

Output-Evaluation versus Prozeß-Evaluation

Professor Dr. Detlef Müller-Böling

CHE Centrum für Hochschulentwicklung, Gütersloh

1. Evaluation - Mode oder Notwendigkeit?

Evaluation ist zu *dem* Modewort der hochschulpolitischen Entwicklung geworden. Vor vier Jahren in Deutschland im Zusammenhang mit Hochschulen noch völlig ungebräuchlich, scheint sich damit nunmehr für viele die Lösung aller Probleme des tertiären Bildungsbereichs zu verbinden.

Selbstverständlich ist die Begutachtung und die Bewertung in Hochschulen kein unbekannter Vorgang, dennoch verbirgt sich hinter diesem Sammelbegriff nunmehr eine Vielzahl von unterschiedlichen neuen Ansätzen.

2. Funktionen von Evaluationen

Evaluationen haben im wesentlichen zwei Funktionen (auch Kieser u.a. 1996). Sie sollen einmal der

- Qualitätssicherung in den Hochschulen und zum anderen der
- Rechenschaftspflicht der Hochschulen gegenüber der Gesellschaft — inkl. Mittelzuteilung

dienen

Man muß sich natürlich fragen, warum es heute oder gerade jetzt zu einem derartigen Instrument für die Qualitätssicherung kommt und warum nunmehr die Rechenschaftspflicht eine neue Bedeutung erlangt.

2.1. Qualitätssicherung

Es wäre verfehlt anzunehmen, die Hochschulen hätten in der Vergangenheit keine Instrumente der Sicherung ihrer Qualität eingesetzt. Im Wesentlichen handelte es sich dabei um Formen einer ex-ante Steuerung, d.h. durch vorher ergriffene Maßnahmen wird späteres Handeln und Wirken vorherbestimmt. Dabei handelt es sich häufig um Festlegungen über einen Zeitraum von mehr als 30 Jahren, so etwa bei Berufungen, die die Qualität eines Lehrstuhls oder Arbeitsbereichs über Jahre hinweg bestimmen. Ex-ante, d.h. vor Arbeitsbeginn wird die Qualität geprüft, anschließend bestehen keinerlei Eingriffsmöglichkeiten mehr. Dieses Prinzip ist auch leitend für die Mittelzuteilungen, wobei der Staat durch detaillierte Haushaltsvorgaben versucht, auf die Qualität

steuernd einzuwirken. Gleiches gilt für die Genehmigung von Verfahren innerhalb der Hochschule in Form von festgelegten Ordnungen (z. B. die Genehmigung von Prüfungsordnungen, Studienordnungen etc.). Auch hier versucht der Staat Einfluß zu nehmen auf zukünftiges Handeln der Hochschulangehörigen.

Das Prinzip der ex-ante-Steuerung ist kennzeichnend für die kontinental-europäischen Staaten. Dieses System hat unbestreitbare Vorteile und Erfolge in der Vergangenheit erzielt. Es sichert

- eine große Homogenität in der Qualität, ohne allerdings zwingend hohe Qualität zu garantieren (dabei ist die Gleichwertigkeit und Einheitlichkeit der Abschlüsse allerdings eine Fiktion, für die sehr viel Koordinationsaufwand betrieben wird.),
- eine große individuelle Freiheit derjenigen, die berufen sind,
- eine Finanzverantwortung des Staates im Hinblick auf die Alimentierung der Hochschulen.

Die Nachteile dieses Systems bestehen in

- der hohen Inflexibilität:
Bis Prüfungsordnungen entsprechend neuen gesellschaftlichen Anforderungen oder einem europäischen Wettbewerb geändert werden, vergehen Jahre;
- einem individuellen Freiheitsmißbrauch:
Dabei handelt es sich weniger um tatsächliche Rechtsbrüche, etwa durch die Mißachtung von Lehrdeputaten oder durch zu geringe Anwesenheit. Der Mißbrauch besteht vielmehr in der fachlichen Spezialisierung und Nischenpolitik von Lehrenden, so daß Lernende nicht mehr umfassend (aus)gebildet werden (akademischer Individualismus);
- den Problemen, sofern der Staat seiner Finanzverantwortung nicht mehr nachkommen kann.

Die frühere ex-ante-Steuerung basiert auf Haushaltsvorgaben, gesetzlichen Vorschriften über die Organisationsstruktur, Genehmigungsvorbehalten hinsichtlich Fachbereichen, Studiengängen und Berufungen. Der Staat hat versucht, durch Fixierung wesentlicher Rahmenbedingungen im vorhinein Qualität zu sichern. In einer stabilen Umwelt mit ausreichenden finanziellen Ressourcen hat sich dieses Verfahren durchaus bewährt. Bei instabilen Umweltverhältnissen (veränderte Anforderungen der Gesellschaft an die Ausbildungsinhalte, explodierendes Wachstum der Bildungsnachfrage) und stagnierenden oder zurückgehenden Ressourcen erweist sich dieses Steuerungsverfahren als nicht mehr erfolgreich.

Die ex-ante-Steuerung reicht nicht mehr aus, weil

- die Festlegungen bei Berufungen über einen Zeitraum von 25 Jahren den Erfordernissen der rasanten Umweltentwicklung nicht mehr gerecht werden, darüber hinaus unterschiedliche Leistungsstufen im Verlauf eines Wissenschaftlerlebens bestehen oder Fehler nicht zu perpetuieren sind,
- die Abstimmungsprozesse zur Koordinierung der Homogenität zu langsam gehen und
- der Staat seiner Finanzverantwortung nicht mehr in ausreichendem Maß gerecht wird.

Dies ist der Grund, warum weltweit eine Tendenz zur ex-post-Steuerung der Qualität eingesetzt hat, die an den Ergebnissen, an den Leistungen der Hochschulen ansetzt und danach Belohnungen oder Bestrafungen verteilt, mittels beruflichem Aufstieg, Mittelzuweisungen oder -kürzungen sowie Ansehensgewinn oder -verlust.

Dabei ist das, was man unter Qualität im Hochschulbereich zu verstehen hat, außerordentlich schwer zu definieren. Ich will es daher auch eher mit einem Vergleich versuchen, der die Problematik aufzeigt. Danach ist Qualität der Hochschule mit der Liebe vergleichbar:

- nicht faßbar, aber doch vorhanden,
- erlebbar, aber nicht quantifizierbar,
- stets flüchtig, so daß man sich immer erneut und dauerhaft um sie bemühen muß.

2.2. Rechenschaftspflicht

Am 8. Januar 1994 schrieb der englische Economist:

"Vor dreißig Jahren waren die Universitäten unbestreitbar die am meisten verhätschelten Institutionen der Welt. Nun sind die Universitäten überall in der Defensive. Es besteht kein Vertrauen mehr seitens der Regierungen. Sie müssen sich Sorgen machen über zurückgehende Finanzmittel und den Verlust von Einfluß. Nicht mehr und nicht weniger als ein allgemeiner Pendelrückschlag gegen die akademische Welt ist auf dem Vormarsch."

Und in der Tat genau dies ist Fakt: Während in früheren Jahren grundsätzlich angenommen wurde, das in die Hochschulen investierte Geld sei sinnvoll angelegt und die dort tätigen Wissenschaftler würden Gutes damit tun, ist dieses Vertrauen in die Hochschulen deutlich geschwunden. Die Gründe hierfür mögen vielfältig sein:

- Einmal spielt sicherlich der Mangel in den öffentlichen Kassen eine wichtige Rolle.
- Zum zweiten aber dürfte generell auch das Vertrauen in die Leistungsfähigkeit der Wissenschaft zur Lösung der politischen, ökonomischen und ökologischen Probleme gesunken sein, nicht zuletzt auch wegen der vielfältigen Lösungen, die in der Regel für ein Problem angeboten werden.

Dies ist, wie das Zitat aus einer britischen Zeitung belegt, ein weltweites Problem. Insofern fragt die Gesellschaft, fragt der Staat nicht nur in Deutschland nach den nachweisbaren Relationen zwischen dem, was in die Hochschulen hineingesteckt wird und dem, was bei ihnen herauskommt. Allerdings ist diese Diskussion in Deutschland zeitversetzt (zur internationalen Entwicklung Müller-Böling 1995). Im übrigen scheinen die deutschen Hochschulen besonders kritikresistent zu sein. Ihre geringe Selbstkritikfähigkeit unterstützt dieses Phänomen zusätzlich.

Während die Funktion der Qualitätssicherung letztlich nach innen gerichtet ist auf Entscheidungs- oder Organisationsentwicklungsprozesse innerhalb der Hochschulen, ist die Funktion der Rechenschaftspflicht nach außen gerichtet auf den Nachweis von Leistungen und auf die Kontrolle der Gesellschaft bzw. des Trägers der Hochschule.

Es scheint mir notwendig zu sein, die bestehenden Evaluationsverfahren und -instrumente im Hinblick auf diese beiden Funktionen, die durchaus unterschiedlicher Natur sind, abzuprüfen, um letztlich beurteilen zu können, welche Verfahren für welche Funktion besonders geeignet sind.

was Fachbereich Forschergruppe nicht Individuum

3. Evaluationsverfahren

Unter dem Begriff Evaluation lassen sich mittlerweile eine Vielzahl von Verfahren subsumieren, die alle zu einer Bewertung von Leistungen führen. Dabei sind es im wesentlichen ex-post-Analysen, die an bereits erbrachten Leistungen ansetzen und diese dann bewerten. Eine Ausnahme bildet die Forschungsförderung, stellvertretend erläutert am Beispiel der Deutschen Forschungsgemeinschaft.

3.1. Quantitative vs. qualitative Ansätze

Die quantitative Methode der performance indicators versucht über intersubjektiv überprüfbare Leistungskennzahlen Aspekte des Inputs (Studentenzahlen), des Throughputs (Zwischenprüfungen, Dropout-Quote) und des Outputs (Absolventenzahlen, Studiensemester, Promotionen, Publikationen) statistisch zu erfassen. Dagegen werden bei den qualitativen peer reviews die von der akademischen Gemeinschaft entwickelten und akzeptierten Leistungsstandards in einem letztlich subjektiven Bewertungsprozeß zur Beurteilung der Ausbildungssituation, der Qualität der Lehr- und Lernprozesse oder der Forschungsergebnisse herangezogen.

Allerdings kommen auch die peers ohne gesicherte quantitative Datengrundlage nicht aus. Umgekehrt ergeben die Leistungskennzahlen ohne eine qualitative Interpretation ebensowenig (höheren) Sinn. Grundsatz muß daher sein, beide Verfahren einzusetzen.

Insbesondere bei den quantitativen Verfahren ergeben sich zwei Probleme, die in der empirischen Forschung unter den Begriffen der Validität und der Reliabilität diskutiert werden. Welcher erfassbare Indikator kann Leistung repräsentieren? Dies ist eine Frage der Validität. Weitestgehend akzeptiert sind hier beispielsweise Drittmittel, zumindest soweit sie aufgrund von Bewertungsverfahren verteilt werden. Ein Problem der zuverlässigen Messung dieses Indikators ist dann, welche Drittmittel sollen von wem in Deutschland wie erhoben werden. Dabei handelt es sich um die Frage der Reliabilität der Messung. Hier gibt es - wie das Pilotprojekt "Profilbildung" der HRK oder eine kürzliche Aufstellung des Wissenschaftsrates gezeigt haben - große Erhebungsunterschiede, die erst vereinheitlicht werden müssen, bevor diese Daten belastbar sind. Selbstverständlich ist, daß diese Zahlen dann nur Fach-zu-Fach verglichen werden können.

Die im folgenden beschriebenen Verfahren sind jeweils eher quantitativ oder qualitativ orientiert, ohne selbstverständlich jeweils die andere Seite auszuschließen.

3.2.	Deutsche Forschungsgemeinschaft	qualitativ
3.3.	Lehrberichte <i>Forschungsberichte</i>	qualitativ
3.4.	Selbstbericht + Peer Review <i>Un. Do.</i>	qualitativ
3.5.	Gutachten Unternehmensberater	qualitativ
3.6.	Hörerbefragungen	quantitativ
3.7.	Absolventen <i>Absolventen</i> befragungen <i>Kollegen</i>	quantitativ
3.8.	bibliometrische Messungen	quantitativ
3.9.	Rankings	quantitativ

3

3.2. Evaluation der Deutschen Forschungsgemeinschaft (DFG)

Die Gutachter versuchen sich im Vorhinein ein Bild von der Qualität eines Forschungsprojektes zu machen. Aufgrund dieser ex-ante-Bewertung erfolgt die Mittelzuteilung. Abschließend erfolgt eine Projektbewertung, die allerdings keine Konsequenzen für das laufende Projekt, etwa durch Rückforderung von Geld, hat. Allerdings geht diese ex-post-Bewertung ein in eine gegebenenfalls neue ex-ante-Bewertung für ein weiteres Projekt.

3.3. Lehrberichte als Form der Selbstevaluation

Die Selbstevaluation ist das tragende Steuerungselement innerhalb einer Hochschule. Die deutschen Hochschulen sind aufgebaut auf

- dem Kollegialitätsprinzip, gekoppelt mit Hierarchiefreiheit
- der Selbstorganisation einschließlich der Selbstkontrolle.

Diese Selbstkontrolle funktioniert bei einer auf gemeinsamen Wertvorstellungen und Qualitätsstandards beruhenden Organisationskultur zumindest auf der Ebene der Fachbereiche, ist aber angesichts des immensen Wachstums in den 60iger und 70iger Jahren verloren gegangen. Darüber hinaus ist diese Selbstkontrolle bei den Professoren durch die Gruppenuniversität mit ihren knappen Mehrheitsverhältnissen verlorengegangen, weil man sich im Zweifel auch um die schwarzen Schafe der Zunft scharen mußte, um Abstimmungs-niederlagen zu vermeiden.

Aus diesem Grund scheint eine Fremdkontrolle auch bei Selbstevaluation notwendig zu sein. Fremdkontrolle darf aber nicht verwechselt werden mit Fremdorganisation!

Aus staatlicher Sicht sind Lehrberichte ein Versuch, von außen die Effizienz und Effektivität des Hochschulsystems mit einem aus dem inneren herauskommenden Kontrollinstrument zu sichern. In der gemeinsamen Empfehlung von KMK und HRK zur Umsetzung der Studienreformen aus dem Jahr 1993 und in den 10 Thesen des Wissenschaftsrates wird empfohlen, Lehrberichte zu erstellen. In verschiedenen Länderprogrammen zur Qualitätssicherung der Lehre und in den meisten Hochschulgesetzen der Länder sind Lehrberichte inzwischen verbindlich vorgeschrieben. In Baden-Württemberg, Nordrhein-Westfalen, Hessen, Sachsen u. a. liegen Lehrberichte aus zahlreichen Hochschulen und Fächern (zum Teil flächendeckend) vor. Ebenfalls in den Hochschulgesetzen sind die Lehrberichte im Saarland sowie in Rheinland-Pfalz verankert. Hinsichtlich Inhalt und Struktur weisen die vorliegenden Lehrberichte große Unterschiede auf. Dies liegt nicht zuletzt an den großen Differenzen in den gesetzlichen Vorgaben. Während Nordrhein-Westfalen (Dekan) und Baden-Württemberg (Fakultätsrat) lediglich Verantwortliche für die Erstellung des Lehrberichts festsetzen, geben Niedersachsen, Sachsen, Sachsen-Anhalt, Schleswig-Holstein, Mecklenburg-Vorpommern und Hessen teilweise detailliert die Inhalte vor. (Bisher) nicht gesetzlich vorgeschrieben sind Lehrberichte in Bayern, Bremen, Brandenburg und Thüringen.

Lehrberichte sollen das instrumentelle Rückgrat der Evaluationsprojekte und anderer Aktivitäten zur Qualitätssicherung an den Hochschulen in der Bundesrepublik bilden. Von ihnen wird mehr Transparenz in Bereich des Lehr- und Studienbetriebs der Hochschulen, Fakultäten und Fächer erwartet. Ihr systemischer Ansatz soll das Blickfeld über die einzelne Lehrveranstaltung und deren

Beurteilung hinaus auf Studienabschnitte, Studiengänge sowie Rahmenbedingungen der Ausbildung, von Studien- und Prüfungsordnungen bis hin zu Ressourcen und Lehrmanagement erweitern. Lehrberichte sollen außerdem über Maßnahmen zur Studienreform, Qualitätssicherung und -verbesserung sowie über Meinungen der Lehrenden und Studierenden zur Ausbildung informieren.

Es zeigt sich, nicht nur hier, daß Transparenz allein keine kontrollierende Wirkung hat. Da die inhaltliche Zielsetzung der bisherigen Lehrberichte letztlich ebenso diffus ist wie ihr Adressatenkreis (Ministerium, Öffentlichkeit, zukünftige/jetzig Studierende?), ist ihre Wirkung äußerst begrenzt. Darüber hinaus kann zuviel Information ebensoviel Intransparenz hervorrufen, wie keine Information. Insofern sind Lehrberichte tatsächlich eher als ein Instrument der Selbstkontrolle eines Fachbereichs anzusehen, das, wenn die Selbstkontrolle denn funktioniert, auch eine positive Wirkung entfalten kann.

3

3.4. Selbstbericht und Peer Review

Auf diesen Überlegungen aufbauend wird neben dem Selbstbericht nunmehr immer häufiger auch in Deutschland die externe Begutachtung durch Peers adaptiert. Im Wesentlichen orientieren sich alle Aktivitäten in Deutschland am niederländischen Modell, in dem anhand eines vorher fixierten Fragenkatalogs Selbstberichte der Fachbereiche erstellt werden, die externen Gutachtern als Grundlage für Fragen bei ihren Besuchen dienen, aus denen dann anschließend ein Fremdbericht erstellt wird.

3.4.1. Pilotprojekt an den Universitäten Hamburg und Kiel (Geschäftsstelle HIS)

1993 hat haben die Behörde für Wissenschaft und Forschung in Hamburg sowie das Ministerium für Wissenschaft, Forschung und Kunst in Schleswig-Holstein an den Universitäten Hamburg und Kiel die Evaluation der Fächer Germanistik und Biologie initiiert und von HIS organisieren lassen (Reissert 1994).

3.4.2. Nordverbund

Vorbild: Das niederländische Verfahren. Keine Verknüpfung mit der Mittelzuweisung; kein Vergleich zwischen den Hochschulen;

Um in dieses starre System institutionell, fachbezogen und regional Bewegung zu bringen, gründeten die norddeutschen Universitäten Bremen, Hamburg, Kiel, Oldenburg und Rostock einen Verbund zur Evaluation von Studium und Lehre. Das Ziel ist, durch die Entwicklung kommunikativer Strukturen zwischen den Vertretern eines Faches innerhalb der Universität und im Rahmen des Verbundes die Qualität zu sichern. Das Instrument dazu ist eine Evaluation nach dem niederländischen Modell, verknüpft mit einer auswertenden Konferenz aller Beteiligten. Nun geht es darum, Strukturen der Qualitätssicherung aufzubauen, die gewährleisten, daß die Empfehlungen in den Gutachten umgesetzt werden.

(Lüthje im Buch zu unserer Eröffnungstagung) Orientiert am niederländische Modell; Zusammenschluß der Universitäten HB, HH, Kiel, Oldenburg, Rostock. Pro Jahr Evaluierung von zwei Fächern mit dem Ziel der Verbesserung der Qualität in Studium und Lehre; Teilnahme an der Evaluierung ist freiwillig; Auswertung der Ergebnisse und Evaluation der Evaluation (um das Verfahren kontinuierlich zu verbessern): dem dient die 'auswertende Tagung', an der alle

Beteiligten teilnehmen; Interne Evaluation im SS, Begutachtung und Tagung im WS Gemeinsam mit dem CHE. Verfahrensentwicklung zur Umsetzung von Empfehlungen auf der Basis schriftlicher Regelungen (z.B. Verträgen) zwischen Fachbereich und Hochschulleitung)

3.4.3. HRK EU

HRK als nationale Agentur im Rahmen des europäischen Pilotprojekts zur Qualitätssicherung und -bewertung; Ziele: Erprobung einer Methode, die allgemein angewendet werden kann; Sensibilisierung für die Notwendigkeit von Maßnahmen der Qu.-Sicherung; Einbringung einer europäischen Dimension; erhöhte Transparenz; beteiligte Hochschulen: TU Dresden(und FH Osnabrück mit dem Fach Maschinenbau/-wesen; U/GH Essen und FH Hannover mit dem Fach Kommunikationsdesign Grundsätze: Selbstbewertung; Einhaltung gemeinsamer Leitlinien; Bewertung durch externe Peers; Veröffentlichung eines Berichts; Prüfung der Berichte durch einen europäischen Ausschuss

Die Hochschulrektorenkonferenz will mit dem Pilotprojekt Profilbildung die Fundamente für eine regelmäßige Berichterstattung legen, sowohl quantitativ als auch qualitativ Transparenz schaffen, der Forderung nach Rechenschaft Nachdruck verleihen und den Wettbewerb zwischen den Hochschulen fördern. Wie im Verfahren für die Universitäten wird der Selbstdarstellung der Fachhochschule und der einzelnen Studiengänge breiten Raum eingeräumt und der verbale Teil durch einen statistischen Teil ergänzt. Dieser umfaßt eine differenzierte Aufbereitung der Daten zu den Studierenden, dem wissenschaftlichen Personal, der Forschung, der Bibliotheksaustattung etc. Das Problem ist, daß mit diesen Daten etwas untersucht wird, was eigentlich gar nicht existiert: Das Profil einer Hochschule inmitten einer Hochschulandschaft, die durch eine politisch so gewollte Homogenität geprägt wird. Die grundgesetzlich garantierte Gleichheit der Lebensverhältnisse durchzieht viele Bereiche unseres Lebens als quasi ideologischer Anspruch.

3.4.4. Wissenschaftsrat

(Hofmann im Buch zu unserer Eröffnungstagung; außerdem gibt es die Empfehlung des Wissenschaftsrats zur Lehrevaluation) Orientiert am niederländische Modell; Ziel: Erprobung des Evaluationsverfahrens und Test auf die Tragfähigkeit; Fächer: Physik (U Freiburg, TU Berlin, U Saarnrücken, U Osnabrück, U Leipzig) und Betriebswirtschaftslehre (FH Frankfurt, FHW Berlin; U München, U Köln, U Kiel, GH Siegen); quantitative Kennziffern sind keine Indikatoren sondern Rohdaten für daraus abgeleitete Frage der Gutachter Begehung der Fachbereiche im WS 94/95 und im SS 95 (jeweils 1,5 Tage)

3.4.5. Niedersachsen

Nach zweijähriger Diskussion im Vorfeld habe sich die LHK durchgerungen, der Einrichtung der Evaluationsagentur zuzustimmen. Die Agentur versteht sich als Selbsthilfeeinrichtung der niedersächsischen Hochschulen, ist also von ihrem Selbstverständnis vorrangig den Hochschulen verpflichtet.

Die Evaluation zielt primär auf Lehre und Studium.

Die Evaluation verläuft in einem zweistufigen Verfahren:

- 1. die Selbstevaluation, daran anschließend***
- 2. die externe Evaluation (peer review), die mit Unterstützung der Evaluationsagentur den endgültigen Evaluationsbericht erarbeiten.***

Im Unterschied zu dem Nordverbund sind

1. die Evaluationsberichte zweigeteilt und umfassen erstens eine Einschätzung des Faches A in Niedersachsen und zweitens die Einschätzung des Faches an der Hochschule.
 2. die Abschlußberichte am Ende eine Einschätzung einzelner Fachbereiche und Standorte "in aggregierter Form", was in der Summe einer Profilbeschreibung entspricht.
- Angestrebt ist eine für die interessierte Öffentlichkeit brauchbare Publikation auf Basis der Abschlußberichte.

In der ersten Runde werden 5 Fächer evaluiert:

- Elektrotechnik, getrennt für die Unis/FHs (daher zweimal gezählt)
- Chemie
- Sozialpädagogik (Uni)/Sozialwesen (FHs)
- Geschichte

insgesamt würden 29 Fachbereiche an 19 Hochschulen in der ersten Runde evaluiert.

Zeitplan: Die Selbsteinschätzung beginne im SS 96, die externe Evaluation im WS 96/97. Es sei ein laufendes Verfahren angestrebt, sodaß die zweite Runde im WS 96/97 mit der Selbsteinschätzung beginne. Auf dieser Basis wolle man in fünf bis sechs Jahren alle Fächer in Niedersachsen evaluiert haben.

Die von den Fachbereichen erstellten Berichte zur Selbstevaluation gingen an die "peers", ansonsten hätten die Fachbereiche das "Copyright". Die Fächer erhalten für die im Rahmen der Selbstevaluation erbrachten Eigenleistungen bis zu 30.000 DM von der Agentur. Die Daten liegen im Prinzip auf der Grundlage der Hochschulstatistik den Fachbereichen vor. Die Gutachten der peers wiederum gingen an die Fachbereiche/Unileitung, nicht hingegen an das Ministerium/die Lenkungsgruppe.

Die Gutachten und Selbstberichte werden als Evaluationsberichte zusammengeführt, die unter den Fächern ausgetauscht werden können, sofern diese einem Tausch zustimmen. Der Abschlußbericht wird von ZEvA verfaßt; dies heißt insbesondere eine Aggregation und Anonymisierung der Ergebnisse.

Die Lenkungsgruppe ist der Agentur beigeordnet und steuert den Evaluationsprozeß. Vorsitzender (ohne Stimmrecht) ist Herr Seidel, Präsident Uni Hannover und zugleich der Leiter der Agentur.

Sie umfaßt fünf stimmberechtigte Mitglieder:

1. Vorsitzender der LHK (oder Stellvertreter) derzeit Herr Künzel, Uni Osnabrück
2. amtierendes oder früheres Mitglieder einer Uni (benannt von der LHK), derzeit Frau Ehmke, Vizepräsidentin TU Braunschweig
3. amtierendes oder früheres Mitglieder einer FH, derzeit Herr Jaudzims, Präsident FH Hannover
4. Vertreter des MWK (inkl. Stellvertreter ohne Stimmrecht), derzeit Palandt/Valentien
5. eine Persönlichkeit, die über Erfahrungen in der Evaluation im nat./int. Bereich verfügt (benannt von der LHK), derzeit Herr Schnitzer, HIS.

Wichtig ist die Lenkungsgruppe insbesondere bei der Benennung der jeweils zu evaluierenden Fächer, wo bislang unterschiedliche Vorstellungen seitens der in der Lenkungsgruppe vertretenen Interessengruppen bestanden hätten (MWK - Hochschulen). Herr Reuke betont die etwas unterschiedlichen Interessen, die Staat bzw. Hochschulen mit der Agentur verfolgen: der Staat sehe in der Agentur primär ein Planungs- und Rechenschaftsinstrument, die Hochschulen sehen die Rolle der Agentur vor allem im Bereich Qualitätssicherung der Lehre. Die Konstruktion der Lenkungsgruppe trägt dem Rechnung

und dient der Ausbalancierung ungerechtfertigter Ansprüche. Dies zeigt sich auch bei der Benennung der Fächer, wo im ersten Durchlauf u. a. die Befürchtungen der Universität Hildesheim (Fach Informatik) berücksichtigt worden ist, die Evaluationsergebnisse könnten die Schließung des Standortes legitimieren.

Zur Zusammensetzung der peer-groups (ehrenamtlich, zwischen drei bis fünf Personen):

- *Frauenanteil*
- *eine "fach-ferne" Person*
- *keine Niedersachsen*
- *Studierende in den Fächern, wo diesbezüglich ein Konsens hergestellt werden kann und die kontinuierliche Mitarbeit während eines Evaluationsdurchlaufs gesichert ist.*
- *Optional: "Abnehmer-Seite" (Wirtschaft ..), Mitglied aus dem Ausland*

Die Fachbereiche haben ein Vorschlagsrecht zur Besetzung der peer-group. Die Berufung erfolgt dann durch die ZEvA im Benehmen mit den Hochschulen.

was geschieht danach?

nach Abschluß des Evaluationsverfahrens legen die Fächer ein differenziertes Maßnahmenprogramm vor (Adressaten: MWK und Evaluationsagentur).

nach zwei Jahren liefern die Fächer einen Zwischenbericht über den Stand der Umsetzungen.

Das Maßnahmenprogramm stellt die Grundlage für Verhandlungen mit der Hochschulleitung/dem MWK.

generell ist keine unmittelbare Verknüpfung zwischen Evaluationsergebnissen und der Ressourcenverteilung beabsichtigt.

3.4.6. Universität Dortmund

3.5. Gutachten von Unternehmensberatern

Evaluationen werden in erster Linie auf Lehre und der Forschung bezogen. Notwendig ist es aber auch, die Managementprozesse der Hochschulen mit einzubeziehen, beispielsweise die Fähigkeit zu Entscheidungen zu kommen. Am Beispiel von Berufungen zeigt sich die Problematik, die keineswegs geringer ist als bei der Bewertung von Forschung und Lehre: Der Zeit-, Arbeits- und Gutachtenaufwand muß z. B. ins Verhältnis zur Qualität der Berufenen gesetzt werden. Ohne diesen hohen Anspruch zu erfüllen, haben sich Unternehmensberater insbesondere mit dem Aufbau und dem Ablauf von Hochschulverwaltungen beschäftigt, so in Nordrhein-Westfalen Mummert&Partner (199x) K. Neuvians bitte gutachten hier zitieren!!!!!!!!!!!!

Hayek,

3.6. Hörerbefragungen

- durch Dozenten

3.7. Absolventenbefragungen

3.8. Bibliometrische Messungen

3.9. Rankings

2 4. "Gutachter" bei Evaluationen

Bei den oben angeführten Evaluationsverfahren wird deutlich, daß sie jeweils auf unterschiedliche Personen(gruppen) als diejenigen, die die Bewertung vornehmen, Bezug nehmen. Es sind dies

- Studierende
- Absolventen
- Arbeitsmarkt / *Forschungs markt*
- Öffentlichkeit/Journalisten
- externe Wissenschaftler (peers)
- Unternehmensberater
- Wissenschaftsforscher.

Jeder dieser "Gutachter" ist in der Lage, jeweils aus unterschiedlicher Sicht einen Teil der Beurteilungsfelder sachkundig abzudecken. Insofern haben alle ihre Berechtigung für die unterschiedlichsten Bereiche und Zwecke.

5. Evaluations-Extrema

Es zeigt sich, daß Evaluationen in zwei Extremen betrieben werden können:

- sehr detailliert, aufwendig und gerecht im Urteil, damit kosten- und zeitintensiv wie es augenblicklich in Großbritannien sowohl im Bereich der Forschung wie der Lehre geschieht,
- sehr grob, ungenau und ungerecht, damit wenig kosten- und zeitintensiv wie es in Australien in drei Runden praktiziert wurde (Müller-Böling u.a. 1995).

Das erste Modell, das ich Output-Evaluation nennen möchte, weil es detailliert auf die Forschungs- und Lehrerergebnisse abhebt, erfährt größere Widerstände bei den Beteiligten und dient in erster Linie der Rechenschaftspflicht. Damit ist sie dem Versuch der Manipulation durch die Begutachteten ausgesetzt, die verständlicherweise versuchen, sich im besten Licht darzustellen. Diese Versuchung wird verstärkt, wenn die Evaluation mit Mittelzuweisungen verbunden wird, wie es in England im Bereich der Forschung der Fall ist. Fragwürdig ist die Output-Evaluation im Hinblick auf die qualitätsverbessernden Wirkungen nicht zuletzt aufgrund der geringen Akzeptanz bei den Wissenschaftlern.

Müller-Böling, Detlef, Andreas Barz und Klaus Neuvians: Die jüngste Entwicklung des Australischen Hochschulsystems, in: Wissenschaftsmanagement, 1. Jg. 1995, S. 145 - 148.

Reissert, R.: Evaluation der Lehre - interne Selbstevaluation und externe Begutachtung durch Peers, in: HIS Kurzinformation A8/1994.

Richter, Roland: Das niederländische Modell (Gewerkschaft Erziehung und Wissenschaft, Materialien und Dokumente Hochschulen und Forschung, 73), o.O: 1994.

Webler, Wolff-Dietrich: Das Modell eines Lehrberichts über die Evaluation von Lehre und Studium und erste Ergebnisse, in: Das Hochschulwesen, (1995), 4, 258-266.

Wissenschaftsrat (Hrsg.): Empfehlungen zur Stärkung der Lehre an den Hochschulen durch Evaluation, Bonn 1996.

Zentrale Evaluationsagentur der niedersächsischen Hochschulen (Hrsg.): Leitfaden für die Selbstevaluation von Studium und Lehre an niedersächsischen Hochschulen, Hannover 1996.