Wissenschaftsfreiheit u. a. Ladeur 2005: 764, nunmehr auch grundlegend Gärditz 2009 mit weiteren Nennungen). Denn im Bereich der Lehre an Hochschulen bewegen sich diese Fragen im Spannungsverhältnis zwischen der Forschungs- und Lehrfreiheit der Lehrenden (aus Art. 5 Abs. 3 S. 1 GG) sowie ggf. anderer Grundrechte und grundrechtsähnlichen Rechte, die etwa aus ihrer Beamtenstellung resultieren können (Art. 33 Abs. 5 GG), und des

durch Art. 12 Abs. 1 GG garantierten Anspruchs der Studierenden auf eine gute Ausbildung und ggf. ihrer eigenen „Studierfreiheit“ (zur genauen dogmatischen Herleitung vgl. Glaser 2008). Denn trotz der vorbehaltlosen Gewährung des Grundrechts der Wissenschaftsfreiheit (Satz 2 des Art. 5 Abs. 3 GG wird gemeinhin nicht als eine sogenannte Grundrechtsschranke, sondern als die Klarstellung einer Selbstverständlichkeit betrachtet), muss sich das Grundrecht der Wissenschaftler auf eine freie Lehre mit den erwähnten Rechten der Studierenden auf eine gute Ausbildung durch den Gesetzgeber (!) in einen Ausgleich bringen lassen.

Für diese sogenannte „praktische Konkordanz“ hat das Bundesverfassungsgericht bestimmte Rahmenanforderungen aufgestellt: Es hat eine Vereinbarkeit der Verknüpfung der Bewertung wissenschaftlicher Qualität mit Folgen bei der Mittelverteilung mit der Wissenschaftsfreiheit gemäß Art. 5 Abs. 3 S. 1 GG nur unter der Bedingung bejaht, dass diese aufgrund eines auch hinsichtlich des Verfahrens und der Folgen hinreichend bestimmten Gesetzes erfolge (hinsichtlich des mit der Erhebung und Speicherung bzw. Verarbeitung der im Rahmen der Evaluationen erhobenen Daten erfolgt dieses auch schon aus datenschutzrechtlichen Gründen). Des Weiteren fordert es für die Verfassungskonformität eines solchen Gesetzes, dass „eine wissenschaftsadäquate Bewertung der Leistung hinreichend gewährleistet ist“ (BVerfGE 111, 333, 333). Um die Wissenschaft nicht der Gefahr einer externen Einflussnahme oder gar Lenkung auszusetzen, müssten die für die Evaluation benutzten Bewertungskriterien „hinreichenden Raum für wissenschaftseigene Orientierungen“ (BVerfGE 111, 333, 358) lassen. Das gelte übrigens unabhängig davon, ob solche Kriterien hochschulextern (also z. B. durch ein Ministerium oder im Zusammenwirken mit diesem) oder hochschulintern festgesetzt würden. Beachtlich ist in diesem Zusammenhang der Hinweis des Bundesverfassungsgerichts darauf, dass einer externen Festsetzung „ein erhöhtes Risiko der Vernachlässigung wissenschaftsadäquater Belange“ innewohne, etwa „durch Nutzung der Evaluation zur Erreichung wissenschaftsfremder Zwecke“ (BVerfGE 111, 333, 358).

Es ist in der Qualitätssicherungsliteratur umstritten, ob die Kriterien zur Evaluation von Hochschulen von diesen selbst aufgestellt werden (durch ein Messen an den eigenen Zielen und mithilfe der dafür angebotenen Evaluationskriterien, sogenannter „fitness for“ bzw. „of purpose“-Ansatz) oder ob Hochschulen an externen Kriterien, einer Art Benchmarking, gemessen werden sollen (sogenannter „compliance“-Ansatz). Beide Ansätze haben ihre Vorteile, jedoch ist der compliance-Ansatz auch mit Problemen verbunden, die zum einen die Hochschulautonomie und zum anderen die Motivation der zu Evaluierenden betreffen. Innovationsbereitschaft wird eher dann gefördert, wenn diejenigen, die die Qualitätssicherung betrifft (Hochschulen, ggf. auch Fachbereiche, ähnliches gilt für Forschungsinstitute) ihr Leitbild und ihre Ziele selbst aufstellen und die Maßstäbe, anhand deren die Zieler-

**Konkrete Erfordernisse
verfassungskonformer
Evaluation**

**Qualitätsmaßstab:
fitness of/for purpose
oder compliance?**

reichung überprüft werden kann, zumindest mitbeeinflussen können. Das Bundesverfassungsgericht hat diese Frage in seiner Entscheidung im 111. Band offen gelassen, seine Ausführungen deuten aber darauf hin, dass unter dem Gesichtspunkt der organisationalen Aussagen des Art. 5 Abs. 3 S. 1 GG der „fitness of“- bzw. „fitness for purpose“-Ansatz weniger verfassungsrechtliche Bedenken aufwirft als der (eher heteronome und mit der Gefahr einer gewissen „Versteinerung“ versehene) „compliance“-Ansatz.

Anwendung wissenschaftsadäquater Evaluationskriterien und -verfahren

Evaluationskriterien müssen daher mit besonderer Sorgfalt und nach Möglichkeit in Zusammenwirken von Ministerien und Hochschulen bzw. Hochschulen und Fachbereichen ermittelt werden, um die vom Bundesverfassungsgericht benannten „wissenschaftsadäquaten Kategorien“ zu berücksichtigen. Deren Berücksichtigung sei, so das Bundesverfassungsgericht, insbesondere dann nötig, „wenn die Verteilung öffentlicher Mittel an die Evaluationsergebnisse geknüpft“ werde, da „die Hochschulangehörigen auf den öffentlichen Wissenschaftsbetrieb und dessen Ressourcen angewiesen“ seien (BVerfGE 111, 333, 358). Mithin setzt die Sicherstellung „wissenschaftsadäquater Evaluationskriterien“ nach Ansicht des Bundesverfassungsgerichts auch eine Beteiligung von Wissenschaftlern bei deren Aufstellung in gewissem Umfang voraus (BVerfGE 111, 333, 359).

Gesetzesvorbehalt – Grund, Grenzen und Umfang

Evaluationen setzen ferner ein Gesetz voraus. Die Kriterien für die Evaluation und das Verfahren sowie die zugrunde zu legenden Indikatoren müssen im Gesetz nicht eigens bestimmt sein, dies birgt wohl auch die Gefahr einer Überregulierung. Gleichwohl ist auch im Hinblick auf den Allgemeinen Gleichheitssatz des Art. 3 Abs. 1 GG eine gewisse Vergleichbarkeit gleich behandelter Wissenschaftseinrichtungen zu fordern, es besteht insoweit ein Willkürverbot, das wiederum für eine gewisse Vergleichbarkeit der Indikatoren und Verfahren spräche. Wie weit diese im Einzelfall reichen soll, ist schwierig zu beurteilen. Der Gesetzgeber kann im Rahmen seines Einschätzungs- und Prognosespielraums nach Ansicht des Bundesverfassungsgerichts ein Modell etablieren, in dem die Herausarbeitung von solchen Kriterien einem inneruniversitären Prozess überlassen bleibt, sofern dort keine strukturelle Gefahr für die Möglichkeit besteht, Wissenschaft frei auszuüben (BVerfGE 111, 333, 360). Den Gesetzgeber trifft aber insoweit eine Beobachtungs- und gegebenenfalls eine Nachbesserungspflicht (vgl. auch BVerfGE 95, 267, 314 m. w. N.), sobald Gefahren für die Wissenschaftsfreiheit, zum Beispiel durch die Verwendung wissenschaftsinadäquater Kriterien, auftreten (BVerfGE 111, 333, 360). Eine solche Beobachtungspflicht dürfte auch für die mittelbare Erwähnung von Evaluationen in § 6 Abs. 1 S. 1 Nr. 4 des Berliner Hochschulgesetzes einschlägig sein, wünschenswert wäre eine weitere Konkretisierung.

**Allgemeines
Persönlichkeitsrecht –
Verhältnismäßigkeits-
grundsatz**

Das Allgemeine Persönlichkeitsrecht der Evaluierten (und die es konkretisierenden datenschutzrechtlichen Bestimmungen in den Landeshochschul- bzw. Landesdatenschutzgesetzen) zwingen zu einer möglichst anonymisierten Darstellung, zumal auch mit einem hochschulöffentlichem „*naming and shaming*“, wie es die OECD in ihren PISA-Leistungsvergleichen vornimmt, den betroffenen Lehrenden nicht geholfen ist, diese sollen besser zu Beratungsgesprächen eingeladen werden. Der Bundesgerichtshof hat in einem viel zitierten Urteil zu einem privaten (!) Bewertungsportal namens „spickmich“ (BGH, Urteil v. 23.06.2009, Az. VI ZR 196/0) der Meinungsfreiheit aus Art. 5 Abs. 1 GG der Bewertenden grundsätzlich einen Vorzug gegenüber dem Allgemeinen Persönlichkeitsrecht der Bewerteten gegeben, sofern es um Diskussionen in einer Art ‚Betroffenen-communities‘ geht; eine verfassungsgerichtliche Entscheidung hierzu steht noch aus. Indes ist zu beachten, dass die Verwaltung gemäß Art. 20 Abs. 3 und Art. 1 Abs. 3 GG sowie den Dienstherrn nach Art. 33 Abs. 5 GG in Verbindung mit den Bestimmungen der jeweiligen Landesverfassungen und Landesbeamtengesetze eine Fürsorgepflicht trifft, die strengere Maßstäbe an deren Kundgaben anlegt, als dies für die Meinungsäußerung im privaten Verkehr gilt. Schwierig wird es aber, wenn Dritte aufgrund der neuen Landesinformationsgesetze Akteneinsicht begehren; hier wird wegen der personenbezogenen Daten indes ein Ausnahmetatbestand einschlägig sein.

Das Allgemeine Persönlichkeitsrecht der Evaluierenden schließlich umfasst gewisse urheberrechtliche Aspekte. Insbesondere darf deren Name nicht in Verbindung mit verkürzten oder veränderten Evaluationsberichten in Verbindung gebracht werden, diese Berichte müssen von allen Evaluierenden genehmigt oder den Evaluierenden die Möglichkeit zu einem Sondervotum gewährt werden.

6. Fazit

Evaluationen sind nützliche Instrumente für das Wissensmanagement von Forschenden und Lehrenden wie von Fachbereichen, Hochschulen und Instituten. Sie vollziehen sich indes nicht im rechtsfreien Raum, sondern unter bestimmten rechtlichen Rahmenbedingungen (dazu auch Seckelmann 2009a; Seckelmann 2010), zu denen insbesondere das Demokratieprinzip, die organisationsbezogenen Aussagen der Wissenschaftsfreiheit und das Allgemeine Persönlichkeitsrecht der Evaluierten wie der Allgemeine Gleichheitssatz zu rechnen sind. Wegen der Grundrechtsrelevanz wiegt – gerade unter Hinblick auf den Allgemeinen Gleichheitssatz – die noch unabgeschlossene Diskussion über die Verfahren und Indikatoren zur Leistungsbeurteilung in Forschung und Lehre besonders schwer. Verfassungsrechtlich geboten ist vor allem der Umstand, dass die Evaluationen im Hochschulbereich

auf ein Gesetz zurückzuführen sind und dass die zu wählenden Kategorien *solche der Wissenschaft selbst sind*, die zwischen den Ministerien und den Hochschulen, den Hochschulleitungen und den Fachbereichen (BVerfGE 111, 333, 333) oder zwischen den Fachgemeinschaften und den Wissenschaftsorganisationen ausgehandelt werden können (Knie/Simon 2008: 29; vgl. auch Seckelmann/Lauer/Jug 2006: 201f.). Eine gewisse Standardisierung der Parameter (z. B. durch eine Verpflichtung auf die Grundsätze der Deutschen Gesellschaft für Evaluation, <http://www.degeval.de/>) ist durchaus wünschenswert, ebenso wie die notwendige Flexibilität der Indikatoren, um dem jeweiligen Fach oder der jeweiligen Wissenschaftseinrichtung gerecht zu werden.

Um an die eingangs erwähnte Bemerkung von Stefan Hornbostel anzuknüpfen: Es geht bei Evaluationen nicht um das „Unwort des Jahres“, aber es lässt sich durchaus an einen Satz von Paracelsus denken: Die Dosis und die Anwendung unterscheiden die Medizin vom Gift. Maßvoll und im Sinne der Hochschulautonomie durchgeführte Evaluationen („fitness for/of purpose“-Ansatz) können sich belebend auf die Innovationskraft von Hochschulen auswirken, „Überdosierungen“ und heteronome Kriterien sowie unsachgemäße Anwendungen können dagegen auch Lähmungserscheinungen hervorrufen.

Literatur

Zugriffsdatum für alle elektronischen Quellen: 30.11.2009.

- [1] Arnold, Nathalie/Groß, Thomas (2005): Die Entscheidungsstrukturen der Leibniz-Gemeinschaft, *Wissenschaftsrecht* 38, S. 238-263.
- [2] Beisiegel, Ulrike/Hüttl, Reinhard F./Klein, Carsten/Müller, Bernhard (2007), Fortschritt durch Evaluierung. Die Leibniz-Gemeinschaft hat ihr Evaluierungsverfahren weiterentwickelt. In: *wissenschaftsmanagement*, S. 28-34.
- [3] DeGEVal (2008) (Hrsg.): Standards für Evaluation, 4. Auflage, Mainz: DeGEVal – Deutsche Gesellschaft für Evaluation.
- [4] Ebcinoğlu, Fatma/Jaeger, Michael/Lezczensky, Michael (2008): Evaluation der Sächsischen Hochschulvereinbarung. Gutachten im Auftrag des Sächsischen Staatsministeriums für Wissenschaft und Kunst, HIS: Forum Hochschule 2/2008, Hannover.
- [5] Frey, Bruno S./Osterloh, Margit, Evaluations: Hidden Costs, Questionable Benefits, and Superior Alternatives (October 2006). University of Zurich IEW Working Paper No. 302; Download unter: <http://ssrn.com/abstract=928354>.
- [6] Garfield, Eugene (1979): *Citation Indexing – Its Theory and Application in Science, Technology, and Humanities*. New York/Chichester /Brisbane/Toronto.
- [7] Gärditz, Klaus Ferdinand (2009), *Hochschulorganisation und verwaltungsrechtliche Systembildung*, Tübingen.

- [8] Glaser, Andreas (2008): Die „Studierfreiheit“ – Analyse eines ungeklärten Grundrechts in historischer Perspektive, *Der Staat* 47, S. 213-240.
- [9] Hirsch, J. E. (2005): An index to quantify an individual's scientific research output, *PNAS* 102 (46): 16569-16572 NOV 15 2005, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1283832>.
- [10] Hornbostel, Stefan (2004): Leistungsparameter und Ratings in der Forschung, in: Hochschulrektorenkonferenz (Hrsg.), *Evaluation – Ein Bestandteil des Qualitätsmanagements an Hochschulen* (Beiträge zur Hochschulpolitik), Bonn, S. 173–187.
- [11] Hornbostel, Stefan (2010): (Forschungs-)Evaluation. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), *Handbuch Wissenschaftspolitik*, Wiesbaden, S. 292 – 309.
- [12] Hüther, Otto (2009), Hochschulräte als Steuerungsakteure?, *Beiträge zur Hochschulforschung* 31, S. 50 – 71.
- [13] Hufen, Friedhelm (1995): *Rechtsfragen der Lehrevaluation an wissenschaftlichen Hochschulen*, Bonn.
- [14] Jaeger, Michael/In der Smitten, Susanne (2009): Evaluation der leistungsbezogenen Mittelvergabe an die Berliner Hochschulen. Gutachten im Auftrag der Berliner Senatsverwaltung für Bildung, Wissenschaft und Forschung, HIS: Forum Hochschule 1/2009, Hannover.
- [15] Jansen, Dorothea/Wald, Andreas/Franke, Karola/Schmoch, Ulrich/Schubert, Torben (2007): Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Zum Einfluss von Rahmenbedingungen auf Forschungsleistung, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59 (2007), S. 125-149.
- [16] Jansen, Dorothea (2010): Von der Steuerung zur Governance: Wandel der Staatlichkeit? In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), *Handbuch Wissenschaftspolitik*, Wiesbaden, S. 39-50.
- [17] Kaufmann, Dieter/Hornbostel, Stefan (2008), Die Rolle des Qualitätsmanagements in Hochschulen. Online-Publikation. In: *wissenschaftsmanagement-online* (Einstellungsdatum 27.10.2008, einzusehen unter http://www.wissenschaftsmanagement-online.de/converis/state_of_the_art_inhalt/84).
- [18] Knie, Andreas/Simon, Dagmar (2008): Evaluationen im Governance-Mix. In: *wissenschaftsmanagement*, S. 24-29.
- [19] Knie, Andreas/Simon, Dagmar (2010): Stabilität und Wandel des deutschen Wissenschaftssystems. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), *Handbuch Wissenschaftspolitik*, Wiesbaden, S. 26 – 38.
- [20] Kohler, Jürgen (2008): Sachliche Maßgaben des Entscheidens: Topoi, Kriterien, Standards, Indikatoren. Begriffe und Zusammenhänge; insbesondere Gegenstände, Geltungsgründe und Risiken von Standards. In: Benz, Winfried/Kohler, Jürgen/Landfried, Klaus (Hrsg.), *Handbuch Qualität in Studium und Lehre*, Berlin, Abschnitt D 1.1.
- [21] Kracht, Stefan (2006): Das neue Steuerungsmodell im Hochschulbereich. Zielvereinbarungen im Spannungsverhältnis zwischen Konsens und hierarchischem Verwaltungsaufbau, Baden-Baden.
- [22] Krawietz, Markus (2006): Evaluation der Evaluation. Erfolg und Misserfolg von Evaluationen an deutschen Hochschulen – Die Perspektive der Studierenden, HIS Projektbericht, Hannover.

- [23] Kromrey, Helmut (2005): Zur Verbindung von Akkreditierung und Evaluation. In: Benz, Winfried/Kohler, Jürgen/Landfried, Klaus (Hrsg.), Handbuch Qualität in Studium und Lehre, Berlin, Abschnitt F. 2.2.
- [24] Ladeur, Karl-Heinz (2005): Die Wissenschaftsfreiheit der „entfesselten Hochschule“ – Umgestaltung der Hochschulen nach dem Ermessen des Staates?, Die öffentliche Verwaltung, S. 753-764.
- [25] Lange, Rainer (2010): Benchmarking, Rankings und Ratings. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, 322-333.
- [26] Lange, Stefan (2008): Bewertung und Rechtfertigung von Wissenschaft in praktischer Bewährung – Umstritten sind Zeitpunkt, Methode und Gegenstand der Leistungsmessung nach wie vor. In: wissenschaftsmanagement, S. 12-13.
- [27] Lange, Stefan (2010): Hochschulräte. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 347-360.
- [28] Laqua, Alexander (2004): Der Hochschulrat zwischen Selbstverwaltung und staatlicher Verwaltung. Eine Analyse der Ratsmodelle nach den Landeshochschulgesetzen, Baden-Baden.
- [29] Lütje, Jürgen (2010): Aktivierendes Wissenschaftsmanagement. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 262-279.
- [30] Matthies, Hildegard/Simon, Dagmar (Hrsg.) (2008): Wissenschaft unter Beobachtung – Effekte und Defekte von Evaluationen. Sonderband der Zeitschrift Leviathan. Wiesbaden.
- [31] Meier, Frank/Schimank, Uwe (2002): Szenarien der Profilbildung im deutschen Hochschulsystem. Einige Vermutungen. In: Die Hochschule 2002: 82-91.
- [32] Müller-Böling, Detlef (1996): Evaluationen zur Rechenschaftslegung oder zur Verbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen (Centrum für Hochschulentwicklung, Arbeitspapier Nr. 12), Gütersloh.
- [33] Müller-Böling, Detlef (2004): Hochschule und Profil – zwischen Humboldt und Markt? In: Hochschulrektorenkonferenz (Hrsg.), Profilbildung an Hochschulen. Grundlage für Qualität und Exzellenz, Bonn, S. 15-23.
- [34] Mußnug, Reinhard (1992): Gefährden Lehrevaluationen die Freiheit der Wissenschaft? In: Mitteilungen des Hochschulverbandes 40, S. 253-256.
- [35] Mutz, Rüdiger/Daniel, Hans-Dieter (2008): Nutzung von Lehrevaluationsdaten für die Qualitätssicherung der Evaluationsinstrumente am Beispiel der Universität Zürich, Beiträge zur Hochschulforschung 30, S. 34.
- [36] Neidhardt, Friedhelm (2010): Selbststeuerung der Wissenschaft: Peer Review. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 280 – 292.
- [37] Neufeld, Jörg (2009): Die leistungsorientierte Mittelvergabe (LOM) in der deutschen Hochschulmedizin, <http://www.forschungsinfo.de/iq/agora/LOM/LOM.asp>.
- [38] Power, Michael (1997): The Audit Society: Rituals of Verification, Oxford.
- [39] Promberger, Kurt/Bernhart, Josef/Niederkofler, Carmen (2006): Grundlagen zur Evaluation von Verwaltungsreformen, Wien.

- [40] Rietschel, Ernst-Theodor/Arnold, Erik (2009): Evaluation of the sixth framework programmes for research and technological development, 2002-2006. Report of the expert group (February 2009), download unter: http://ec.europa.eu/research/reports/2009/pdf/fp6_evaluation_final_report_en.pdf.
- [41] Röbbcke, Martina (2010): Akkreditierung. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 334-346.
- [42] Ruffert, Matthias (2006): Grund und Grenzen der Wissenschaftsfreiheit. In: Veröffentlichungen der Vereinigung der Deutschen Staatsrechtslehrer 65, S. 146-216.
- [43] Schenker-Wicki, Andrea (1999): Moderne Prüfverfahren für komplexe Probleme – Evaluation und Performance Audits im Vergleich, Wiesbaden.
- [44] Schubert, Torben/Schmoch, Ulrich (2010): Finanzierung der Hochschulforschung. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 244-261.
- [45] Schulte, Martin (2006): Grund und Grenzen der Wissenschaftsfreiheit. In: Veröffentlichungen der Vereinigung der Deutschen Staatsrechtslehrer 65, S. 110-145.
- [46] Seckelmann, Margrit/Lauer, Felicia/Jug, Andreas (2006): Gewogen und für zu leicht befunden? Zur Bewertung von Wissenschafts- und Forschungsleistungen, Verwaltung und Management 12, S. 195-203.
- [47] Seckelmann, Margrit (2009a): Evaluationen und deren mögliche finanzielle Konsequenzen. In: Christian Berthold/Brigitte Tag/Hanns H. Seidler/Günter Scholz (Hrsg.), Handbuch Praxis Wissenschaftsfinanzierung, Stuttgart, 10. Ergänzungslieferung (November 2009), A-1-13.
- [48] Seckelmann, Margrit (2009b): „Renaissance“ der Gemeinschaftsaufgaben in der Föderalismusreform II. Zu den Gemeinschaftsaufgaben und den Möglichkeiten kooperativen Handelns im Bundesstaat (insbesondere zu den Art. 91 a und b GG n.F. sowie zur Aufnahme der Art. 91 c und d GG in das Grundgesetz). In: Die öffentliche Verwaltung 62, S. 747-757.
- [49] Seckelmann, Margrit (2010): Rechtliche Grundlagen und Rahmensetzungen. In: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden, S. 227-243.
- [50] Sondermann, Michael/Simon, Dagmar/Scholz, Anne-Marie/Hornbostel, Stefan (2008): Die Exzellenzinitiative: Beobachtungen aus der Implementierungsphase (IFQ-Working-Paper No. 5), Bonn.
- [51] Stüber, Jessica (2009), Akkreditierung von Studiengängen. Qualitätssicherung im Hochschulrecht vor dem Hintergrund der internationalen Entwicklungen im Bildungssektor, insbesondere des Bologna-Prozesses, Frankfurt am Main u. a.
- [52] Vedung, Evert (1999): Evaluation im öffentlichen Sektor, Wien/Köln/Graz.
- [53] Vernau, Katrin (2007): Hochschul-Controlling. In: Benz, Winfried/Kohler, Jürgen/Landfried, Klaus (Hrsg.), Handbuch Qualität in Studium und Lehre, Berlin, Abschnitt E 7.4.
- [54] Walach, Harald (2009): Hurra – wir haben eine neue Religion! Über Qualitätssicherung, Forschung & Lehre, 342 – 345.
- [55] Winter, Martin/Reil, Thomas (2002): Auf dem Weg zu einem integralen Qualitätsmanagementsystem an Hochschulen. In: Dieselben: Qualitätssicherung an Hochschulen. Theorie und Praxis, Bielefeld, S. 6-14.

A 1.3 Einforderung von Qualität: Politische Rahmenbedingungen und Ziele

Deutsche und europäische Hochschulen im Wandel

- [56] Wissenschaftsrat (1985): Empfehlungen zum Wettbewerb im deutschen Hochschulsystem, Köln.
- [57] Wissenschaftsrat (2008): Steuerungsgruppe der Pilotstudie Forschungsrating im Auftrag des Wissenschaftsrats: Forschungsleistungen deutscher Universitäten und außeruniversitärer Einrichtungen in der Soziologie, Köln (10.4.2008), download unter http://www.wissenschaftsrat.de/texte/pilot_ergeb_sozio.pdf.

Information zur Autorin:

Dr. iur. **Margrit Seckelmann**, M. A., ist Geschäftsführerin des Deutschen Forschungsinstituts für öffentliche Verwaltung in Speyer. Nach Studien der Rechtswissenschaften und der Geschichte und Germanistik in Heidelberg und Berlin (abgeschlossen mit dem Ersten Juristischen Staatsexamen und dem M. A. in Geschichte an der Freien Universität) absolvierte sie ein Referendariat in Berlin und Brüssel, das sie mit dem Zweiten Juristischen Staatsexamen abschloss. Daran anschließend promovierte sie in Frankfurt am Main (Doktorvater: Prof. Dr. Dres. h. c. Michael Stolleis) an der Goethe-Universität über „Industrialisierung, Internationalisierung und Patentecht im Deutschen Reich, 1871-1914“ (erschieden bei Klostermann, Frankfurt am Main 2006). Diese Arbeit entstand während ihrer Beschäftigungszeit am Max-Planck-Institut für Europäische Rechtsgeschichte im Projekt „Recht in der Industriellen Revolution“. Die Autorin ist weiterhin wissenschaftlich, u. a. auf dem Gebiet des Wissenschaftsrechts, tätig. Neben der Dissertation seien folgende ausgewählte Publikationen erwähnt: „Rechtliche Grundlagen und Rahmensetzungen“, in: Dagmar Simon/Andreas Knie/Stefan Hornbostel (Hrsg.), Handbuch Wissenschaftspolitik, Wiesbaden 2010, S. 227-243; „Institutionalisierte Sachverständige in Wissenschaft und Medizin? Zur rechtlichen Bewertung von Ethikkommissionen“, Wissenschaftsrecht 2008, 188-204; „Gewogen und für zu leicht befunden? Zur Bewertung von Wissenschafts- und Forschungsleistungen“, Verwaltung und Management 12 (2006), 195-203 (zusammen mit Felicia Lauer und Andreas Jug).

Kontakt:

Dr. iur. Margrit Seckelmann, M. A.

Deutsches Forschungsinstitut für öffentliche Verwaltung Speyer

Freiherr-vom-Stein-Str. 2

67324 Speyer

E-Mail: seckelmann@foev-speyer.de