

Wie Qualität sichern? Wie Qualität kontrollieren?

Parlamentarisches Hearing zur Wissenschaftspolitik
Kommission für Wissenschaft, Bildung und Kultur

Bern 27./28. August 1998

1 Vorbemerkung

Qualitätssicherung ist „in“, und Evaluation als ein mögliches Instrument der Qualitätssicherung und -kontrolle ist sogar zu *dem* Modewort der hochschulpolitischen Entwicklung geworden. Auffallend ist dabei, daß in diesem Zusammenhang Erwartungen oder Forderungen eine große Rolle spielen, die letztlich einen Paradigmenwechsel in drei wesentlichen Bereichen des Hochschulsystems widerspiegeln:

- Hochschulen müssen ein deutlicheres Qualitätsbewußtsein entwickeln,
- Hochschulen sind in einem stärkeren Maße als bislang gegenüber der Öffentlichkeit rechenschaftspflichtig und
- Hochschulen sind nach Leistung zu finanzieren.

Qualitätssicherung, Rechenschaftspflichtigkeit und Mittelverteilung sind somit die mit Evaluation verbundenen neuen Anforderungen oder Erwartungen, die in ihren zentralen Veränderungen jeweils einen Paradigmenwechsel symbolisieren und im folgenden kurz nachgezeichnet werden sollen.

2 Drei Paradigmenwechsel

2.1 Qualitätssicherung

Qualitätssicherung bzw. Qualitätsmanagement müssen im allgemeinen folgende Anforderungen erfüllen:

- Qualität muß *geplant* werden, indem auf Fachbereichs- und Hochschulebene gemeinsame Ziele und Standards definiert werden;
- Qualität muß *organisiert* werden, indem Regeln für qualitätssicherndes Handeln formuliert werden;

- Qualität muß *gesteuert* werden, indem Maßnahmen zur Prüfung der Zielerreichung und, wenn erforderlich, zur Gegensteuerung entwickelt werden;
- und schließlich erfordert Qualität auch, daß Mitarbeiter zu qualitätsbewußtem und -förderndem Verhalten *motiviert* werden.

Nun wäre es aber verfehlt zu glauben, Hochschulen hätten bislang keine Qualitätssicherung betrieben. Mehr oder weniger informelle Verhaltensnormen, gemeinsame Wertvorstellungen, eine akademische Kultur des intellektuellen Austauschs oder das für Wissenschaftler leitende Kriterium der Exzellenz sind akzeptierte und praktizierte Elemente der Qualitätssicherung innerhalb der „scientific community“, denen komplementäre Instrumente der staatlichen Hochschulsteuerung gegenüberstehen.

Neu ist jedoch, daß Qualitätssicherung zunehmend als Managementaufgabe auf allen Ebenen der Hochschulen verstanden wird, die ein deutlicheres Qualitätsbewußtsein, eine systematischere Herangehensweise und eine integrative Sicht der damit verbundenen Aspekte verlangt. Neue Formen der Qualitätssicherung im Hochschulbereich sind daher zu entwickeln, die die gerade erwähnten Anforderungen erfüllen.

Verschiedene Entwicklungen der vergangenen Jahrzehnte haben zu diesem Perspektiven- oder Paradigmenwechsel geführt:

1. Das bislang in kontinental-europäischen Staaten vorherrschende System der Ex-ante-Qualitätssicherung ist geeignet für eine stabile Umwelt mit ausreichenden finanziellen Ressourcen. Bei instabilen Umweltverhältnissen und stagnierenden oder zurückgehenden Ressourcen erweist sich dieses Steuerungsverfahren als nicht mehr erfolgreich. Daher hat weltweit eine Tendenz zur ex-post-Steuerung der Qualität eingesetzt, 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.
2. Eine weitere grundlegende Veränderungsnotwendigkeit ergibt sich aus einem anderen Wissenschaftsverständnis heraus:
 - Wenn Forschung immer weniger der Logik einzelner Disziplinen folgt, sondern immer häufiger in Anwendungskontexten und temporären, trans-disziplinären Netzwerken stattfindet, dann kann Qualitätssicherung nicht mehr ausschließlich auf der Basis von (in der Regel disziplinenorientierten) Peer Reviews betrieben werden.
 - Der Umsetzungsprozeß von Ergebnissen aus der Grundlagenforschung in den Anwendungsbereich folgt immer weniger dem traditionellen linearen Muster. Vielmehr liegen Anwendungsbezüge in wachsendem Maße bereits der Definition von Forschungsvorhaben zugrunde.

Wenn diese Beobachtungen stimmen, dann sind zusätzliche Kriterien und veränderte Verfahren zur Qualitätssicherung erforderlich, um verschiedenen Interessen, Ansprüchen und Erwartungen auch *außerhalb* einer Disziplin und *außerhalb* der Hochschulen als Institutionen der (Grundlagen-)Forschung hinreichend Rechnung zu tragen. Dabei müssen Aspekte wie die Wettbewerbsfähigkeit von Forschung, das Verhältnis zwischen dem betriebenen (finanziellen) Aufwand und den erzielten bzw. zu erwartenden Ergebnissen ebenso an Bedeutung gewinnen wie die Verantwortung und Rechenschaftspflichtigkeit von Forschung gegenüber ihrem weiteren gesellschaftlichen Umfeld stärker zu berücksichtigen sind.

2.2 Rechenschaftspflicht

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:

- Das Vertrauen in die Leistungsfähigkeit der Wissenschaft zur Lösung der politischen, ökonomischen und ökologischen Probleme ist gesunken, nicht zuletzt auch wegen der vielfältigen Lösungen, die in der Regel für ein Problem angeboten werden.
- Mit der Hochschulexpansion hat sich die Einbindung von Hochschulen in weitere soziale, politische, wirtschaftliche Zusammenhänge verstärkt, so daß die autonome, in sich abgeschlossene, nur nach ihren eigenen Kriterien und Werten funktionierende Gelehrtenrepublik einer zunehmend gesellschaftlich rechenschaftspflichtigen Institution gewichen ist.
- Finanzielle Mittel stehen in den öffentlichen Kassen nicht mehr unbegrenzt zur Verfügung.

Vor diesem Hintergrund fragt die Gesellschaft bzw. der Staat in allen industrialisierten Ländern zunehmend nach den nachweisbaren Relationen zwischen den in die Hochschulen fließenden (Finanz-)Mitteln einerseits und den damit erzielten Ergebnissen andererseits.

Im Zusammenhang mit der allgemeinen Forderung nach Rechenschaft wird jedoch auch die Hoffnung nach mehr Transparenz für individuelle Entscheidungen verbunden. Diese Entscheidungen können sich auf die Wahl des Studienortes durch Abiturienten, die Auswahl von Bewerbern durch die Wirtschaft oder aber auch auf die Mittelzuweisung an die Hochschule durch den Staat beziehen.

2.3 Mittelverteilung

In Umkehrung des Wunsches zu wissen, was mit dem in die Hochschulen fließenden Geld geschieht (value for money), wird immer stärker gefordert, das Geld dahin fließen zu lassen, wo hochwertige Leistungen erzeugt werden. Die leistungsbezogene Finanzierung der Hochschulen ist ein grundlegender Paradigmenwechsel in der Steuerung der Hochschulen bzw. im Verhältnis Staat - Hochschule: Gerade auch in der Mittelzuweisung, die sich volumenmäßig am stärksten im Stellenhaushalt niederschlägt, ist die bisherige Steuerung eine ex-ante-Steuerung, die das Angebot an Studienplätzen bzw. Forschungskapazitäten festlegt, ohne direkten Bezug zur Nachfrage. Eine volumen- und leistungsorientierte Finanzierung der Hochschulen setzt nunmehr an der Nachfrage und an den erbrachten Leistungen an und steuert die Hochschulen damit ex-post. Eine derartige ex-post-Steuerung in der Mittelverteilung erfordert Evaluationen der Leistungen.

3 Evaluationsverfahren als Instrumente der Qualitätssicherung, Rechenschaftspflicht und Mittelverteilung: Versuch einer Klassifizierung

3.1 Methoden: quantitativ vs. qualitativ

In methodischer Hinsicht sind quantitativ und qualitativ orientierte Verfahren zu unterscheiden. Die **quantitative Methode** stützt sich auf Leistungsindikatoren (performance indicators) und versucht, über intersubjektiv überprüfbare Kennzahlen Aspekte des Inputs (Studentenzahlen), des Throughputs (Zwischenprüfungen, Dropout-Quote) und des Outputs (Absolventenzahlen, Studiensemester, Promotionen, Publikationen) statistisch zu erfassen.

Dabei sind insbesondere zwei Probleme zu berücksichtigen, 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. Ein Problem der zuverlässigen Messung dieses Indikators ist dagegen, welche Drittmittel sollen von wem in Deutschland wie erhoben werden. Dabei handelt es sich um die Frage der Reliabilität der Messung. Hier gibt es 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.

Dagegen werden bei **qualitativen Bewertungen**, die in der Regel über Peer Reviews erfolgen, 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. Ein Grundsatz muß daher sein, beide Verfahren aufeinander abgestimmt einzusetzen.

3.2 Typen: Output- vs. Prozeßorientierung

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

- Als **Output-Evaluation**: Sie ist in der Regel sehr detailliert, aufwendig aber tendenziell gerecht im Urteil, damit kosten- und zeitintensiv wie es augenblicklich in Großbritannien sowohl im Bereich der Forschung wie der Lehre geschieht. Weil sie detailliert auf die Forschungs- und Lehrergebnisse abhebt, erfährt sie größere Widerstände bei den Beteiligten und dient in erster Linie der Rechenschaftspflicht. Zugleich 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. Fragwürdig ist die Output-Evaluation im Hinblick auf die qualitätsverbessernden Wirkungen nicht zuletzt aufgrund der geringen Akzeptanz bei den Wissenschaftlern.
- Als **Prozeß-Evaluation**: Sie ist sehr grob, ungenau und eher ungerecht, damit wenig kosten- und zeitintensiv. Sie zielt auf die qualitätssichernden Maßnahmen und nicht auf die Qualität an sich, verzichtet auf eine detaillierte Analyse zugunsten einer häufigeren auf

die Veränderungen ausgerichteten Evaluation. Sie scheint zumindest in Australien sehr erfolgreich zu sein, da sie in Verbindung mit positiven Anreizen in Form von Mittelzuweisungen eine Vielzahl von Organisationsentwicklungsprozessen zur Qualitätsverbesserung in den Hochschulen bewirkt hat.

Es wird offensichtlich, daß man die Evaluationsverfahren sehr viel stärker im Hinblick auf die Zielsetzungen und auch die Kosten hinterfragen muß. Wenn es um Rechenschaft geht, ist bei der Komplexität des Gegenstandes eine aufwendige, detaillierte und genaue Evaluation notwendig. Dies ist insbesondere dann der Fall, wenn auch Mittel „gerecht“ verteilt werden sollen.

Wenn es allerdings um die Veränderung im Sinne einer Qualitätsverbesserung sowie um den Anstoß von Maßnahmen zur Organisationsentwicklung innerhalb der Hochschule geht, reicht eine kostensparende grobe Analyse aus, die den Anstoß zu Veränderungsmaßnahmen gibt. Dabei kommt es auf die Validität der Ergebnisse weniger an. Dies gilt insbesondere dann, wenn keine Mittel verteilt oder gar gekürzt werden. Positive Anreize in Form von zusätzlichen (!) Mitteln können den Prozeß der Qualitätsverbesserung erheblich beschleunigen, wie sich am australischen Beispiel zeigt. Dies wäre auch innerhalb der Hochschulen realisierbar durch zusätzliche Mittelzuweisungen an Fachbereiche mit ausgeprägtem Qualitätsbewußtsein.

3.3 „Gutachter“ bei Evaluationen

So unterschiedlich Evaluationen hinsichtlich ihrer Funktion, Methodik und Zielsetzung sein können, so unterschiedlich sind auch die möglichen „Gutachter,“ die eine Bewertung vornehmen. „Gutachter“ können - je nach „stakeholder“-Interesse - sein:

- Studierende
- Absolventen
- Arbeitsmarkt
- Öffentlichkeit/Journalisten
- externe Wissenschaftler (Peers)
- Unternehmensberater
- Wissenschaftsforscher.

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

4 Evaluationen in Deutschland: Verfahren, Probleme, Ergebnisse

Nach diesen eher allgemeineren Überlegungen und Unterscheidungen möchte ich nunmehr auf konkrete Evaluationsmaßnahmen und -projekte aus der deutschen Hochschullandschaft eingehen, um die unterschiedlichen Dimensionen, Zielsetzungen und Verfahren anhand konkreter Beispiele zu verdeutlichen.

4.1 Begutachtung der Deutschen Forschungsgemeinschaft (DFG)

Evaluationen der DFG folgen dem Modell der ex-ante „Peer Review“; sie ist die Basis für Mittelzuweisungen. Sofern überhaupt eine Ergebnisbewertung nach Beendigung des Projekts erfolgt, hat diese in der Regel keine Konsequenzen für das laufende Projekt, etwa durch Rückforderung bewilligter Mittel. Allerdings geht diese ex-post-Bewertung ein in eine gegebenenfalls neue ex-ante-Bewertung für ein weiteres Projekt.

Nun ist angesichts der eingangs skizzierten Veränderungen im Wissenschafts- und Forschungsbetrieb - aber auch angesichts der Belastung der Fachgutachter durch eine stetig steigende Zahl von Forschungsanträgen - die Frage durchaus berechtigt, wie es mit der Leistungsfähigkeit dieser Form des Peer-Review durch Fachgutachter steht und inwieweit künftig Modifikationen wünschenswert oder erforderlich wären. Denn die Gefahr, daß durch Peer Reviews disziplinäre Strukturen und Forschungsansätze immer wieder reproduziert werden, innovative Ansätze, die jenseits von disziplinären Grenzen und in trans-disziplinären Kontexten entstehen, dagegen unberücksichtigt und ohne Förderung bleiben, ist nicht einfach von der Hand zu weisen. Zu fragen ist darüber hinaus aber auch, inwieweit die ausschließliche ex-ante-Qualitätssicherung durch Fachgutachter durch ein kontinuierliches Audit und im Rahmen aufeinander abgestimmter externer Reviews und Evaluationen ergänzt werden müßte. Dafür spricht, daß inzwischen

- ganz erhebliche Summen z.T. über Jahre hinweg an Projekte und Forschungsstrukturen (z.B. SFBs) gebunden sind, was bei knappen Mitteln und möglicherweise rasch sich veränderten Forschungsprioritäten zu Engpässen in der Förderung führen kann;
- vor diesem Hintergrund eine einmalige ex-ante-Evaluation als alleiniges Element der Qualitätssicherung u.U. nicht mehr ausreicht;
- und schließlich der Druck zur Rechenschaftslegung über den Mitteleinsatz deutlich gestiegen ist.

4.2 Lehrberichte als Form der Selbstevaluation

Die Selbstevaluation ist das bisher tragende Steuerungsinstrument innerhalb der 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 verlorengegangen. Darüber hinaus ist diese Selbstkontrolle bei den Professoren durch die Gruppenuniversität mit ihren knappen Mehrheitsverhältnissen außer Kraft gesetzt worden, weil man sich im Zweifel auch um die schwarzen Schafe der Zunft scharen mußte, um Abstimmungs-niederlagen zu vermeiden.

Aus staatlicher Sicht sind Lehrberichte ein Versuch, von außen die Effizienz und Effektivität des Hochschulsystems mit einem aus dem Inneren der Hochschule herauskommenden Kontrollinstrument zu sichern. Sie sollen das instrumentelle Rückgrat der Evaluationsprojekte und anderer Aktivitäten zur Qualitätssicherung an den Hochschulen bilden. Von ihnen wird mehr Transparenz im Bereich des Lehr- und Studienbetriebs der Hochschulen, Fakultäten und Fächer erwartet. Ihr systemi-

scher 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 allerdings nicht nur bei diesem Instrumentarium, daß Transparenz allein keine kontrollierende Wirkung hat. Drei Gründe dafür seien genannt:

- Die inhaltliche Zielsetzung der bisherigen Lehrberichte ist letztlich ebenso diffus wie ihr Adressatenkreis (Ministerium, Öffentlichkeit, zukünftige/jetzig Studierende?)
- Zuviel Information kann ebenso viel Intransparenz hervorrufen wie keine Information.
- Staatlich angeordnete Lehrberichte haben sich immer dem Verdacht zu erwehren, daß sie als Instrument der Mittelkürzung bzw. der Ressourcenumleitung und -umwidmung herangezogen werden, was sich direkt auf die Freimütigkeit der Berichterstattung seitens der Hochschulen auswirken kann.

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.

4.3 Selbstbericht und Peer Review nach dem niederländischen Modell

Auf diesen Überlegungen aufbauend, wird neben dem Selbstbericht auch in Deutschland immer häufiger auf die externe Begutachtung durch Peers nach dem niederländischen Modell zurückgegriffen: Anhand eines vorher fixierten Fragenkatalogs werden Selbstberichte der Fachbereiche erstellt, die den externen Gutachtern als Grundlage für Fragen bei ihren Besuchen dienen, aus denen dann anschließend ein Fremdbericht erstellt wird.

Dieses Evaluationsmodell bringt durchaus Probleme mit sich:

- Fachbereiche haben erhebliche Schwierigkeiten, die erforderlichen Daten zu beschaffen und aufzubereiten.
- Es fällt ihnen außerordentlich schwer, fachbereichsbezogene Ziele zu formulieren. In der Regel kommt es im Verlauf des Verfahrens - häufig erstmals seit einer Vielzahl von Jahren - zu einer Zieldiskussion im Fachbereich, die oftmals auch nicht annähernd befriedigend abgeschlossen werden kann.
- Bei den Fremdberichten stellt sich in erster Linie das Problem der Folgerungen oder Konsequenzen, also der Umsetzung von Evaluationsergebnissen.

Dem Problem des Versandens von Evaluationsergebnissen stellen sich explizit die Verfahrensvarianten im Nordverbund, das Modell der Universität Dortmund sowie das Verfahren, das vom CHE Centrum für Hochschulentwicklung in Zusammenarbeit mit der Wirtschaftswissenschaftlichen Fakultät der Universität Münster entwickelt und angewandt wurde.

- **Nordverbund**

In bewußter Absetzung von der staatlichen Seite hat sich ein Verbund norddeutscher Universitäten gebildet, der die Evaluationen staatsfern und autonom organisiert hat. Das Ziel der Universitäten Bremen, Hamburg, Kiel, Oldenburg und Rostock liegt darin, 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 niederländischem Modell, verknüpft mit einer auswertenden Konferenz aller Beteiligten. Wesentliches Element ist die Freiwilligkeit sowohl der Universität wie der Fachbereiche.

Aus den Evaluationen im Nordverbund konnten zwei wichtige Erfahrungen zur Umsetzung der Gutachterempfehlungen gewonnen werden. Erstens erscheint es nützlich, eine auswertende Konferenz zusammen mit den Mitgliedern der Fachbereiche und den Peers durchzuführen, die auf Erfahrungsaustausch und Maßnahmen ausgerichtet ist. Zweitens muß von den Hochschulleitungen eine Verantwortung im Hinblick auf die Umsetzung der Empfehlungen übernommen werden, die bis hin zu schriftlichen Vereinbarungen zwischen Hochschulleitungen und Fachbereichen geführt hat.

- **Universität Dortmund**

Im Zuge der Einführung eines neuen Modells für die hochschulinterne Mittelverteilung, das auf quantitativen Leistungsparametern basierte (z.B. Drittmittelvolumen, Anzahl der Prüfungen) wurde an der Universität Dortmund nach einer Entsprechung für die qualitative Beurteilung gesucht. Ausgehend vom Akademischen Senat wurde daher ein Verfahren der Evaluation der Fachbereiche eingeführt, das ebenfalls das niederländische Verfahren mit Selbstbericht und Peer Review adaptiert, allerdings einige Besonderheiten aufweist.

1. Es handelt es sich um ein rein internes Verfahren, das lediglich durch jeweils zwei externe Fach-Gutachter angereichert wird. Die übrigen drei Gutachter sind emeritierte Hochschul-lehrer der Universität Dortmund, die aufgrund ihrer internationalen Reputation und der im bisherigen akademischen Leben an der Universität Dortmund bewiesenen Objektivität und Unabhängigkeit ausgewählt wurden.
2. Im Gegensatz zu den meisten Verfahren in Deutschland werden Forschung, Lehre und die Organisation des Fachbereichs zusammen beurteilt. Alle übrigen Evaluationen nach dem niederländischen Modell in Deutschland konzentrieren sich auf die Bewertung der Lehre.
3. Da es sich um ein rein hochschulinternes Verfahren handelt, werden auch nur unterschiedliche Fachbereiche gleichzeitig bewertet, so daß keine Fachvergleiche hergestellt werden können.
4. Das gesamte Evaluationsverfahren wird begleitet von einer Arbeitsgruppe "Review-Verfahren", geleitet durch einen Prorektor, die einerseits den Ablauf metaevaluiert, zum anderen für die Schlußfolgerungen aus den Evaluationen zuständig ist. Damit soll gesichert werden, daß Erkenntnisse der Evaluation nicht versanden.

- **Wirtschaftswissenschaftliche Fakultät der Universität Münster**

Die Initiative für dieses Evaluationsprojekt ging von den Betroffenen und Evaluierten selbst aus, die sich von einer Evaluation Orientierungshilfen für die Identifizierung von Stärken und Schwächen

ihrer Fakultät erwarteten. Entscheidend dabei war, daß die Evaluation ein wesentliches Element des strategischen Planungsprozesses der Fakultät darstellte, die sich im Vorfeld der Begutachtung intensiv über ihre künftige strategische Orientierung verständigt und ein neues Leitbild - aufbauend auf bestimmten Kernkompetenzen und Erfolgspotentialen - entwickelt hatte.

Die Evaluation stand somit in einem engen Zusammenhang mit einem umfassenden, fakultätseigenen Organisationsentwicklungsprozeß. Für die Umsetzung der Befunde, die von einer fünfköpfigen, international besetzten Gutachtergruppe im Rahmen einer Peer Review erzielt und der Fakultät übermittelt wurden, bestehen sehr günstige Voraussetzungen, da das gesamte Projekt vom Veränderungswillen der Fakultät selbst getragen wurde.

Aber auch die Hochschulleitung wurde als Partner in das Projekt einbezogen. Ihr Interesse lag in der Erprobung eines Instrumentariums, das auch für andere Fakultäten innerhalb der Universität geeignet ist. Darüber hinaus ist beabsichtigt, daß Fakultät und Hochschulleitung zu bestimmten Fragen Zielvereinbarungen abschließen, durch die eine für beide Seiten gültige Verbindlichkeit für die Umsetzung der Gutachterempfehlungen hergestellt werden kann.

4.4 Befragungen von Studierenden und Absolventen

Befragungen werden isoliert oder auch in Zusammenhang mit dem Verfahren des Selbstberichts mit Peer Review oder auch bei den Rankings eingesetzt. Im wesentlichen konzentriert man sich auf zwei Zielgruppen:

- **Hörerbefragungen/Lehrveranstaltungskritiken**

Bis auf einige Ausnahmen an der TU Berlin, der Ruhr-Universität Bochum und Universität Mannheim sind sie durch die Dozenten selbst induziert und werden von diesen - als individuelles feedback - ausgewertet und interpretiert. Darüber hinaus werden an etlichen Fachbereichen von studentischer Seite Hochschullehrer beurteilt und gelegentlich negativ herausgestellt. Die Wirkungen derartiger studentischer Veranstaltungskritik sind durchaus zweifelhaft, auch wenn sie in Abhängigkeit von der individuellen Psyche des Hochschullehrers vielleicht hier und da zu Verhaltensänderungen führen.

- **Absolventenbefragungen**

Neben Studentenbefragungen werden gelegentlich auch Absolventenbefragungen durchgeführt. Aus den rückblickenden Beurteilungen der Absolventen können durchaus Rückschlüsse auf die Arbeitsmarktorientierung der Ausbildung gezogen werden. Wie bei den Studentenbefragungen werden derartige Befragungen wohl häufiger durchgeführt als öffentlich bekannt ist. Die Schlußfolgerungen, die aus den Ergebnissen gezogen werden, sind jedoch über den jeweiligen Fachbereich hinaus weitgehend unbekannt. Generell kann aber gesagt werden, daß sowohl Hörerbefragungen als auch Absolventenbefragungen, wenn sie denn systematisch durchgeführt werden, wichtige Elemente eines Qualitätssicherungssystems darstellen können. Voraussetzung dafür ist allerdings, daß die Ergebnisse seitens der Betroffenen ernstgenommen werden und in Maßnahmen zur Veränderung und Verbesserung einmünden.

4.5 Bibliometrische Messungen

Ein Instrument der quantitativen Evaluation im Forschungsbereich stellen bibliometrische Messungen dar. Ihr Ziel ist es, wissenschaftliche Leistungen mit Hilfe von Anzahl und Niveau der Publikationen sowie dem internationalen Einfluß dieser Veröffentlichungen zu erfassen. Sofern die entsprechenden Indikatoren angemessen gehandhabt und wissenschaftlich aufbereitet werden, sind sie in der Lage, wesentliche Einschätzungen über den nationalen oder internationalen Stand der Forschung, insbesondere von Fachbereichen oder Instituten, zu geben. Allerdings ist ihre Anwendbarkeit von fächerspezifischen Publikationskulturen und Zitationspraktiken abhängig. In Deutschland sind bisher lediglich naturwissenschaftliche Institute analysiert worden.

4.6 Rankings

Das Problematische an den meisten der bisher durchgeführten Rankings ist, daß versucht wird, verschiedene Dimensionen von Qualität in eine einzige Gesamtnote zusammenzufassen, die dann die Qualität einer gesamten Hochschule oder einzelner Fachbereiche wiedergeben soll. Qualität ist aber mehrdimensional und zudem abhängig von der Perspektive und dem Interesse des jeweiligen Betrachters. *Die* beste Hochschule oder *den* besten Fachbereich kann es daher nicht geben, lediglich die beste Hochschule *für* ein bestimmtes Interesse, eine bestimmte Neigung oder einen bestimmten Bedarf.

Vor diesem Hintergrund will der „Vergleichende Studienführer,“ den das CHE gemeinsam mit der Stiftung Warentest erarbeitet und vor kurzem veröffentlicht hat (1998), eine differenzierte Gegenüberstellung von Studienmöglichkeiten und -bedingungen an unterschiedlichen Hochschulstandorten ermöglichen, zunächst für die Studienbereiche Chemie und Wirtschaftswissenschaften. Er ist gedacht als Orientierungshilfe bei der Wahl einer passenden Hochschule und enthält daher systematische Informationen über Studienfächer, Fachbereiche, Hochschulen und deren Standorte. Studienbedingungen und Lehrleistungen, einschließlich der für die Lehre relevanten Forschung werden in einem Vergleich nach verschiedenen Kriterien wissenschaftsgerecht abgebildet.

4.7 Benchmarking

Der systematisch gegenübergestellte Vergleich von Strukturen oder Prozessen soll im Benchmarking zu Qualitätsverbesserungen oder Effizienzgewinnen führen. Seit 1996 besteht ein Benchmarking Club Technischer Universitäten, moderiert durch das CHE, in dem der strikt vertrauliche Austausch von Informationen zur Verbesserung der Entscheidungsqualität von Hochschulleitungen betrieben wird. Für einen ersten Zyklus wurden zwei Untersuchungsbereiche ausgewählt: interne Verteilungsverfahren für „Laufende Mittel für Forschung und Lehre“ und ein „Strukturdatenvergleich ausgewählter Fächer“. Durch die strikte Vertraulichkeit ist die Reliabilität der Daten gewährleistet. Der Vergleich zeigt den Hochschulen ihre Stärken und Schwächen auf und kann in hochschulinterne Entscheidungen umgesetzt werden, auch wenn nicht unbedingt ein „best practice“ ermittelt werden kann.

4.8 Gutachten von Beratern

Evaluationen werden in erster Linie auf Lehre und Forschung bezogen. Notwendig ist es aber auch, die Managementprozesse der Hochschulen mit einzubeziehen, beispielsweise ihre Fähigkeit zur Entscheidungsfindung. Am Beispiel von Berufungen zeigt sich die Problematik, die keineswegs geringer ist als bei der Bewertung von Forschung und Lehre: Der Zeit-, Arbeits- und Gutachteraufwand 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. Die Problematik derartiger Gutachten besteht darin, daß sie von staatlicher Seite jeweils unmittelbar zur Mittelkürzung genutzt werden.

In bewußter Absetzung davon hat das CHE Centrum für Hochschulentwicklung bei einer kürzlich abgeschlossenen Evaluation der Hochschulverwaltung von drei thüringischen Hochschulen einen grundlegend anderen Weg gewählt, der die Einbindung der Mitarbeiter bei der prozeßorientierten Analyse und Neugestaltung von Verwaltungsaufgaben im Bereich „Studium und Lehre“ sowie „Ressourcenbewirtschaftung“ in den Mittelpunkt stellte. Damit sollte bewirkt werden, daß bereits während der Evaluation Organisationsentwicklungsprozesse zur Qualitätsverbesserung in der Leistungserstellung eingeleitet werden.

4.9 Modelle zur Mittelverteilung

Letztendlich stellen die Modelle zur Mittelverteilung, die zwischenzeitlich in einigen Bundesländern eingeführt wurden bzw. werden, ebenfalls implizit über die verwendeten Leistungsindikatoren Evaluationen dar.

Als generelle Erkenntnisse eines Vergleichs inländischer und ausländischer Modelle lassen sich hinsichtlich des Evaluationsaspekts ableiten:

1. Umfassende Modelle der Mittelverteilung sind ein „Mehrkomponentencocktail“ mit einem Volumen-, einem Leistungs- und einem Innovationsteil. Dies gilt für alle betrachteten Modelle. Die Volumen- und Leistungskomponenten werden i.d.R. über formelgebundene Zuweisungen realisiert. Der Innovationsanteil hingegen impliziert zumeist diskretionäre Zuweisungen, da gezielte Innovationsanreize über rückwärtsgewandte Indikatormessungen nur schwer zu vermitteln sind. Unterschiede bestehen sowohl im Mischungsverhältnis des Cocktails als auch darin, daß die finanziell honorierten Aufgaben und Leistungen der Hochschulen ganz verschiedene Dinge beinhalten können. Letzteres zeigt die Abhängigkeit der Modellgestaltung von den zugrundeliegenden politischen Zielen.
2. Es herrscht hohe Übereinstimmung in bezug auf die Anforderungen, die an im Rahmen der formelgebundenen Mittelzuweisung verwendete Indikatoren gestellt werden. Die wesentlichen Anforderungen sind:
 - Möglichkeit zur zeitnahen Erhebung,
 - geringer Erhebungsaufwand,
 - Vergleichbarkeit,
 - Nachvollziehbarkeit,
 - Transparenz.

3. Eine Abwägung dieser Anforderungen führt zu einem begrenzten Kanon an praktikablen Indikatoren, die in den betrachteten Modellen Anwendung finden. Unterschiede ergeben sich in der Operationalisierung der Indikatoren. Die gängigen Indikatoren für formelgebundene Mittelzuweisungen sind:
- Studierende,
 - Absolventen,
 - Promotionen,
 - Drittmittel,
 - bibliometrische Indikatoren,
 - Studienzeiten,
 - Personal.
4. Leistungsorientierte Mittelverteilungen sind ein Balance-Akt zwischen der Setzung von finanziellen Anreizen und der Bewahrung finanzieller Stabilität. Zu starke Instabilität bringt das System in Gefahr oder ist allenfalls dann tragbar, wenn die Hochschule finanzielle Schwankungen bei den staatlichen Geldern aus anderen Einnahmequellen kompensieren kann. Die erörterten Modelle wenden unterschiedliche Verfahren zur Stabilisierung an:
- die Glättung von Zufallsschwankungen durch gleitende Mehrjahresdurchschnitte der Indikatorwerte;
 - die Verwendung von stabilen, volumenbezogenen Indikatoren (z.B. Studienplatzkapazitäten beim Modell für die niedersächsischen Fachhochschulen);
 - die Integration „historischer“ Komponenten (z.B. Niederlande);
 - Sonderprogramme zur diskretionären Kompensation von Einbußen (z.B. Dänemark);
 - die Verwendung von Zielvereinbarungen über Indikator-Soll-Werte (z.B. Finnland).

Die bisherigen Ansätze in Deutschland (Nordrhein-Westfalen, Rheinland-Pfalz und Niedersachsen) zeigen im Verhältnis zu ausländischen Modellen einige systematische Unterschiede:

1. Die ausländischen Modelle sind sehr viel ganzheitlicher und umfassender angelegt. In Deutschland herrschen im Moment Ansätze vor, bei denen unter weitgehender Beibehaltung der traditionellen Verfahren der staatlichen Finanzierung lediglich in Teilbereichen eine formelgebundene Mittelzuweisung vorgenommen wird. So werden beispielsweise bestimmte Prozentsätze der Titelgruppe für Forschung und Lehre nach Formeln verteilt. D.h. das neue Verfahren wird nur auf die Mittel angewandt, die im Rahmen des gegenwärtigen Systems (u.a. unter Beibehaltung von Stellenplänen) entscheidungsverfügbar sind. Selbst das für deutsche Verhältnisse sehr umfassende Reformvorhaben des Landes Niedersachsen ist insofern partiell, als daß es Fachhochschulen und Universitäten (noch) nicht integriert. Partielle Modelle der nordrhein-westfälischen bzw. rheinland-pfälzischen Ausprägung sind stark leistungsbezogen (da die volumenbezogene Komponente über traditionelle Verfahren abgedeckt wird). Im Ausland stellen dagegen die neuen Zuweisungsmodelle eine systematische Ergänzung zum umfassenden Globalhaushalt dar. In einem solchen autonomen und anreizgesteuerten Gesamtmodell braucht man keine haushaltsrechtlichen Detailsteuerungen mehr. Zur leistungsbezogenen tritt die aufgaben-/volumenbezogene Finanzierung, die ebenfalls über Indikatoren realisierbar ist. Zudem enthalten umfassende Modelle immer auch sy-

stematische diskretionäre Elemente, die in Deutschland bisher nicht in Globalzuweisungen eingehen.

2. Im Vergleich zu den ausländischen Modellen läuft die Diskussion in Deutschland in Richtung höherer „Zielhaltigkeit“ der Modelle. So sollen mit den Modellen in Deutschland Studienzeitverkürzung ebenso wie Frauenförderung realisiert werden. In den ausländischen Formelmodellen spielen diese politisch gesetzten Ziele eine eher geringe Rolle, wobei hier nach anfänglich hohen Erwartungen seitens der Politik teilweise auch bereits bewußte Rücknahme der politischen Einflußnahme praktiziert wird.

5 Evaluationsverfahren, Adressaten und Funktionen

Die derzeit diskutierten und eingesetzten Instrumente der Qualitätssicherung dienen gleichzeitig der Kontrolle und Rechenschaft gegenüber der Gesellschaft ebenso wie der Mittelverteilung, also sehr unterschiedlichen Zwecken und Funktionen.

Die einzelnen Verfahren sind jedoch für die unterschiedlichen Funktionen und Zwecke sehr unterschiedlich geeignet. Von daher sollte sehr genau nach Funktion und Zielgruppe gefragt werden und mit großer Sensibilität die Aussagefähigkeit und Wirkungsmöglichkeit der jeweiligen Verfahren für die angestrebten Zielsetzungen beurteilt werden.

Für die im vorigen Abschnitt vorgestellten, in Deutschland augenblicklich eingesetzten Instrumente kann die Zuordnung wie folgt vorgenommen werden:

Evaluationsverfahren	Zielgruppe/Adressat	Funktion
Begutachtung DFG	Staat/DFG	Mittelverteilung
Lehrberichte	Hochschule	Qualitätssicherung
Selbstbericht und Peer Review	Hochschule	Qualitätssicherung / Strategie
Befragungen Studierende Absolventen	Öffentlichkeit/Hochschule	Leistungstransparenz / Qualitätssicherung/(interne) Mittelverteilung
Bibliometrische Messung	Öffentlichkeit/Hochschule	Leistungstransparenz / Mittelverteilung
Ranking	Öffentlichkeit/Hochschule	Rechenschaft / Leistungstransparenz / Qualitätssicherung
Benchmarking	Hochschule	Qualitätssicherung
Gutachten von Beratern	Staat/Hochschule	Mittelverteilung (-kürzung) / Qualitätssicherung
Modelle zur Mittelverteilung	Staat/Hochschule	Mittelverteilung, extern und intern