
Manfred Hennen, Hrsg.

**Evaluation –
Erfahrungen und Perspektiven**

Mainzer Beiträge zur Hochschulentwicklung, Bd. 4

Hrsg.: Zentrum für Qualitätssicherung und –entwicklung (ZQ)

Manfred Hennen, Hrsg.
Evaluation – Erfahrungen und Perspektiven

Mainzer Beiträge zur Hochschulentwicklung, Bd. 4
Hrsg.: Zentrum für Qualitätssicherung und –entwicklung (ZQ)

Mainz 2002

Nachdruck und Verwendung in elektronischen Systemen nur mit vorheriger schriftlicher
Genehmigung

ISSN: 1616-5799
ISBN: 3-935461-03-8

Inhalt

1. Katja AHLSTICH und Tanja URICH-NEITZERT : Grundlagen der Evaluation	1
2. Elisabeth SPRINGER: Evaluation nach dem „Mainzer Modell“	21
3. Uwe SCHMIDT : Ausgewählte Ergebnisse der Evaluation unter systemischer Perspektive	47
4. Manfred HENNEN: Die Zukunft der Evaluation	82
5. Ulrich DRUWE: Evaluation – Neues Steuerungsmodell – Wettbewerb	146
6. Harald PAULSEN und Jürgen MARKL Evaluation 2000 des Fachbereichs Biologie – ein Rückblick	158
Literaturverzeichnis	163

Vorwort

Mit diesem vierten Band der Reihe ‚Mainzer Beiträge zur Hochschulentwicklung‘ legen wir sechs Arbeiten vor, die sich des Themas der Evaluation annehmen. Sie beruhen alle auf Erfahrungen, die mit dem sogenannten ‚Mainzer Modell‘ der Hochschulevaluation gemacht worden sind. Unterschiedliche Perspektiven zu diesem Thema sollen zusammengeführt werden. Die Gemeinsamkeit des Themas bei unterschiedlichem Zugang der einzelnen Beiträge erlaubt nach Ansicht des Herausgebers von sechs Kapiteln zu sprechen, deren Anordnung einem Prinzip der Abfolge von Dokumentation, grundsätzlichen Erörterungen und Bewertung innerhalb der Hochschule folgt. Dies entspricht dem Anliegen der ‚Mainzer Beiträge zur Hochschulentwicklung‘, in denen sich Dokumentationen von Verfahren, theoretische Erörterungen und hochschulpolitische Einschätzungen abwechseln und ergänzen sollen. Band 1 und 3 belegen Evaluationsverfahren der Universität Mainz; Band 2 und 4 befassen sich mit der theoretischen und politischen Einordnung von Maßnahmen der Hochschulentwicklung.

Zu den einzelnen Beiträgen: Im ersten Kapitel ‚Grundlagen der Evaluation‘ skizzieren Katja Ahlstick und Tanja Urich-Neitzert die Entwicklung der Evaluation zu einem Instrument, das in Deutschland während des zurückliegenden Jahrzehntes in der Gestaltung der Hochschulen und anderer Institutionen einen Aufstieg ungewöhnlicher Art erfahren hat: Aus einer Waffe der hoheitlichen Aussonderung, die nur ausnahmsweise gegen fehlgeleitete oder leistungsschwache Organisationen eingesetzt worden ist, hat sich ein Verfahren zur Mobilisierung der selbstverantwortlichen Reorganisation u.a. von Fakultäten entwickelt. Eine derartige Karriere ruft Skepsis und Erklärungsbedarf hervor. Die Autorinnen verweisen auf einen verstärkten nationalen sowie internationalen Wettbewerb der Universitäten und auf die zunehmende Finanzknappheit der Länder. Sie schildern den Prozess zunehmender Professionalisierung der Evaluationsverfahren wie er sich in der wachsenden Zahl von Evaluationsagenturen und Evaluationsnetzwerken niederschlägt. Begriffen wie Qualitätssicherung und Rechenschaftspflicht gegenüber der Öffentlichkeit wird eine besondere Bedeutung zugewiesen.

Dieses Kapitel liefert einen Überblick über verschiedene Ansätze innerhalb der Evaluationsverfahren und deren bisherige Entwicklung an den deutschen Hochschulen. Es ermöglicht einer Leserschaft, die im tagtäglichen Geschäft mit der Problematik der Evaluation im Einzelnen nicht vertraut sein kann, einen Einstieg in die Thematik des Bandes.

Im zweiten Kapitel stellt Elisabeth Springer das Mainzer Modell der Evaluation vor. Das an der Johannes Gutenberg-Universität entwickelte Verfahren hat Entwicklungen und Modifikationen vor allem durch den jahrelangen Austausch zwischen den Mitgliedern und Fachvertretern der Fachbereiche und wissenschaftlichen Einrichtungen einerseits und dem vom Senat eingerichteten Zentrum für Qualitätssicherung und –entwicklung (ZQ) andererseits erfahren. Der Artikel betont die Prioritäten und Schwerpunkte des Mainzer Modells.

Einer dieser Schwerpunkte liegt in der Aufhebung einer scheinbaren Unvereinbarkeit zweier Seiten ein und derselben Philosophie: Einerseits sind die Fächer prinzipiell für das Verfahren der Evaluation zuständig, andererseits ist die Begleitung des internen Evaluationsprozesses im Vergleich zu anderen Agenturen sehr beratungsintensiv und beruht auf Erörterungen von Grundsatzfragen in multiplen Gesprächsrunden. Die eigentliche Evaluationsarbeit beginnt nach unserer Auffassung jenseits einer standardisierten Erhebungs- und Befragungstechnik. Das soll in dem Selbstverständnis von ZQ als einer intermediären und treuhänderisch verpflichteten Instanz zum Ausdruck kommen. Die wirklich heiklen Fragen beim Geschäft der organisationalen Bestandsaufnahme und Reorganisation vertragen sich nicht mit einer exklusiven anonymen Mechanik der Erhebung; es geht um die Gewinnung oder Erzeugung von Urteilen und deren Konfrontation zum Zwecke der Entscheidung. Es liegt in dieser Logik, dass die Autorin auch nach den Grenzen des Verfahrens fragt.

Uwe Schmidt macht in Kapitel 3 unter dem Titel ‚Ausgewählte Ergebnisse der Evaluation unter systemischer Perspektive‘ den bisher in dieser Form noch nicht unternommenen Versuch einer gegliederten Zusammenfassung der konkurrierenden Ansprüche und Voraussetzungen, denen Hochschulen und vergleichbare Systeme ausgesetzt sind. Diese Interdependenzen können nicht durch unterschiedliche und jeweils einseitige Krisenszenarien abgebildet werden. Die Hochschuldispute leiden

chronisch an der Einseitigkeit der hervorgehobenen Einzelaspekte unter Vernachlässigung ihrer Wechselbeziehung. Der Diagnose der Unterausstattung ist beispielsweise nicht mit dem Befund von Defiziten in der Lehre als abschließende Folgerung zu begegnen. Schmidts Bemühungen gelten deshalb einem theoretischen Bezugssystem, das erlaubt, sich ein Bild sowohl von der Vielfalt der Evaluationsergebnisse zu machen als auch von der Notwendigkeit ihrer Verortung in einem austarierten Zusammenhang. Hochschulentwicklung verlangt Augenmaß für die Gleichzeitigkeit gegensätzlicher Unabdingbarkeiten. Die Bewältigung von Komplexität besteht zunächst im Erkennen eines Zusammenhangs oder Gleichgewichts in der Vielfalt. Schmidts Beitrag engagiert sich in der Verknüpfung von theoretischen Ausführungen mit praktischen Evaluationserfahrungen an der hiesigen Universität. Um mit kontrastierenden hochschulpolitischen Positionen umgehen zu können, wird theoretisches Werkzeug wie in diesem Fall ein sozialwissenschaftlicher Systembegriff benötigt.

Meine eigenen Ausführungen im 4. Kapitel über die Zukunft der Evaluation schließen Überlegungen zur erwähnten Karriere des Verfahrens und der dafür verantwortlichen Hintergründe ein. Evaluation ist dort zu einem subsidiären Steuerungsinstrument geworden, wo der Schutz von funktional notwendiger Autonomie in Koordinations- und Verantwortungsblockaden umzuschlagen pflegt. Diese Sichtweise erfordert eine Neubestimmung der Definition von Evaluation und Qualität. Ich wende mich gemäß der zuvor angedeuteten Grenzen von Standardverfahren der Befragung gegen eine verlockende, aber trügerische Hoffnung auf Technisierung von Verständigungs- und Entscheidungsprozessen. Meine Transaktionsanalyse hebt hervor, in welchem Ausmaß wir in komplexen Bereichen der Handlungssynchronisation in Organisationen auf Urteilssynthese angewiesen sind. Evaluation wird entsprechend dieser Einschätzung auf eine Annäherung an Organisationsentwicklung sowie auf Organisationsberatung auch in Teilbereichen hinauslaufen. Ich versuche mich in einer Darstellung von synthetisch erzeugter und verfahrensgestützter Dialogfähigkeit unter Bedingungen der genannten Verständigungsblockade. Dabei rücke ich die Vorstellung einer Zentrierung auf natürliche Sprache als Medium der Urteils- und Entscheidungsfassung in den Vordergrund.

Kapitel 5 steuert einen Beitrag von Harald Paulsen und Jürgen Markl mit dem Titel ‚Evaluation 2000 des FB Biologie – ein Rückblick‘ bei. Die Verfasser beurteilen und kommentieren die vor einem Jahr abgeschlossene Evaluation des Fachbereiches Biologie, die sie als verantwortliche Funktionsträger und Professoren mit großem Einsatz und viel Verantwortungsbewusstsein begleitet haben. Aus der zeitlichen Distanz stellen sie fest, dass die Evaluation einerseits fachbereichsintern neue Kräfte freigesetzte und ins Stocken geratenen Diskussionen neue Impulse versetzte, zum anderen aber auch dem Fachbereich nach außen hin zu einem angemesseneren und positiveren Bild verhalf. Die kritische Beschreibung der Evaluation im Fachbereich Biologie ist einer Kooperation mit den Fächern zum Zweck der Überarbeitung sowie der Verankerung des Verfahrens im Vollzug der Wissenschaft nach dem Mainzer Modell äußerst dienlich.

Der letzte Beitrag soll die thematischen Erörterungen abrunden. Vizepräsident Ulrich Druwe äußert sich zum Thema: ‚Evaluation – Neues Steuerungsmodell – Wettbewerb‘. Er bettet das Thema der Evaluation nach dem Mainzer Modell in die aktuellen und die zu erwartenden Entwicklungen der Universität Mainz und der deutschen und europäischen Hochschulsituation ein. Vor allem seit der europolitischen Bologna-Erklärung von 1999, die auf Reorganisation und vor allem auf eine Stärkung institutioneller Autonomie und Selbstverantwortung der Hochschulen zielt, erhält die zentrale These Druwes Bedeutung, dass Wettbewerb ein zentraler Faktor in der weiteren Entwicklung der Hochschulen sein wird. In diesem Wettbewerb wird Evaluation nach seiner Prognose eine zunehmend wichtige Rolle spielen.

Als Herausgeber dieses Bandes danke ich den Autorinnen und Autoren für ihre Beiträge und für die den Hauptverpflichtungen und Hauptneigungen ihres Arbeitslebens abgerungene Zeit. Für ZQ ist vorliegende Systematisierung ein Gewinn, weil wir aus der Erfahrung mit Evaluation wissen, dass nichts praktisch wirksamer ist als eine erfahrungsgestützte und theoretisch fundierte Vergewisserung.

Frau Dr. Weiß hat die Bearbeitung der Drucklegung übernommen. Erst diese anspruchsvolle Arbeit bringt Unstimmigkeiten der Gedanken und

die Notwendigkeit der Korrektur zu Tage. Dafür sei ihr kollegial und herzlich gedankt.

Manfred Hennen im Januar 2002

Grundlagen der Evaluation

Katja Ahlstick und Tanja Urich-Neitzert

1. DEFINITIONEN UND ZIELE DER EVALUATION.....	2
2. BEDEUTUNG UND ENTWICKLUNG DER EVALUATION	4
3. BEDEUTUNG UND ENTWICKLUNG DER EVALUATION AN DER HOCHSCHULE.....	9
4. ZIELE UND FUNKTIONEN DER EVALUATION AN DER HOCHSCHULE.....	11
5. EVALUATIONSVERFAHREN AN DEUTSCHEN HOCH- SCHULEN.....	15

1. Definitionen und Ziele der Evaluation

In den letzten 40 Jahren hat sich ein sozialwissenschaftliches Arbeitsfeld etabliert, das sich wie keine andere Disziplin um die systematische Anwendung sozialwissenschaftlichen Wissens auf die Entwicklung, Implementierung und Bewertung von Maßnahmen zur Lösung gesellschaftlicher Probleme bemüht und als **Evaluation** oder zum Teil auch als empirische **Evaluationsforschung** bezeichnet wird. Wir entscheiden uns hier für die Bezeichnung Evaluation im Sinne eines Konzeptes der ‚wissenschaftlich gestützten Evaluation‘, um von der davon abzugrenzenden Evaluationsforschung nur dann zu sprechen, wenn diese einer quasi-experimentellen Forschungslogik folgt. Evaluationsforschung untersucht die tatsächlichen Effekte von Interventionsprogrammen, die zur Lösung definierter gesellschaftlicher Probleme implementiert wurden. Zu diesem Zweck folgt sie einer Theorie und verbietet während der Zeit des Messvorganges jede Form der Intervention um die kausale Zurechnung nicht zu stören. In diesem Fall könnte man deshalb berechtigter Weise auch von ‚Programmwirkungstheorien‘ sprechen.¹

Evaluation ist ein Teilbereich der empirischen Forschung. Sie ist die Anwendungsvariante wissenschaftlicher Forschungsmethoden auf eine spezielle Gruppe von Fragestellungen. Darüber hinaus ist sie Teil eines komplexen Entscheidungsprozesses und beinhaltet Management und politische Tätigkeit. Sie ist somit eine dynamische Kraft der öffentlichen Verwaltung und Politik.

Das Konzept der Evaluation stammt aus dem angloamerikanischen Sprachraum. Es gibt weder eine allgemein anerkannte Definition von Evaluation noch gibt es eine allgemein gültige Theorie, die die verschiedenen Konzepte, Überlegungen und Lösungsversuche vollständig erfasst. Dies liegt u.a. darin begründet, dass die Evaluation ein interdisziplinärer Forschungszweig ist und die Teildisziplinen unverbunden nebeneinander ihre Forschungen betreiben. Im wissenschaftlichen Sprachgebrauch wird Evaluation deshalb mit einer Vielzahl von verschiedenartigen sozialwissenschaftlichen Aktivitäten und Konzepten beschrieben wie z.B. mit Erfolgskontrolle, Effizienzforschung, Begleitforschung, Bewährungskontrolle, Bewertungsforschung, Wirkungsforschung, Wirkungskontrolle, Qualitätskontrolle, Qualitätssicherung oder Kosten-Nutzen-Analyse. Sie werden z.T. synonym im Sinne einer spezialisierten Form von Evaluation verwendet.

¹ Vgl. hierzu den Beitrag von Manfred Hennen in diesem Band, 91.

Trotz aller Unterschiede gibt es gewisse Gemeinsamkeiten: Weitgehende Einigkeit besteht in der Unterscheidung zwischen Evaluation und Evaluationsforschung. *Ganz allgemein betrachtet, bedeutet Evaluation den Prozess der Beurteilung des Wertes eines Produkts, Prozesses oder eines Programms.*² Solche Bewertungen finden häufig ohne systematische Verfahren oder datengestützte Beweise der Folgerichtigkeit statt. *Evaluationsforschung meint dagegen „die systematische Anwendung sozialwissenschaftlicher Forschungsmethoden zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme“.*³ Sie bestimmt folglich das Ausmaß, in dem z.B. ein Programm Ziele erreicht, und beschreibt Ursachen, warum Ziele oder Teilziele nicht erreicht wurden. Die allgemeinen Kennzeichen sozialwissenschaftlicher Evaluation in der bestehenden Definitionsvielfalt stellen Wottawa und Thierau⁴ zusammen:

- Alle mit Evaluation zusammenhängenden Tätigkeiten haben etwas mit der Bewertung von Handlungsalternativen zu tun.
- Evaluation ist ziel- und zweckorientiert. Sie hat primär das Ziel, praktische Maßnahmen zu überprüfen, zu verbessern oder über sie zu entscheiden.
- Evaluationsmaßnahmen werden dem aktuellen Stand wissenschaftlicher Techniken und Forschungsmethoden angepasst.

Es ist Aufgabe der Evaluation Antwort auf z.B. folgende Fragen zu finden:

- Worin bestehen die Probleme und wie groß sind sie?
- Welche Maßnahmen kann man ergreifen, um die Probleme zu lösen?

² Vgl. hierzu Jürgen BENGEL, Uwe KOCH (1988): *Evaluationsforschung im Gesundheitswesen*, 323; Heinrich WOTTAWA, Heike THIERAU (1998): *Lehrbuch Evaluation*, 13 sowie Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 14.

³ Peter H. ROSSI, Howard E. FREEMAN, Gerhard HOFFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*, 3. Vgl. auch Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 14; Jürgen BENGEL, Uwe KOCH (1988): *Evaluation im Gesundheitswesen*, 323; Jürgen BORTZ, Nicola DÖRING (2002): *Forschungsmethoden und Evaluation für Sozialwissenschaftler*, 102 sowie Heinrich WOTTAWA, Heike THIERAU (1998): *Lehrbuch Evaluation*, 13.

⁴ Heinrich WOTTAWA, Heike THIERAU (1998): *Lehrbuch Evaluation*, 14.

- Welche Zielgruppe sollte für die Maßnahmen in Betracht gezogen werden?
- Erreicht das Programm die Zielgruppe?
- Wird die Maßnahme so implementiert wie geplant?
- Ist die Intervention effektiv?
- Wie ist die Kosten- und Nutzenrelation?

2. Bedeutung und Entwicklung der Evaluation

Der Begriff der Evaluation entstand in seiner methodischen Ausformulierung Anfang der 60er Jahre in den USA in Verbindung mit den Reformprogrammen unter den Präsidenten Kennedy und Johnson, von denen der Feldzug gegen die Armut wohl das bekannteste Programm darstellt. Mit den Programmen „War on Poverty“ und „Great Society“ wurde von Beginn an der Auftrag verbunden, die Wirkungen dieser Programme wissenschaftlich zu ermitteln. In den USA ist **Evaluation** institutionell verankert und fester Bestandteil der universitären, insbesondere der postgraduierten Ausbildung.⁵

Die Anfangszeit der Evaluation ist charakterisiert durch objektivistische Theorien. Sie legten besonderen Wert auf die Entdeckung von Programmeffekten und waren in wissenschaftlicher (experimenteller) Methodik und Logik verwurzelt. Die bedeutendsten Theoretiker dieser Zeit zwischen 1965 und 1975 waren Michael Scriven und Donald T. Campbell.

Für Scriven besteht Evaluation darin, wissenschaftlich fundierte Werturteile zu fällen. Die Aufgabe einer systematischen und rationalen Urteilsfindung über den Wert eines Evaluationsgegenstandes ergibt sich aus der gesellschaftlichen Notwendigkeit, möglichst vorurteilsfreie, objektive und über den Interessen der Betroffenen stehende Sichtweisen im Gegensatz zu den üblichen subjektbezogenen Beurteilungsverfahren der Praxis einzuführen. Die von Scriven (1967) vorgenommene Unterscheidung zwischen *formativer* und *summativer Evaluation* hat innerhalb der zahlreichen Modelle bis heute die größte Verbreitung gefunden: „Summative

⁵ Vgl. hierzu Peter H. ROSSI, Howard E. FREEMAN, Gerhard HOFFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*, 4ff. sowie Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 22f., 30.

Evaluationen werden geplant, um den vollen Umfang der Effekte zu erfassen, die zu einem Zeitpunkt in der Vergangenheit durch das Objekt der Evaluation hervorgerufen wurden. ... Formative Evaluationen haben eine in die Zukunft gerichtete Perspektive und trachten danach, die verbesserungswürdigen Bestandteile eines Programms zu identifizieren“.⁶ Sie beschäftigen sich mit Problemen der alltäglichen Programmgestaltung. Im Hinblick auf die einzelnen Stadien, die ein Programm von der Entwicklung über die Umsetzung bis hin zu seinen möglichen Wirkungen durchläuft, bezieht sich die formative Evaluation auf die Beurteilung der Programmkonzeption und Programmdurchführung, während die summative Evaluation vornehmlich die Bewertung der Programmwirkung zum Ziel hat. Nach Scriven ist die formative Evaluation „non-evaluativ“, da sie keine Urteile darüber abverlangt, ob das zu evaluierende Programm gut oder schlecht ist. Er entwickelt ein Konzept der summativen Evaluation, nach der die Bewertung eines Programms in einem vierstufigen Prozess erfolgt: In einem ersten Schritt müssen Wertkriterien angegeben werden, in einem zweiten Schritt Vergleichsstandards, in einem dritten Schritt müssen empirische Daten für jedes Kriterium und Alternativen gesammelt werden, bevor im vierten Schritt die Bewertung der Alternativen erfolgt.⁷

Während Scriven das Ziel verfolgt, Evaluationen objektiv zu halten und eher die Logik in der Zuweisung von Werten betont, konzentriert sich Campbell (1969) auf wissenschaftliche (experimentelle) Methodik. Er fordert die umfassende Anwendung randomisierter Experimente.⁸

Mitte der 70er Jahre begann eine qualitative Gegenbewegung. Was die verschiedenen Richtungen in der qualitativen Evaluation hauptsächlich unterscheidet, „is not methods, but rather whose questions are addressed and which values are promoted“.⁹ Das *Pragmatische Paradigma* entstand als Reaktion auf das so empfundene Versagen des experimentellen Ansatzes in der Evaluation. Diese Richtung wird dominiert von Michael Quinn Patton (1977), der in seinen Arbeiten die Interessen der mit der

⁶ Thomas D. COOK, Georg E. MATT (1990): *Theorien der Programmevaluation – Ein kurzer Abriss*, 17.

⁷ Vgl. hierzu Thomas D. COOK, Georg E. MATT (1990): *Theorien der Programmevaluation – Ein kurzer Abriss*; Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 14 sowie Romano GROHMANN (1997): *Das Problem der Evaluation in der Sozialpädagogik*, 66ff.

⁸ Vgl. Thomas D. COOK, Georg E. MATT (1990): *Theorien der Programmevaluation – Ein kurzer Abriss*, 19 f. und Romano GROHMANN (1997): *Das Problem der Evaluation in der Sozialpädagogik*, 63ff.

⁹ Jennifer C. GREENE (1994): *Qualitative Program Evaluation: Practice and Promise*, 533.

Umsetzung von Programmen beauftragten Personen in den Mittelpunkt stellte. Er präferiert die Anwendung qualitativer *und* quantitativer Verfahren und sieht den Evaluationsforscher eher als Künstler denn als Wissenschaftler, da der Evaluationsforscher für den Einzelfall die angemessenen Methoden ermittelt. Vertreter des *interpretativen Paradigmas* verstehen sich als Teil des „interpretative turn in social science“ und betonen die Einmaligkeit einer jeden Evaluation. Sie bedienen sich fast ausschließlich qualitativer Methoden. Die *kritisch-normative* Richtung geht über die Bewertung einzelner Programme hinaus und versucht „to illuminate the historical, structural, and value bases of social phenomena and, in doing so, to catalyze political and social change toward greater justice, equity, and democracy“.¹⁰

Seit den 80er Jahren stehen u.a. mit den Arbeiten von Peter H. Rossi und Lee J. Cronbach *integrative Ansätze* im Vordergrund. Sie versuchen eine Synthese zwischen den sich ausschließenden Konzepten zu leisten.¹¹

Die Evaluationsverfahren sind auf das spezifische Informationsbedürfnis und auf die kontextuellen Bedingungen abzustimmen und nicht nach der paradigmatischen evaluationstheoretischen Position. Je nachdem, welche Informationsbedürfnisse bei der Evaluation vorliegen und abgestimmt auf die Art des zu evaluierenden Projekts und auf die Entwicklungsphase, in der es sich befindet, müssen mehrere und ganz unterschiedliche Fragestellungen als relevant erachtet werden, die dann mit den jeweils geeigneten Verfahren zu beantworten sind: „Es ist schlicht sinnlos zu fragen, welcher Ansatz der ‚bessere‘ sei, wenn nicht zugleich die Untersuchungsfrage genannt wird. Entscheidend ist die Verschränkung von Methodik und Untersuchungsziel: Wer den einen gegen den anderen Ansatz ausspielt, trägt nur zu einer unnötigen und letztendlich nicht zu rechtfertigenden Polarisierung in der Sozialforschung bei“.¹²

Die Evaluation hat sich in den USA nicht nur in der Politik und Privatwirtschaft sondern auch im Bereich der Universitäten etabliert. Dies

¹⁰ Jennifer C. GREENE (1994): *Qualitative Program Evaluation: Practice and Promise*, 533.

¹¹ Vgl. Peter H. ROSSI, Howard E. FREEMAN, Gerhard HOFFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*; Huey-Tsyh CHEN (1990): *Theory Driven Evaluators*; Huey-Tsyh CHEN, Peter H. ROSSI (1992): *Using Theory to Improve Program and Policy Evaluations*; Peter H. Rossi, Howard E. Freeman, Mark W. Lipsey (1999): *Evaluation – A Systematic Approach* sowie Lee J. CRONBACH (1980): *Toward Reform of Program Evaluation* und Lee J. CRONBACH (1982): *Designing Evaluations of Educational and Social Programs*.

¹² Peter H. ROSSI, Howard E. FREEMAN, Gerhard HOFFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*, 188.

zeigt sich an der Gründung mehrerer Berufsorganisationen und der Festlegung anerkannter Standards durch das *Joint Committee on Standards for Educational Evaluation*.

Seit 1981 kodifizieren die Program Evaluation Standards des Joint Committee on Standards for Educational Evaluation in den USA und in englischsprachigen Ländern den Grundkonsens der Disziplin. Für den fortschreitenden Professionalisierungsprozess von Evaluation geben die Evaluationsstandards wichtige Impulse. Die Diskussion über Leitlinien und Standards der Evaluation wurde mit der deutschen Übersetzung auch den deutschsprachigen Interessenten zugänglich gemacht. Als Einrichtung hierfür steht die Deutsche Gesellschaft für Evaluation (DeGEval).

Die amerikanischen ‚Program Evaluation Standards‘ sind das Ergebnis eines Pionierprojekts aus dem Jahr 1975. Ziel war die Entwicklung von Standards, die dazu beitragen, nützliche, durchführbare, ethische und fundierte Evaluation von Programmen, Projekten und Materialien in den Bereichen Bildung und Erziehung zu ermöglichen. Diese Standards werden angewendet, überprüft und verbessert, um die Praxis der Evaluation voranzubringen. Sie wurden auch auf andere Kontexte übertragen. Der Kanon von Evaluationsstandards wird definiert als „ein Prinzip, auf das sich die in der Evaluation professionell tätigen Praktiker geeinigt haben, und dessen Beachtung dafür sorgt, dass die Qualität und die Fairness einer Evaluation verbessert werden“.¹³

Da keine klare Definition vorliegt, was eine anerkannte Evaluation von bildungspädagogischen Programmen ausmacht, hat das Joint Committee Kenntnisse über Programmevaluation zusammengetragen. Die Standards wollen „die Evaluatorinnen und alle weiteren, an Evaluation beteiligten Personen dabei unterstützen, qualitativ hochstehende Evaluationsprojekte zu realisieren, den Auftraggebern und weiteren Adressaten einer Evaluation dabei behilflich sein, deren Qualität einzuschätzen und den mit der Aus- und Weiterbildung in Evaluation befaßten Personen ein didaktisch aufbereitetes Material zur Verfügung zu stellen“.¹⁴

Im Kern werden 30 Standards formuliert. Sie sind in vier Kategorien eingeteilt: der *Nützlichkeit*, „Evaluation dient den praktischen Informationsbedürfnissen bestimmter Adressaten“; der *Durchführbarkeit*, „Evaluation ist realistisch, wohlüberlegt, diplomatisch und kostenbewußt“; der *Korrektheit*, „Evaluation wird rechtlich und ethisch korrekt durchgeführt

¹³James R. SANDERS (2000): *Aspekte der Entwicklung und Verbreitung der Evaluationsstandards*, 25.

¹⁴Ebd., 8.

und schenkt dem Wohlergehen der an der Evaluation beteiligten wie auch durch deren Ergebnisse betroffenen Personen die gebührende Aufmerksamkeit“ und der *Genauigkeit*, „stellt zur Güte und/oder Verwendbarkeit des Evaluationsgegenstandes systematisch gewonnene Informationen bereit“¹⁵. Die Standards werden als Mittel angesehen, um die genannten Ziele zu erreichen.

Die internationale Übertragung der Standards und die Übertragung auf Kontexte außerhalb des Bildungsbereichs wurde wissenschaftlich untersucht und als gegeben betrachtet. Deshalb wird die Anwendung der Standards auf Evaluationen in Deutschland empfohlen.¹⁶ Jedoch ist anzuführen, dass die Evaluationsstandards bisher in Deutschland in keinem nennenswerten Umfang eingesetzt wurden.

In Deutschland setzte die Entwicklung der Evaluation erst Anfang der 70er Jahre im Zuge der Reformprogramme der sozialliberalen Koalition ein. Im Gegensatz zu den USA konnte sie sich bis heute nicht als eigenständiger Dienstleistungsbereich etablieren.¹⁷ Die Evaluationstätigkeit selbst gewann zwar an Umfang, jedoch werden diese Arbeiten überwiegend von Markt- und Meinungsforschern, Mitarbeitern des betrieblichen Rechnungswesens oder Controllern durchgeführt.¹⁸

Die Gründe der Entwicklungsdefizite liegen in der Skepsis der deutschen Administration gegenüber der Evaluation, der institutionellen Zersplitterung der Fachressorts auf administrativer Ebene und v.a. der disziplinären Segmentierung der Evaluation: Eine über Fachgrenzen hinausgehende Wissensakkumulation findet kaum statt, die Weiterentwicklung und Standardisierung von Methoden wird erschwert, und die Ausbildung von Evaluationsforschern und -anwendern in den Administrationen ist häufig unzureichend.¹⁹ Darüber hinaus fehlt nach Stockmann eine Vermittlung des Wertes von Evaluationen.²⁰ Eine Lösung

¹⁵ James R. SANDERS (2000): *Aspekte der Entwicklung und Verbreitung der Evaluationsstandards*, 237.

¹⁶ Thomas WIDMER, Wolfgang BEYWL (2000): *Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder*, 257.

¹⁷ Vgl. hierzu Peter H. ROSSI, Howard E. FREEMAN, Gerhard HOFFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*, 7 sowie Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 30.

¹⁸ Sebastian BAMBERG, Harald GUMBEL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 30.

¹⁹ Reinhard STOCKMANN (1998): *Kleine Entwicklungsgeschichte der Evaluationsforschung. Nachholende Entwicklung in Deutschland*, 33.

²⁰ Ebd., 32.

sieht er darin, wegweisende über Ressortgrenzen hinaus bekannte Studien, die entsprechende Schlüsselergebnisse vermitteln könnten, zu erstellen.

3. Bedeutung und Entwicklung der Evaluation an der Hochschule

In der Hochschulforschung ist Evaluation ein Sammelbegriff für unterschiedliche Bewertungs- und Analyseverfahren, die von der Hörer- oder Absolventenbefragung über die Durchführung von Einzelstudien bis hin zu dem Modell der sogenannten internen und externen Evaluation reichen. Evaluationsverfahren beziehen sich auf Lehre und Forschung. Evaluation ist fach- bzw. programmorientiert: Sie setzt auf der Fakultäts- bzw. Fachbereichsebene an. Nur in seltenen Fällen, z.B. bei Verfahren von ZEvA und Nordverbund²¹, ist Evaluation auf regionaler Ebene hochschulübergreifend organisiert, d.h. vergleichbare Studienprogramme werden gemeinsam evaluiert.

In den Hochschulen gewinnen Maßnahmen zur Sicherung der Qualität der Lehre erst Anfang der 90er Jahre an Bedeutung. Begründet liegt diese Entwicklung v.a. in finanziellen Engpässen der Länder, im verstärkten nationalen und internationalen Wettbewerb der Hochschulen und in der Verantwortung und Rechenschaft der Bildungsträger gegenüber den Studierenden und der Öffentlichkeit sowie gegenüber den staatlichen Institutionen als Finanzgeber. Die Qualitätssicherung durch Evaluation steht aber auch im Zusammenhang mit anderen wichtigen Entwicklungen in den Hochschulen wie etwa der langfristigen Strategieplanung, der Mittelverteilung, der Reform von Leitungsstrukturen und Organisationsentwicklung. Evaluation wird dadurch selbst zu einem Instrument grundlegender Hochschulentwicklung. Davon ist auch die Akkreditierung von Bachelor- und Master-Studiengängen betroffen, die sich auf Evaluationsergebnisse stützt.

Standards für Hochschulabschlüsse und ihre gegenseitige Anerkennung wurden bisher durch Vereinbarungen zwischen den Ländern, vertreten durch die Kultusministerkonferenz (KMK) und den Hochschulen, vertreten durch die Hochschulrektorenkonferenz (HRK), über Bestimmungen für Prüfungsordnungen und fächerspezifische Rahmenprüfungsordnungen sichergestellt. Zur Umsetzung der Studienstrukturreform wurde 1993 die Einführung qualitätssichernder Maßnahmen durch die KMK und HRK

²¹ Diese Institutionen werden im Folgenden kurz erläutert.

empfohlen. 1994 und 1995 haben die HRK und der Wissenschaftsrat²² zwei Pilotprojekte durchgeführt, mit dem Ziel, die Diskussion über Evaluationsverfahren auf nationaler Ebene auszuweiten, allgemein anerkannte Regeln für die Qualitätssicherung herauszuarbeiten und standardisierte Indikatoren der Leistungsmessung zu begründen. Die HRK betreute als Teil eines europäischen Pilotprojekts zur Qualitätsbewertung im Hochschulbereich die Evaluation zweier Fachbereiche. „Im Ergebnis dieser Pilotprojekte haben HRK und der Wissenschaftsrat Empfehlungen zur Entwicklung von Qualitätssicherung in den Hochschulen mit sehr ähnlichen Ergebnissen verabschiedet“.²³

Mittlerweile wurden in den Hochschulgesetzen fast aller Bundesländer die Fachbereiche zu regelmäßigen Lehrberichten verpflichtet, was eine gemeinsame Informationsgrundlage der Fachbereiche ermöglicht.

Eine obligatorische Evaluation in systematischer Form findet jedoch nur in einigen Bundesländern bzw. Hochschulen statt. Das im August 1998 novellierte Hochschulrahmengesetz (HRG) verpflichtet die Hochschulen u.a. zur Qualitätssicherung in Lehre und Forschung (§6). Die Länder haben dies bereits in ihren Hochschulgesetzen aufgegriffen bzw. entsprechende Änderungen eingeleitet.²⁴

Die föderative Struktur in Deutschland und die Zuständigkeit der Länder für das Politikfeld Bildung erschweren die Festlegung allgemein anerkannter Regeln für die Qualitätssicherung und eine einheitliche Vorgehensweise der Evaluation von Lehre und Forschung an den Universitäten in Deutschland.

Es existieren im Wesentlichen von Hochschulen getragene Agenturen, Netzwerke und Verbände für Evaluation. Zu nennen sind:

- Verbund Norddeutscher Universitäten (Nordverbund), seit 1994. Hier haben sich die Universitäten Bremen, Hamburg, Kiel, Oldenburg, Greifswald und Rostock zusammengeschlossen.

²² Der Wissenschaftsrat berät die Bundesregierung und die Regierungen der Länder. Er besteht aus Wissenschaftlern, anerkannten Persönlichkeiten des öffentlichen Lebens und Vertretern des Bundes und der Länder und übernimmt eine doppelte Vermittlungsfunktion, nämlich zwischen Wissenschaft und Politik sowie zwischen Bund und Ländern.

²³ Vgl. HRK, Hrsg. (2000): *Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen unter besonderer Berücksichtigung der Lehre*, 7ff.

²⁴ Vgl. HRK, Hrsg. (2000): *Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen unter besonderer Berücksichtigung der Lehre*, 8.

-
- Zentrale Evaluations- und Akkreditierungsagentur (ZEVA) in Hannover, von der Landeshochschulkonferenz (LHK) Niedersachsen getragen und vom Land gefördert, seit 1995.
 - Geschäftsstelle Evaluation der Universitäten Nordrhein-Westfalen mit Sitz an der Universität Dortmund, seit 1996.
 - Geschäftsstelle Evaluation der Fachhochschulen Nordrhein-Westfalen mit Sitz an der FH Gelsenkirchen, seit 1997.
 - Lehrevaluation in der Universitätspartnerschaft (LEU), Evaluationsverbund im Rahmen der Universitätspartnerschaft Halle – Jena – Leipzig mit einer Koordinationsstelle an der Universität Halle, seit 1999.
 - Zentrum für Qualitätssicherung und -entwicklung an der Johannes Gutenberg-Universität Mainz, seit 1999.
 - Stiftung Evaluationsagentur Baden-Württemberg, seit 2000.
 - Evaluationsnetzwerk Wissenschaft (ENWISS), auf Initiative der TU Darmstadt gegründeter Verbund mit insgesamt dreizehn Fachhochschulen und Universitäten aus Hessen, Rheinland-Pfalz, Saarland und Thüringen, seit 2001.

Die Agenturen und Netzwerke verfügen i.d.R. über eine Geschäftsstelle. Sie stellt die Infrastruktur für den Evaluationsprozess bereit und ist verantwortlich für Vorbereitung und Administration, Einhaltung des Zeitplans, Vollständigkeit von Datenangaben der zu evaluierenden Fächer und Fachbereiche, Weiterleitung der Selbstevaluationsberichte an die externen Gutachter (Peers), Organisation der externen Evaluation vor Ort und Veröffentlichung der Abschlussberichte.

Evaluation wird auf Anfrage einzelner Fachbereiche auch von externen Einrichtungen betreut. Hier ist die HIS GmbH (Hochschul-Informationssystem) in Hannover, das Centrum für Hochschulentwicklung (CHE) in Gütersloh und das Interdisziplinäre Zentrum für Hochschuldidaktik in Bielefeld (IZHD) anzuführen.

4. Ziele und Funktionen der Evaluation an der Hochschule

Ziel der Evaluation an der Hochschule ist die Qualitätssicherung und -entwicklung von Lehre und Forschung. Stärken und Schwächen der zu evaluierenden Fächer und Fachbereiche sind im Hinblick auf die angestrebten Ziele und Standards herauszuarbeiten. Leistungsstandards sollen besser eingeschätzt, Qualitätsmängel erkannt und behoben und Transparenz verstärkt werden. Evaluation soll dazu beitragen, die internen Selbstkontrollmechanismen zu stärken, um zu systematischen Strategien der Qualitätssicherung- und verbesserung zu gelangen.²⁵ Evaluation beinhaltet u.a. Bereiche wie Studienordnung, Personalstruktur, Sonderforschungsbereich und Drittmittelprojekte, Profilbildung, Kommunikation, Kooperation und/oder die Einführung neuer Studiengänge. Die Lehrevaluation unterstützt eine mögliche Akkreditierung neuer Studiengänge: Die Bewertung der zur Zeit bestehenden Studiengänge, Fragen nach Schwerpunkten in Lehre und Forschung und Überlegungen bezüglich Sinn und Umsetzung neuer Studiengänge werden vorgenommen.

Evaluationen sind im Hinblick auf die Eigenverantwortlichkeit der Fachbereiche bei der Mittelverteilung und der Verteilungskämpfe unter den Fachbereichen, den Hochschulen und Fachhochschulen eine Chance. Denn häufig versäumen Fächer die positive Außendarstellung z.B. zentraler Forschungsprojekte, Sonderforschungsbereiche, attraktiver Studiengänge oder einzigartiger Einrichtungen wie Bibliotheken oder Archive. Die Profilierung von Fächern und Fachbereichen auf nationaler und internationaler Ebene durch zunehmenden Wettbewerb ist dementsprechend ein weiteres Ziel.

Evaluationen haben im Wesentlichen zwei Funktionen: Qualitätssicherung in den Hochschulen und die Sicherung der Rechenschaftspflicht der Hochschulen gegenüber der Gesellschaft. Während die Funktion der Qualitätssicherung nach innen, auf Entscheidungs- und Organisationsentwicklungsprozesse innerhalb der Hochschule gerichtet ist, richtet sich die Funktion der Rechenschaftspflicht nach außen, auf den Nachweis von Leistungen und auf die Kontrolle der Gesellschaft bzw. des Trägers der Hochschule.

Die Rechenschaftspflicht hat in den letzten Jahren eine neue Bedeutung erlangt, wie z.B. die in den Medien publizierten Rankings von Fächern

²⁵ Vgl. HRK (Hrsg.) (2000): *Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen unter besonderer Berücksichtigung der Lehre*, 11.

deutscher Hochschulen zeigen. Dies ist vor allem auf die mangelnden Finanzmittel der öffentlichen Kassen zurückzuführen: Mangelnde Finanzmittel führen zu Verteilungskämpfen und zwangsläufig zu einer verstärkten Rechenschaft gegenüber der Öffentlichkeit und den staatlichen Institutionen in ihrer Funktion als Finanzgeber. Während in der Vergangenheit grundsätzlich angenommen wurde, Hochschulinvestitionen seien sinnvoll angelegt, ist gegenwärtig ein Vertrauensverlust in die Leistungsfähigkeit der Wissenschaft zur Lösung politischer, ökonomischer und ökologischer Probleme zu beobachten.²⁶ Und je stärker eine Institution in das Blickfeld des Interesses rückt, umso eher werden Defizite sichtbar und Organisationselemente und Maßnahmen in Frage gestellt. Die Diskussion zu nachweisbaren Relationen zwischen Kosten und Nutzen der Hochschulen in Deutschland wird im internationalen Vergleich zeitversetzt und mit weniger Selbstkritik geführt.²⁷

Die Qualitätssicherung an den Hochschulen hat in den letzten Jahren eine neue Form erhalten: Während die Hochschule in der Vergangenheit durch eine *ex-ante*-Steuerung geprägt war, besteht heute die Tendenz zu einer *ex-post*-Steuerung.

Bei dem Verfahren der *ex-ante*-Steuerung versucht der Staat durch Fixierung wesentlicher Rahmenbedingungen im Vorhinein die Qualität zu sichern. Durch zuvor ergriffene Maßnahmen wird späteres Handeln und Wirken vorherbestimmt. Dabei handelt es sich häufig um Festlegung über einen Zeitraum von mehreren Jahrzehnten: Berufungen beispielsweise bestimmen die Qualität und das Profil eines Lehrstuhls über Jahre hinweg, wobei vor der Besetzung in einem Berufungsverfahren die wissenschaftliche Qualität geprüft wird. Anschließend bestehen keinerlei Eingriffsmöglichkeiten mehr. Die *ex-ante*-Steuerung basiert auf Haushaltsvorgaben, gesetzlichen Vorschriften über die Organisationsstruktur, Genehmigungsvorbehalten hinsichtlich Berufungen, Fachbereichen und Studiengängen.²⁸

Das Prinzip der *ex-ante*-Steuerung ist kennzeichnend für die kontinental-europäischen Staaten. Die Vorteile dieses Systems liegen in der Siche-

²⁶ Vgl. Detlef MÜLLER-BÖLING (1997): *Evaluation zur Rechenschaftslegung oder Qualitätsverbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen*, 91.

²⁷ Vgl. Roland RICHTER (1996): *Qualitätsmanagement im europäischen Hochschulwesen – Drei Varianten über ein Thema*.

²⁸ Vgl. Detlef MÜLLER-BÖLING (1997): *Evaluation zur Rechenschaftslegung oder Qualitätsverbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen*, 89f.

zung einer Homogenität der Qualität, in der Freiheit der berufenen Professoren und darin, dass eine Finanzverantwortung des Staates im Hinblick auf die Alimentierung der Hochschule besteht. Die Nachteile liegen im individuellen Freiheitsmissbrauch, der finanziellen Abhängigkeit der Hochschulen vom Staat und vor allem in der hohen Inflexibilität: Die Festlegung bei Berufungen über einen Zeitraum von mehreren Jahrzehnten wird den Erfordernissen der rasanten Umweltentwicklung nicht mehr gerecht und die Abstimmungsprozesse zur Koordinierung der Homogenität schreiten zu langsam voran, so die Hauptkritik des Steuerungsverfahrens.²⁹ Folglich hat sich das Steuerungsverfahren in einer stabilen Umwelt mit ausreichenden finanziellen Ressourcen bewährt; bei instabilen Umweltverhältnissen und veränderten Anforderungen der Gesellschaft an die Ausbildungsinhalte, explosionsartiger Bildungsnachfrage oder zurückgehenden Ressourcen erweist sich dieses Steuerungsmodell als nicht mehr erfolgreich.³⁰

Deshalb ist eine Tendenz zur ex-post-Steuerung sichtbar. Bei diesem Verfahren erfolgen Belohnung oder Sanktionierung entsprechend der Ergebnisse und Leistungen. Dies kann mittels beruflichem Aufstieg, Mittelzuweisungen, Mittelkürzungen, Ansehensgewinn oder -verlust erfolgen. Das Hauptproblem jedoch besteht darin, Qualität im Hochschulbereich zu definieren.

In der hochschulpolitischen Diskussion kommt dem Qualitätsbegriff zentrale Bedeutung zu. Leider wird er allzu oft als Schlagwort benutzt, ohne dass in jedem Fall eine hinreichende begriffliche Abgrenzung dessen, was unter Qualität zu verstehen ist, geleistet wird. In einem allgemeinen Sinn wird Qualität in dem vorliegenden Band als die *Abweichung bzw. Übereinstimmung zwischen definierten Zielen und der jeweiligen messbaren bzw. beschreibbaren ‚Wirklichkeit‘* verstanden. So definiert, lässt sich Qualität nicht als eindimensionale Bezugsgröße benennen, sondern ist vor allem von der Ebene sowie der zeitlichen Perspektive der zugrunde gelegten Ziele abhängig. Folglich muss als weiteres Merkmal von Qualität das Maß der *Kompatibilität von Zielen und Teilzielen* auf unterschiedlichen Ebenen einer Einrichtung, einer Organisation oder eines Programms formuliert werden. So ist es beispielsweise denkbar, dass zwar die mittelfristigen Ziele eines Faches zu einem hohen Maß umgesetzt werden können, gleichzeitig aber die Anbindung an über-

²⁹ Vgl. Detlef MÜLLER-BÖLING (1997): *Evaluation zur Rechenschaftslegung oder Qualitätsverbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen*, 90.

³⁰ Vgl. ebd.

greifende und langfristige Ziele der Hochschule nicht gelingt. Die Messung, Entwicklung und Sicherung von Qualität muss diese Gleichzeitigkeit von zum Teil sich widersprechenden Qualitätsmerkmalen abbilden.

5. Evaluationsverfahren an deutschen Hochschulen

Unter Evaluation lassen sich mittlerweile eine Vielzahl von Verfahren subsumieren, die alle der Bewertung von Leistungen dienen. Größtenteils handelt es sich um die zuvor beschriebenen Verfahren der ‚ex-post-Analysen‘: Objekt der Evaluation bilden bereits erbrachte Leistungen, die im Nachhinein beurteilt werden. Die Art und Weise, in der die Evaluation erfolgt, kann jedoch sehr unterschiedlich sein und ist in der Regel abhängig von den finanziellen Ressourcen sowie von der Personalkapazität.

Im Folgenden werden die momentan an deutschen Hochschulen am häufigsten zum Einsatz kommenden Verfahren vorgestellt. Dies sind:

- Befragungen von Studierenden beziehungsweise Veranstaltungsbewertungen durch Studierende,
- gesetzlich vorgeschriebene Lehrberichte und
- interne und externe Evaluation.³¹

Mit dem Einsatz von **Veranstaltungsbewertungen** begann Anfang der 90er Jahre die Evaluationsdiskussion im Hochschulbereich in Deutschland. In der Regel mittels eines standardisierten Fragebogens durchgeführt, wird mit diesem Verfahren versucht, zu einer Einschätzung von Lehrveranstaltungen aus Sicht der Studierenden zu gelangen. Dabei werden zentrale Dimensionen wie Aufbau der Veranstaltung, Angemessenheit von Stoffmenge und Schwierigkeitsgrad, Didaktik, Durchführung, Auftreten der Lehrperson und soziales Klima und Lernerfolgseinschätzung erhoben.

Häufig werden diese Hörerbefragungen von Studierenden (Fachschaft oder auch AStA) selbst angeregt. Es gibt aber durchaus auch Lehrende, die eine Seminar- oder Vorlesungskritik durch Studierende initiieren, und

³¹ Vgl. *Evaluationsverfahren*, <http://evanet.his.de/evanet/knowhow/kh.grund/kh.evalarten.html> (Homepage des EvaNet – Das Evaluations-Netzwerk ist eine offene Plattform für den Erfahrungsaustausch über die Evaluation und Qualitätssicherung von Studium und Lehre an deutschen Hochschulen. Die Homepage wird von HIS im Auftrag der HRK bereitgestellt.)

diese Art der Veranstaltungsbewertung sogar bereits seit geraumer Zeit durchführen lassen.

Diese Art der Befragung lässt sich zwar einfach, schnell und im Vergleich mit anderen Verfahren kostengünstig durchführen, ist jedoch hinsichtlich der Aussagekraft beziehungsweise der Validität der Ergebnisse mit einer Reihe von Erhebungs- und Deutungsproblemen verbunden.³² In den meisten Fällen werden die Fragebögen vom Lehrenden oder der betreffenden Fachschaft beziehungsweise dem Fachbereich selbst häufig eher „laienhaft“ und nur selten unter sozialwissenschaftlicher Anleitung entwickelt. Weiterhin sind beispielsweise die befragten Studierenden sogenannte Alltags-Evaluatoren: „Jeder einzelne von ihnen bewertet *irgend etwas* (was er mit dem in der Frage angesprochenen Sachverhalt ad hoc assoziiert) *irgendwie* (‘alles in allem‘ oder ‚aus aktueller Erfahrung‘ oder ‚mit Blick auf das Wesentliche‘ oder ...) *unter irgendwelchen Gesichtspunkten* (Nutzen für sein Studium oder vermuteter Nutzen für den angestrebten Beruf oder aktuelles persönliches Interesse oder abstrakt-verallgemeinertes Interesse der Studierenden oder ...)“³³ Darüber hinaus stellt zum Beispiel die Auswahl der zu Befragenden ein Problem dar. Werden die Teilnehmer von Wahl- und Wahlpflichtveranstaltungen – wie üblich – in der Mitte des Semesters befragt, bedeutet dies, dass durch die bis zu diesem Zeitpunkt erfolgte Selbstselektion die Bewertung positiv beeinflusst sein wird, denn diejenigen Studierenden, die der Veranstaltung und/oder dem Lehrenden gegenüber eine eher negative Einschätzung zukommen lassen, nehmen zu dem Datum an der Lehrveranstaltung bereits nicht mehr teil. Befragt man Studierende, die eine Pflichtveranstaltung ohne Wahlmöglichkeit zwischen alternativen Angeboten besuchen, fällt deren Urteil systematisch negativer aus. Die Gültigkeit der Ergebnisse wird ferner durch den Zeitpunkt des Einsammelns der Fragebögen mitbestimmt. Geschieht dies nicht unmittelbar in der entsprechenden Veranstaltung, sondern wird um Rückgabe zu einem späteren Termin gebeten, so fördert dies eine Neigung zur Polarisierung der abgegebenen Einschätzungen. So setzt die Befragung ein aktives Beteiligungsverhalten seitens der Befragten voraus, wozu sich insbesondere diejenigen Studierenden aufgefordert sehen, die „etwas zu sagen haben“, sei es außerordentliches Lob oder ausgesprochene Kritik. Demgegenüber ist die üblicherweise zahlenmäßig stärkste Gruppe der „einigermaßen Zufriede-

³² Helmut KROMREY (2001): *Qualität und Evaluation im System Hochschule*, 247ff.

³³ Helmut KROMREY (2001): *Qualität und Evaluation im System Hochschule*, 248. (Kursiv gesetzte Hervorhebung durch den Autor).

nen“ aufgrund fehlender Motivation der aktiven Beteiligung an der Umfrage eindeutig unterrepräsentiert.³⁴

Insgesamt lässt sich festhalten, dass das Verfahren der Veranstaltungsbeurteilung ein Instrument der Akzeptanzmessung und somit ein Element (neben anderen) der Ermittlung der Qualität der Lehre darstellt, aber keineswegs einen Ersatz der Evaluation nach unserem Verständnis bietet.³⁵

In einigen Bundesländern sind die Hochschulen durch die betreffenden Hochschulgesetze zur Abgabe von **Lehrberichten** verpflichtet. Diese Berichte haben zum einen die Funktion der Rechenschaftslegung über die Leistungen der einzelnen Hochschule im Bereich von Studium und Lehre, sollen zum anderen aber auch das Wissen über Studienbedingungen sowie Möglichkeiten der Reorganisation fördern. Sofern in den Hochschulgesetzen der Bundesländer überhaupt Angaben zu Form und Aufbau von Lehrberichten gemacht werden, unterscheiden sie sich hinsichtlich der Ausführungsvorgaben – was die verwendeten Indikatoren oder die zu erhebenden Sachverhalte betrifft – jedoch erheblich.³⁶

So ist beispielsweise unterschiedlich, in welchem Umfang quantitative Daten und Indikatoren einbezogen oder in welchen Intervallen die Lehrberichte verfasst werden müssen. In einigen Bundesländern gilt die Maßgabe eines Gesamtlehrberichts pro Hochschule, während in anderen jedes Fach einer Hochschule einen eigenen Lehrbericht anfertigen muss, der eventuell später – wie zum Beispiel in Rheinland-Pfalz im Abstand von zwei Jahren – zu einem Gesamtbericht zusammengefasst wird. Weiterhin weichen die Berichte hinsichtlich der Vorschriften über die Art der Veröffentlichung und der Einbeziehung des jeweilig zuständigen Ministeriums voneinander ab.

Als Manko dieses Verfahrens gilt die Tatsache, dass die Hochschulgesetze häufig nicht thematisieren, welche Ziele, Adressaten und Konsequenzen die Lehrberichte haben sollen.

³⁴ Ebd., 253f.

³⁵ Ebd., 254: „Die Erhebung und Auszählung bewertender (also ‚evaluierender‘) Aussagen ist noch keine Evaluation im wissenschaftlichen oder methodologischen Sinne, sondern ein Verfahren des Sammelns der persönlichen ‚Alltags-Evaluationen‘ der Befragten. ... Umfrageforschung ist nicht mehr – aber auch nicht weniger – als ein bewährtes Instrument der Informationserhebung.“

³⁶ HRK (Hrsg.) (2000): *Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen unter besonderer Berücksichtigung der Lehre*, 7.

Das Verfahren der **internen und externen Evaluation**³⁷, das zur Zeit in der Bundesrepublik bevorzugt eingesetzt wird, kann entweder ‚verordnet‘ sein oder auf freiwilliger Basis und auf Anfrage der Fächer und Fachbereiche³⁸ durchgeführt werden. Die interne und externe Evaluation wird üblicherweise in drei aufeinander aufbauenden Schritten vollzogen.

Den Kern der *internen Evaluation* bildet die Erstellung des *internen Evaluationsberichts*, auch Selbstreport oder Selbstdarstellung genannt, die üblicherweise in der Verantwortung des Fachs selbst liegt. Dabei soll eine systematische Untersuchung von Lehre und Studium mit dem Ziel, ein detailliertes Profil von Stärken und Schwächen der Ausbildung erstellen zu können, geleistet werden. Zur Analyse werden zum einen *quantitative Daten* wie Verwaltungsangaben herangezogen.

Hierzu gehören sowohl *statistisch erfasste Studierendendaten* hinsichtlich

- Stand und Entwicklung der Studierenden (Zahl der Studienanfänger, der Studierenden insgesamt nach Fachsemestern und Zahl der Absolventen)
- Studiendauer (Fachstudiendauer bis Vordiplom/Zwischenprüfung, Fachstudiendauer bis zum Abschlussexamen, sowie Dauer der Abschlussarbeit)
- Alter der Studierenden bei Studienbeginn und bei erfolgreich abgelegtem Examen
- Erfolg und Misserfolg (Gesamtnote und Notenverteilung im Vordiplom/Zwischenprüfung sowie Abschlussexamen Erfolgs-/ Misserfolgsquoten, Notenverteilung für zentrale Prüfungen) und
- Angaben zu den Zu- und Abgängen der Studierenden durch Fach- und Hochschulwechsel, Schwund und Abbruch

als auch *Personal- und Ressourcendaten* des Fachs.

³⁷ Das im Folgenden beschriebene Verfahren orientiert sich am Vorbild des sogenannten „Niederländischen Modells“, nach dem sich auch die Empfehlungen von HIS ausrichten.

³⁸ Für ‚Fach beziehungsweise Fachbereich‘ wird im folgenden Text der Einfachheit halber die Bezeichnung ‚Fach‘ verwendet.

Zum anderen stützt sich der interne Evaluationsbericht auf *qualitative Daten*. Hier sind zunennen: die Ergebnisse aus den *moderierten Gruppengesprächen* mit Studierenden in Grund- und Hauptstudium, mit nichtwissenschaftlichen Mitarbeitern, wissenschaftlichen Mitarbeitern und Professoren, in denen die bestehende Situation in Studium und Lehre angesprochen und die Diskussion um Vorschläge zur Verbesserung derselben angeregt wird, sowie die *Befragung* von Absolventen eines Fachs.

Auf der Grundlage der oben genannten Daten erstellt das Fach nun den *internen Bericht*, der drei Teilbereiche umfasst. Zunächst werden die existierenden Rahmenbedingungen, die zukünftig zu erwartenden Fortschritte und Entwicklungen sowie Vorhaben und Zielsetzungen dargelegt. Daran anschließend wird auf der Basis der durch die systematische Bestandsaufnahme erfassten Daten über die Studien-, Lehr- und Forschungspraxis die momentane Situation beschrieben. Abschließend soll durch den Abgleich zwischen den ersten beiden Bereichen in der Art eines ‚Soll-Ist-Vergleichs‘ ein Stärken- und Schwächenprofil erstellt und potentielle Maßnahmen zur Verbesserung der festgestellten Schwächen vorgeschlagen werden. Die Formulierung von Schwächen und in gleichem Maße auch von Stärken fällt den Fächern erfahrungsgemäß nicht leicht, da ihnen diese Art der Selbstbeschreibung oft nicht vertraut ist und zunächst in einem zuvor (selten öffentlich geleisteten) Reflexionsprozess erarbeitet werden muss.

Nachdem die Selbstevaluation fertiggestellt und im Fach verabschiedet ist, findet die *externe Evaluation* statt. Ihr Ziel ist die Ergänzung der Innenschau des Fachs durch die Überprüfung der Qualität von Studium, Lehre und Forschung aus der Perspektive externer Gutachter, der sogenannten *Peers* (Gleichrangige – zu verstehen als die Ebenbürtigkeit von den Lehrenden eines Fachs und den externen Gutachtern). Gewöhnlich setzen sich Peers aus Hochschullehrern und gegebenenfalls auch aus in der Praxis tätigen Personen zusammen. Bei einigen externen Evaluationen wirken sogar Fachstudenten aus anderen Hochschulen mit. In der Regel besteht die externe Gutachtergruppe aus drei bis sechs, bei größeren Fachbereichen auch aus mehr Personen. Gewöhnlich wird die Auswahl der externen Fachgutachter vom Fach selbst getroffen, wobei auf Kriterien wie Fachgebiet und fachliche Schwerpunkte und damit einhergehend Vertrautheit mit dem zu evaluierenden Fach und dem Hochschultyp, wissenschaftliche Position und Sachkompetenz und schließlich auch berufliche Position und Tätigkeitsfeld geachtet wird.

Die Begutachtung erfolgt auf der Grundlage des internen Berichts, den die Experten zur Vorbereitung bekommen, und eines üblicherweise zweitägigen Vor-Ort-Besuchs der Peers. Aufgabe der externen Gutachter ist zum einen die *problemorientierte Lektüre des Selbstberichts*, die auf den Besuch vor Ort vorbereiten soll. Zum zweiten sollen sie während ihrer Anwesenheit vor Ort einen *kritischen Blick auf die Studien-, Lehr- und Forschungssituation im Fach* mittels Gesprächen mit den am Evaluationsprozess beteiligten Gruppen und einer Vor-Ort-Begehung, die einen Überblick über Räumlichkeiten und Ausstattung geben soll, werfen und während dieser Zeit die zentralen Kritikpunkte sowie Anregungen und Empfehlungen zu Verbesserungen erarbeiten und präsentieren (*mündliches Gutachten*). Auf diesen den Vor-Ort-Besuch abschließenden Einschätzung wird dann im Anschluss bei der Erstellung des *schriftlichen Gutachtens*, das in der Regel von dem Experten erarbeitet wird, der zu Beginn des Besuchs als der für die Erstellung verantwortliche Gutachter bestimmt wurde, aufgebaut.

Nach der Verabschiedung des *externen Gutachtens* im Fach führen Fach und Vertreter der Hochschulleitung auf der Grundlage dieses sowie des *internen Berichts* Gespräche mit dem Ziel einer schriftlich fixierten **Zielvereinbarung**, in der die zur Behebung der formulierten Kritikpunkte vorgesehenen Maßnahmen sowie der für deren Durchführung geplante Zeitraum verbindlich festgeschrieben werden. Mit der beiderseitigen Unterzeichnung dieser Zielvereinbarung ist das Evaluationsverfahren abgeschlossen.

Der folgende Artikel fasst das zuletzt angeführte Verfahren, das als das sogenannte „Mainzer Modell“ an der Johannes Gutenberg-Universität Anwendung findet, zusammen und beschreibt die bis jetzt gesammelten Erfahrungen.

Evaluation nach dem „Mainzer Modell“

Elisabeth Springer

1. GENESE DER EVALUATION AN DER UNIVERSITÄT MAINZ	22
2. FUNKTION UND ORGANISATION DES ZQ	24
3. STAND DER EVALUATION AN DER UNIVERSITÄT MAINZ.....	27
4. ENTWICKLUNG UND AUSBAU DES EVALUATIONS - VERFAHRENS	29
5. FUNKTION UND ZIELSETZUNG DES EVALUATIONS - VERFAHRENS	40
6. ERGEBNISSE UND ERFAHRUNGEN DER EVALUATIONS - VERFAHREN	43

1. Genese der Evaluation an der Universität Mainz

Seit dem Wintersemester 1995/96 wird an der Johannes Gutenberg-Universität in Mainz die Evaluation von Fächern und Fachbereichen durchgeführt. Im folgenden sollen das zugrundeliegende Verfahren, dessen Weiterungen und Modifikationen sowie die daraus abzuleitenden Schlußfolgerungen und Ergebnisse beschrieben werden.

Der hochschulinterne Stellenwert von Evaluation manifestiert sich u.a. in der Bereitstellung erheblicher Mittel zur Förderung von Studium und Lehre, der Einrichtung eines Strukturfonds sowie der Überführung des „Projektes zur Förderung von Studium und Lehre“ in das „Zentrum für Qualitätssicherung und -entwicklung“ (ZQ) und muss im Zusammenhang mit dem in Mainz gewählten Weg, die externe politische und gesellschaftliche Forderung nach Transparenz und Offenlegung im Hochschulbereich universitätsintern produktiv umzusetzen, gesehen werden.

Zeitlich vor und parallel zu den vom ZQ durchgeführten Evaluationen, die hier Gegenstand sein sollen, wurden und werden bspw. Veranstaltungsevaluationen in verschiedenen Fächern und Fachbereichen mit unterschiedlicher Ausgestaltung durchgeführt. So werden seit über zehn Jahren in der Medizin Studierendenbefragungen durchgeführt, aber auch in der Physik, Mathematik oder der Politikwissenschaft und Publizistik hat sich das Mittel der Einzelveranstaltungsbefragung seit längerer Zeit etabliert. Auch umfassendere Erhebungen das Fach bzw. die Studien- und Lehrsituation betreffend wurden bereits zu einem frühen Zeitpunkt an der Mainzer Universität z.B. in den Fächern Pädagogik, Germanistik sowie Biologie initiiert.¹ Bis zu den ersten Fachevaluationen 1995 durch das ZQ (damals Projekt zur Förderung von Studium und Lehre, PSL) hat es insofern in Mainz zeitlich und inhaltlich einen Vorlauf gegeben, der u.a. das grundsätzliche Bedürfnis der Fächer nach Standortbestimmung, Artikulationskompetenz und notwendiger Reform dokumentiert.²

Mit der Einrichtung des „Modellprojektes zur Förderung von Studium und Lehre“ konnten seit 1992 Projektinitiativen in den Fächern gefördert werden, deren Ergebnisse den zukünftigen Überlegungen hinsichtlich der Anforderungen an ein Evaluationsverfahren zugute kamen. Dies gilt nicht

¹ vgl. hierzu Uwe SCHMIDT, Simon HÄUSER (2001): *Problembereiche und Reformansätze in Studium und Lehre*.

² vgl. hierzu Manfred HENNEN, Elisabeth SPRINGER (1994): *Modellprojekt zur Förderung von Studium und Lehre*.

nur für die im Rahmen des o.g. Modellprojektes zur Förderung von Studium und Lehre explizit geförderten Evaluationsprojekte, die überwiegend Veranstaltungsbefragungen zum Gegenstand hatten. Auch andere Projektinitiativen, die sich mit der Verbesserung der Studien- und Lehrsituation auseinandersetzten, haben beispielsweise aufgrund der positiven Erfahrungen mit den Gesprächen vor Ort und der Methode des responsiven Verfahrens mittelbar zu der Entscheidung beigetragen, universitätsinterne Evaluationen durchzuführen und für die Ausgestaltung des Evaluationsverfahrens in Mainz Pate gestanden.

Insbesondere im Verlauf derjenigen Projekte, die Evaluation von Einzelveranstaltungen zum Gegenstand hatten, wurde deutlich, dass Evaluation per Fragebogen zweifelsohne als Möglichkeit zur Mobilisierung genutzt werden kann. Insbesondere in den auf studentische Initiativen zurückzuführenden Projekten wurden jedoch auch die Grenzen eines solchen Verfahrens deutlich. Die über mehrere Semester hinweg erhobenen und dokumentierten Monita waren bekannt, konnten jedoch durch die Fächer selbst nicht angegangen und beseitigt werden. Dies führte in zwei Fällen gar zu der Überlegung, die Evaluation von Einzelveranstaltungen einzustellen, da die Konsequenzlosigkeit der erhobenen Ergebnisse nachvollziehbarerweise zu einem erheblichen Motivationsverlust bei den Beteiligten führte.

Die hochschulinternen Überlegungen, die von den Fächern bereits initiierten Projekte wie Veranstaltungsbefragungen, Reorganisationen des Curriculums, Verbesserung der Informationssituation für Studierende etc. in einem umfassenderen Verfahren zusammenzufassen, trafen mit dem Interesse des heutigen Ministeriums für Wissenschaft, Weiterbildung, Forschung und Kultur (damals Ministerium für Bildung, Wissenschaft und Weiterbildung) zusammen, Evaluation landesweit an den Hochschulen einzuführen. Der Forderung der Johannes Gutenberg-Universität nach einer Form der Evaluation, die das Verfahren in der Zuständigkeit der Fächer insofern belässt, als die Fächer über Verfahrensform und Evaluationsgegenstand sowie Veröffentlichung der Ergebnisse entscheiden, wurde seitens des Ministeriums nachgekommen. Mit dieser Zusicherung gelang es, die dringend benötigte Motivation der Fächer für das Verfahren sicherzustellen.

Die Erfahrungen, die bereits im Rahmen der o.e. universitätsinternen Modellprojekte gesammelt werden konnten, führten mit dazu, dass die Fächer Mathematik und Politikwissenschaft im Wintersemester 1995/96 die auf Freiwilligkeit basierende Entscheidung trafen, eine Evaluation im Rahmen eines durch das Ministerium für Bildung, Wissenschaft und

Weiterbildung geförderten Pilotprojektes durchzuführen, an dem sich neben der Universität Mainz auch andere Hochschulen des Landes beteiligten.

Die Pilotevaluation in den beiden Fächern wurde von Mitarbeitern des Hochschulinformationssystems in Hannover (HIS) durchgeführt und vom ZQ begleitet, um hierdurch Erfahrungen für weitere universitätsinterne Evaluationsverfahren sammeln zu können und einer Fremdbestimmung entgegenzuwirken. Das zu diesem Zeitpunkt von HIS bereits an anderen Hochschulen praktizierte Verfahren der internen und externen Evaluation wurde im Anschluss an die Pilotphase vom ZQ, das von der konstruktiven Kooperation profitiert hat, in weiten Teilen übernommen, angepasst und weiterentwickelt (s. 4.).

2. Funktion und Organisation des ZQ

Bevor unter Punkt 4 auf das konkrete Verfahren der Evaluation nach dem „Mainzer Modell“ eingegangen werden kann, soll zunächst Organisation und Funktion des ZQ vorgestellt werden. ZQ wurde 1999 durch den Senat der Johannes Gutenberg-Universität mit der Rechtsstellung einer zentralen Betriebseinheit eingerichtet, nachdem die Aufgabenbereiche des seit 1992 bestehenden „Projektes zur Förderung von Studium und Lehre“ (PSL) kontinuierlich erweitert worden waren. Diese Überführung in eine wissenschaftliche, fachbereichsübergreifende Einrichtung stellt die Grundlage dafür dar, dass eine für Beratungstätigkeit im weitesten Sinne erforderliche neutrale, d.h. hier weder einem Fachbereich noch der Zentralverwaltung bzw. der Hochschulleitung zuzuordnende Perspektive eingenommen werden kann. Die überwiegend beratende Tätigkeit des ZQ benötigt diese Positionierung als intermediäre, um Neutralität bemühte treuhänderische Instanz. Eine unparteiische Perspektive der Beratenden ist jedoch nicht nur im Rahmen von Evaluationen zur Mobilisierung der Fächer und Motivation der Verantwortlichen von Bedeutung, auch für die Betreuung von Modellprojekten oder Initiativen im Rahmen der Nachwuchsförderung hat sich diese Perspektive bewährt.

Das ZQ, dem in allen grundsätzlichen Angelegenheiten ein entscheidungsbefugter Senatsausschuss unter Leitung des Vizepräsidenten für Studium und Lehre zur Seite gestellt ist, hat den Auftrag, qualitätssichernde Maßnahmen, die den notwendigen Weiterentwicklungen und Anpassungen im Bereich von Studium und Lehre dienen, durchzuführen.

Konkret sind derzeit folgende Arbeitsbereiche und Aufgabenschwerpunkte im ZQ angesiedelt:

- Den Schwerpunkt der Arbeit des ZQ bildet die *Evaluation von Studium, Lehre und Forschung*, worauf im folgenden noch ausführlich eingegangen wird.
- Die wissenschaftliche Begleitung von *Programmevaluationen* hat zum Gegenstand, neu eingeführte Maßnahmen bzw. Projekte auf ihre Wirksamkeit hin zu überprüfen. ZQ ist zur Zeit in ein gemeinsam vom Bund und dem Land Rheinland-Pfalz gefördertes Projekt im Bereich des Schulmusikunterrichtes eingebunden. Daneben wird eine Evaluation des vom Fachbereich Chemie der Mainzer Universität und dem Ada-Lovelace-Projekt initiierten Schüler/innen-Labors durchgeführt.
- Das seit April 1999 bestehende „*Forum wissenschaftlicher Nachwuchs*“ ist mit der Entwicklung und Koordinierung von Initiativen zur Förderung des wissenschaftlichen Nachwuchses befaßt, um die berufliche Qualifikation künftiger Hochschullehrerinnen und -lehrer umfassend zu fördern. Dieses Förderprogramm berücksichtigt gezielt auch diejenigen Bereiche, die traditionell in der Qualifikation des wissenschaftlichen Nachwuchses keine systematische Berücksichtigung finden. Hierzu zählen u.a. Fragen der Didaktik und Rhetorik, Fragen der Verwaltung einer Universität resp. eines Institutes, Fragen der Mitarbeiterführung und der Hochschulpolitik ebenso wie die Rolle und Verantwortung des Lehrenden.
- Die Betreuung von *Modellprojekten zur Förderung von Studium und Lehre* war, wie bereits o.e., 1992 Ausgangspunkt der Arbeit des ZQ (vorm. PSL). Die Initiative für die Einzelprojekte geht von Lehrenden und Studierenden der Universität aus und hat die Zielsetzung, modellhaft neue Wege in Studium und Lehre zu erproben. ZQ hat neben der wissenschaftlichen Begleitung auch die Aufgabe, die aus diesen Modellprojekten gewonnenen Erfahrungen zu dokumentieren. Ein erster Ergebnisbericht ist 1994 erschienen, ein vorläufiger Abschlußbericht liegt seit Juni 2001 vor.
- Im Bereich der Hochschulforschung befaßt sich ZQ derzeit mit der Auswertung einer Längsschnittstudie zu besonderen Formen der Studienbetreuung. Im Rahmen dieses Projektes wurden Studierende aus vier verschiedenen Fachbereichen über 10 Semester lang durch

einen Vertrauensdozenten ihres Faches in Studienfragen beraten und betreut. Im Gegenzug hatten die Studierenden sich bereit erklärt, im Rahmen von Befragungen per Fragebogen und Interview, Informationen über die Studiensituation zu liefern, um die Erhebung systematischer Studienhindernisse und Schwierigkeiten zu unterstützen. Die Publikation der Ergebnisse erfolgt in Kürze. Neben diesem Projekt findet eine breit angelegte Untersuchung zu Fragen der Studienkultur und Studienmotivation statt. Für diese Untersuchung wurden mehr als 2.000 Studierende ausgewählter Fachbereiche befragt. Bei der Auswahl fanden solche Fächer Berücksichtigung, bei denen davon auszugehen ist, dass bei den Studierenden signifikante Unterschiede im Hinblick auf Studienzugang und Studienmotivation zu erwarten sind. Erste Teilauswertungen sind bereits erfolgt, Ergebnisse sind im laufenden Jahr zu erwarten.

- Ein letztes Aufgabenfeld berücksichtigt einen Bereich, der mit *Organisationsentwicklung* beschrieben werden kann. Darunter ist im Zusammenhang mit der Evaluation von Fächern im Hinblick auf die Forschungs-, Studien- und Lehrsituation zu verstehen, dass die evaluierten Fächer und Fachbereiche zunehmend auch nach dem offiziellen Abschluss des Verfahrens, der mit der Unterzeichnung der Zielvereinbarungen durch das Fach und die Hochschulleitung erreicht ist, Beratung und Moderation durch ZQ nachfragen. Evaluation verliert, so die Erfahrung beispielsweise im Verlauf der Umsetzung der Zielvereinbarungen, mehr und mehr den ursprünglichen Charakter eines Beobachtungs- bzw. Messinstrumentes und wird zu einer kontinuierlichen Aufgabe, die darin besteht, langfristige Anpassungs- und Entwicklungsleistungen der Fächer und Fachbereiche zu begleiten und zu unterstützen.

Um die im Rahmen der Projekte und Evaluationen gewonnenen Erfahrungen und Ergebnisse sowohl hochschulintern als auch hochschulübergreifend zugänglich zu machen, erfolgt eine systematische Auswertung und Dokumentation in der von ZQ herausgegebenen Schriftenreihe „Mainzer Beiträge zur Hochschulentwicklung“.

Das Gelingen einer Evaluation erfordert nicht nur eine intensive Mitarbeit des evaluierten Faches sowie unterschiedlicher fachinterner Abteilungen und Bereiche (Prüfungsamt, Dekanat, Institut etc.), sondern auch die Kooperation und Mitarbeit bspw. universitätsintern zuständiger Fachabteilungen. Diesem Umstand Rechnung tragend, versucht ZQ, die erforderlichen Kontakte und Verbindungen herzustellen und zwischen

den oft unterschiedlichen Perspektiven und Erfordernissen vermittelnd und klärend tätig zu werden. Darüber hinaus bemüht sich ZQ aber auch universitätsextern bzw. -übergreifend um Kontakte und Kooperationspartner, um den notwendigen Austausch über hochschulpolitische und arbeitsspezifische Fragestellungen und Problembereiche zu fördern.

3. Stand der Evaluation an der Universität Mainz

Wie die folgende Tabelle zeigt, wurden seit dem Beginn der Evaluation im Wintersemester 1995/96 in Mainz 17 Fächer aus 8 verschiedenen Fachbereichen evaluiert.

Die im Verlauf der Evaluationen erfolgten Modifikationen der Verfahrensform gründen nicht zuletzt in den mit jedem neuen Verfahren gesammelten weiterführenden Erfahrungen. Auch die Interessen der Fächer und die Erwartungen an Evaluation haben zu Änderungen und Ausbau geführt. Die heutigen Erfahrungen zeigen, und dies korrespondiert weitgehend mit den Interessen der Fächer, dass es sinnvoll ist, das Verfahren dreigliedrig zu gestalten, d.h. sowohl intern als auch extern zu evaluieren und mit Zielvereinbarungen abzuschließen.

Auch die inhaltliche Expansion hat sich, so die Erfahrungen in mehreren Fächern, bewährt und sollte beibehalten werden. Aus heutiger Sicht lässt sich sagen, dass eine konstruktive und folgengenerierende Evaluation von Studium und Lehre gleichermaßen einen Einblick in die Fach- und Forschungsorganisation erfordert.

Bereits die ersten Erfahrungen mit Evaluation machten deutlich, dass es sinnvoll ist, Evaluationsvorhaben in einen umfassenderen Kontext der Organisationsentwicklung zu stellen. Darunter ist zu verstehen, dass externe Berater Organisationen oder Teile von ihnen darin unterstützen, die eigene Betriebsblindheit und interne Blockaden zu überwinden und bei der Notwendigkeit zur Mobilisierung Hilfestellung zu geben.

So ist zum Beispiel die bei externen Gutachtern beliebte Forderung an die Fächer, eine stärkere Akzentsetzung in der Lehre vorzunehmen, um zur vielbeschworenen Profilbildung eines Faches beizutragen, in der Umsetzung nicht denkbar, ohne Entwicklungen des Personalbestandes und aktueller und zukünftiger Forschungsschwerpunkte mit zu berücksichtigen.

Übersicht über an der Johannes Gutenberg-Universität Mainz durchgeführte und laufende Evaluationen

Fach	Verfahrensform ³	Evaluationsgegenstand ⁴	Stand
Politikwissenschaft	i.+e. Eval.	StuL	Abgeschlossen
Mathematik	i. Eval.	StuL	Abgeschlossen
Physik	i. Eval.	StuL	Abgeschlossen
Chemie	i.+e. Eval., ZV	StuL	Abgeschlossen
Pharmazie	i.+e. Eval., ZV	StuL	Abgeschlossen
Geowissenschaften	i.+e. Eval., ZV	StuL	Abgeschlossen
Geographie	i.+e. Eval., ZV	StuL	Abgeschlossen
Betriebswirtschaftslehre	i.+e. Eval.	StuL	Externe Evaluation
Volkswirtschaftslehre	i.+e. Eval.	StuL	Externe Evaluation
Wirtschaftspädagogik	i.+e. Eval.	StuL	Externe Evaluation
Rechtswissenschaft	i. Eval.	StuL	Interner Evaluationsbericht
Biologie	i.+e. Eval., ZV	StuL/F	Abgeschlossen
Anthropologie	i.+e. Eval., ZV	StuL/F	Abgeschlossen
Ethnologie	i.+e. Eval., ZV	StuL/F	Interner Evaluationsbericht
Afrikanische Philologie	i.+e. Eval., ZV	StuL/F	Interner Evaluationsbericht
Deutsche Philologie	i.+e. Eval., ZV	StuL/F	Zielvereinbarung
Kulturanthropologie/ Deutsche Volkskunde	i.+e. Eval., ZV	StuL/F	Zielvereinbarung

³ Verfahrensform: internes Evaluationsverfahren = i. Eval.; internes und externes Evaluationsverfahren = i.+e. Eval.; internes und externes Evaluationsverfahren incl. Zielvereinbarungen = i.+e. Eval., ZV

⁴ Evaluationsgegenstand: Evaluation von Fachorganisation, Studium und Lehre = StuL; Evaluation von Fachorganisation, Studium und Lehre, Forschung = StuL/F

Der Vollständigkeit halber ist zu sagen, dass auch universitätsextern Evaluationen durch das ZQ unterstützt wurden. Hier sind das im Wintersemester 2000/2001 begonnene Evaluationsverfahren im Fachbereich Architektur/Innenarchitektur der Fachhochschule Kaiserslautern zu nennen, die Evaluation des Studienseminars in Koblenz, die im Wintersemester 2001/2002 angefangen hat, sowie die bereits erwähnten Programmevaluationen "Chemie-Labor" und "Schulmusik-Projekt". Die externe Nachfrage der Arbeit vom ZQ hat sich aus hochschulexternen Kontakten ergeben und wird nicht zuletzt aufgrund der damit einzuwerbenden Drittmittel durch die Hochschulleitung unterstützt.

4. Entwicklung und Ausbau des Evaluationsverfahrens

Das an der Universität Mainz angewandte Verfahren lehnt sich an das im Rahmen des erwähnten Pilotprojektes verwendete Modell von HIS an und ist im Verlauf der unterschiedlichen Fachevaluationen kontinuierlich u.a. aufgrund der Ergebnisse und der Nachfrage modifiziert und weiterentwickelt worden.

Das derzeit praktizierte Verfahren gliedert sich in einen vier- bzw. fünfphasigen Evaluationsprozess, bestehend aus Vorlauf, internem Verfahren, externem Verfahren, Zielvereinbarung und Umsetzung der Zielvereinbarungen. Vor dem offiziellen Beginn des internen Evaluationsverfahrens werden mittlerweile zwischen dem evaluierten Bereich bzw. Fach und dem ZQ schriftlich Zielsetzung, Zuständigkeiten, Zeitplan sowie Art und Umfang der Leistungen fixiert und durch das entsprechende Gremium im Fach (Fachbereichsrat, Bereichsausschuss, Institutsversammlung) verabschiedet.

Dem tatsächlichen Beginn einer Evaluation geht ein zeitlich je nach fachinterner Situation mehr oder weniger langer Vorlauf voraus, in dem geklärt und entschieden werden muss, ob, warum und mit welchem Ziel evaluiert werden soll. Diese **Phase der internen Klärung** ist für den Verlauf und das Gelingen der Evaluation von entscheidender Bedeutung. Diejenigen Fächer, die sich bereits im Vorfeld darüber verständigt haben, was sie von der Evaluation erwarten, welche Arbeit bzw. Belastung während und nach Ablauf einer Evaluation auf sie zukommt, können das Verfahren nicht nur zügiger und mit geringeren Reibungsverlusten, sondern auch für das Fach gewinnbringender durchführen.

In Fächern, die in ihrer Entscheidung zur Evaluation durch eine eher diffuse oder heterogene Motivlage zu charakterisieren sind, und in Fächern, in denen sich weite Teile oder Lager innerhalb des Faches gegen eine Evaluation ausgesprochen haben, bzw. die Entscheidung zur Evaluation nicht aktiv mittragen, wird es auch während des Verfahrens zu Diskussionen kommen, die das Verfahren mindestens verzögern. Die Empfehlung muss daher lauten, die Vorphase der Evaluation so zu gestalten, dass vor Beginn interne Klärungen und Diskussionen erfolgt sind, bzw. bei Vorliegen heterogener bzw. diskrepanter Einstellungen, diese kenntlich zu machen, damit diesem Umstand von Beginn an Rechnung getragen werden kann.

Generell läßt sich sagen, dass ein Fach auf die Klärung der Ziele, Vorstellungen, Wünsche und Motive einer Evaluation Sorgfalt aufwenden sollte, nicht zuletzt, um Frustration und Enttäuschung entgegenzuwirken, die aus einer nicht abgestimmten Vorgehensweise erwachsen können. Die Chancen, die in einer Evaluation liegen, und dies meint nicht allein Zugewinn in Form finanzieller Mittel oder personeller Ressourcen, sondern auch Nutzen bspw. in Form von Profilverdienst oder Gewinn an Klarheit in der Definition der Situation des Faches, sollten nicht unterschätzt werden. Je deutlicher Motive, Wünsche und Erwartungen im Vorfeld formuliert werden, desto zuträglicher ist dies für das gesamte Verfahren.

Das **interne Evaluationsverfahren** läßt sich bei dem in Mainz angewandten Modell in die Bereiche Datenerhebung und interne Berichterstellung untergliedern. Zur Datenerhebung gehört die Gewinnung sowohl quantitativer als auch qualitativer Daten, die über unterschiedliche Zugänge erhoben werden. Hauptbestandteil sind die mit den unterschiedlichen Statusgruppen geführten kumulativen und diskursiven Gruppengespräche. In diesen leitfadengestützten Gruppengesprächen werden Studierende des Grundstudiums, des Hauptstudiums, die Mitglieder der Fachschaft, die nichtwissenschaftlichen sowie die wissenschaftlichen Mitarbeiterinnen und Mitarbeiter und schließlich die Professoren zu ihrer Einschätzung der Studien- und Lehrsituation befragt. Kumulativ sind die Gespräche im Hinblick darauf, dass die Beurteilung der Studienanfänger mit der Einschätzung der fortgeschrittenen Studierenden abgeglichen und durch die Meinung der wissenschaftlichen Mitarbeiter ergänzt wird. Sondermeinungen werden ggf. als solche kenntlich gemacht bzw. werden im weiteren Verlauf nicht mehr weiterverfolgt.

Von jedem Gruppengespräch, wie auch von den nachfolgend erwähnten Einzelgesprächen, wird ein ergebnisorientiertes Verlaufsprotokoll gefertigt und durch die jeweiligen Gesprächsteilnehmer autorisiert. Dieses aufwendige und zeitintensive Verfahren hat sich insofern bewährt, als sich insbesondere die Studierenden, aber auch nichtwissenschaftliche und wissenschaftliche Mitarbeiter durch diese Form der Protokollierung mit ihren Einschätzungen wiederfinden.

Intern wird im ZQ - insbesondere im Hinblick auf personelle und zeitliche Ressourcen - z. Zt. überlegt, inwieweit das Verfahren nicht durch ein ergebnisorientiertes Kurzprotokoll, das auf die Abbildung von Argumentationsgängen und Mindermeinungen verzichtet, abgekürzt werden könnte. Bislang wurde aufgrund des hohen Stellenwerts, den das Protokoll insbesondere bei den o.g. Gruppen einnimmt, dies wird in Rückmeldungen immer wieder deutlich, von einer Veränderung abgesehen. Dennoch sind die Überlegungen damit nicht beendet, gerade im Hinblick auf Evaluationen in großen Fächern bzw. Fachbereichen, die eine ungleich größere Anzahl an Gesprächen erfordern, wird diesbezüglich weiter nachzudenken sein.

Neben den Gruppengesprächen, die sich hauptsächlich auf Lehr- und Studienfragen zentrieren, werden ergänzende Gruppengespräche mit Doktoranden, Habilitanden und Drittmittelbediensteten zu Lehr-, Forschungs- und Organisationsfragen geführt sowie Einzelgespräche zur Lehr- und Forschungssituation und zu Struktur- und Organisationsfragen mit Professoren des Faches. Auch diese Gespräche werden, wie oben beschrieben, protokolliert und autorisiert. Je nach Fachsituation bzw. Erfordernis finden darüber hinaus ergänzende Gespräche (Gruppen- oder Einzelgespräche) statt, beispielsweise mit externen Vertretern von Fachorganisationen.

Als wichtiges weiteres Datenmaterial, das insbesondere für die Erstellung des internen Evaluationsberichtes eine Rolle spielt, sind fachintern oder in Verwaltungseinheiten vorgehaltene Statistiken, Lehr- und Forschungsberichte zu nennen. Darüber hinaus wird beispielsweise mit der Durchführung einer Absolventenbefragung der Versuch unternommen, nicht nur die Aussagen der aktuell Studierenden ggf. zu relativieren oder zu stützen, sondern auch eine retrospektive Einschätzung des Studiums und der Fachsituation aus der Perspektive der Ehemaligen zu erhalten und eine Rückmeldung über Verbleib und Erfolg der Absolventen zu formulieren.

Vor der Beschreibung, wie mit den ermittelten Daten weiter verfahren wird, ist eine Bemerkung zum Stichwort Repräsentativität des Datenmaterials vorzuschicken. Die Erhebung von Daten in Form von Gruppen- bzw. Einzelgesprächen beansprucht keinerlei Repräsentativität in einem empirisch methodischen Sinne. Zwar werden die Studierenden für die Gruppengespräche per Zufallsauswahl ermittelt, d.h. von der Gesamtheit derer, die im 1. bis 5. Semester bzw. im 6. bis 12 Semester eingeschrieben sind, werden je 30 Studierende gelost und angeschrieben, Verzerrungen finden jedoch aufgrund der Tatsache statt, dass von diesen Studierenden wiederum nur ca. ein Drittel bzw. die Hälfte tatsächlich an dem Gespräch teilnimmt.

Über diejenigen Studierenden, die sich zu einer Teilnahme bereiterklären - in der Regel zwischen 10 und 15 Personen - und zu dem anberaumten Termin auch anwesend sind, ist zu sagen, dass es sich um für ihr Studium und ihr Fach engagierte und in ihrer Kritik und ihrem Lob differenzierte Studierende handelt, die das Gesprächsangebot keinesfalls als Plattform zur Pauschalbeschwerde interpretieren. Die im Vorfeld häufig vorgebrachte kritische Vermutung seitens einiger Professoren, die Aussagen der Studierenden seien aufgrund dieser nicht kontrollierbaren Auswahl auch nicht verallgemeinerbar, sind u.E. aufgrund der bisherigen, oben beschriebenen Eindrücke und Erfahrungen nicht zutreffend.

Die Annahme, dass sich insbesondere Studierende, die ihr Studium ernst nehmen und mit Engagement und Interesse vorantreiben, die Zeit für das jeweilige Gruppengespräch nehmen, hat sich ausnahmslos bestätigt. Die Ernsthaftigkeit und Fairness, mit der Studierende in diesen Gesprächen über Defizite und Beklagenswertes Auskunft geben, hat Vorbildcharakter. Insofern sollten auch die Beurteilungen der Studierenden, deren Lob und Kritik in Bezug auf die Studien- und Lehrsituation, Beachtung finden. Zusätzlich ist anzumerken, dass die Urteile der Studierenden häufig in den Gesprächen mit den wissenschaftlichen Mitarbeiter nachdrücklich bestätigt werden und letztlich auch im abschließenden Gespräch mit den Professoren unwidersprochen bleiben.

Repräsentativität, so könnte man sagen, wird bei diesem Vorgehen nicht über die Anzahl der Befragten, als vielmehr aufgrund des kumulativen Verfahrens hergestellt. Nicht nur die Gruppengespräche als solche bauen aufeinander auf, auch das erhobene Datenmaterial insgesamt ergänzt sich und ergibt dadurch eine gewisse Plausibilität. Neben den Gruppengesprächen werden ergänzende Einzelgespräche mit den Professoren des Faches geführt. Bei kleinen bzw. mittleren Fächern können alle Professoren berücksichtigt werden, in großen Fächern, in

denen dies aufgrund der Anzahl sowie des Zeit- und Personalbudgets nicht erfolgen kann, werden neben Funktionsinhabern per Zufallsauswahl ergänzende Personen ausgewählt.

Zusätzlich zu den Gesprächen werden durch das ZQ per Fragebogen Studierendenbefragungen zum Thema Studienmotivation sowie eine Absolventenbefragung durchgeführt und ausgewertet. Ferner wird das in einem Fach verfügbare Informationsmaterial (Lehr- und Forschungsberichte, Strukturplan, Informationen zur Personal- und Sachmittel-Ausstattung, Studien- und Prüfungsordnungen u.a.m.), das die über die Gespräche erhobenen Daten ergänzt und kommentiert, ebenfalls berücksichtigt.

Alle erhobenen Daten werden in dem internen Evaluationsbericht zugänglich gemacht und entsprechend aufbereitet. Die Ergebnisse der Gruppengespräche, dokumentiert über die Protokolle, werden, neben einem 1:1 Abdruck im Anhang des Berichtes, im Bericht selbst vorgestellt und diskutiert. Die autorisierten Protokolle der Einzelgespräche werden aus datenschutzrechtlichen Gründen nicht in dieser Form wiedergegeben. Die Aussagen, Meinungen und Einschätzungen der Gesprächsteilnehmer werden in diesem Fall zu einer Synopse zusammengefaßt, deren Aufgabe es ist, die in den Einzelgesprächen vorgefundene Vielfalt oder Einhelligkeit, Homogenität oder Heterogenität zu den relevanten Fragen (Studium und Lehre, Forschung, Struktur, Profil etc.) gesprächsnah abzubilden.

Parallel zur Berichterstellung sollte eine öffentliche Sitzung stattfinden, an der alle Statusgruppen teilnehmen können und in der die bisherigen Ergebnisse sowie das weitere Verfahren vorgestellt werden. Zunächst soll den einzelnen am internen Verfahren beteiligten Statusgruppen damit eine Rückmeldung über die Ergebnisse gegeben werden, zusätzlich kann damit auch die Zeit bis zur externen Evaluation überbrückt werden, da mit Beendigung der Gespräche bis zum Termin der externen Begehung für die Mehrheit der Mitglieder des Faches keine Evaluationsaktivitäten mehr sichtbar sind.

Darüber hinaus ist der Informationsfluss innerhalb eines Faches aufgrund von Größe und internen Hierarchien nicht in allen Fächern in dem Maße möglich bzw. üblich, wie dies nicht nur für ein gelungenes Evaluationsverfahren wünschenswert und erforderlich wäre. Die öffentliche Vorstellung der Ergebnisse und die Klärung des weiteren Verfahrens für alle Fachmitglieder hilft, in die notwendige Kommunikation auch diejenigen einzubeziehen, die nicht mittels

Gremienmandat über das Verfahren informiert sind, und bietet die Möglichkeit, den notwendigen Dialog im Fach zu stabilisieren.

Nicht in allen Fächern ist es zu einer solchen öffentlichen Präsentation der Ergebnisse gekommen. Die Gründe hierfür liegen überwiegend in einer im jeweiligen Einzelfall nachvollziehbaren Mischung aus Zögerlichkeit, Zeitknappheit und Vermeidung von Konflikten in der Öffentlichkeit. Dennoch bleibt festzuhalten, dass diese einen für das Gesamtverfahren sinnvollen Einzelschritt darstellt.

Dem internen Evaluationsbericht sind primär zwei Funktionen zuzuordnen: zum einen stellt er eine schriftliche Bündelung und Definition der Situation aus der Sicht des Faches dar, zum anderen bietet er die Grundlage für die externe Evaluation für die Fachgutachter, die sich anhand des Berichtes einen ersten Eindruck über das zu evaluierende Fach, dessen Probleme und Perspektiven verschaffen. Insbesondere die zweite Funktion rechtfertigt daher auch die Ausführlichkeit der Berichte. Neben einer knappen Darstellung der Gründe und Ziele der Evaluation beinhaltet ein Evaluationsbericht u.a. Ausführungen zu den Themenbereichen Geschichte und Organisation des Institutes bzw. Fachbereichs, finanzielle, räumliche und personelle Ressourcen, Lehr- und Studienschwerpunkte sowie -ziele, Forschungsschwerpunkte und -ziele, quantitative Daten (u.a. Studienanfänger, Absolventen, Prüfungsstatistiken), Praxis von Lehre und Forschung, Kooperationen mit anderen nationalen und internationalen Institutionen sowie einen zukunftsgerichteten Ausblick über die Entwicklung des Faches, in dem offene Fragen und strittige Punkte benannt werden.

Nur wenn die Fächer eine Zuspitzung ihrer Definition der Situation in der Art betreiben, dass sie benennen können, wie der Stand im Fach ist, welche Fragen unbeantwortet sind, welche Probleme schon mehrfach unter wechselnden Zuständigkeiten ungelöst geblieben sind und zu welchen Punkten sich ein Fach Hilfestellung seitens der Fachgutachter erwartet, können sich diese wiederum gezielter vorbereiten und entsprechende Gedanken machen.

Da die Erstellung eines solchermaßen konzentrierten internen Evaluationsberichtes für die meisten Fächer einen erheblichen Aufwand darstellt, der nicht kostenneutral zu bewältigen ist, hat das ZQ einen Modus entwickelt, der vorsieht, dass das Fach eine Person benennt, die für einen Zeitraum von 4 Monaten in Höhe eines BAT IIa-Gehaltes aus universitären Mitteln finanziert wird, um den Bericht zu erstellen. Da es in dem Bericht nicht allein darum gehen kann, Sachverhalte zu

beschreiben, sondern auch Gewichtungen vorzunehmen oder unterschiedliche Perspektiven zu schildern, wäre eine Mitarbeiterin oder ein Mitarbeiter allein mit dieser Aufgabe überfordert. Daher sieht das Verfahren vor, das zwei Fachprofessoren sich bereiterklären, als Ansprechpartner zu fungieren und die Verantwortung zu übernehmen. ZQ berät die Mitarbeiterin bzw. den Mitarbeiter bei der Berichterstellung und ermutigt zu einer realitätsnahen, ungeschönten Berichterstattung.

Je pointierter die tatsächlichen Fragen, Probleme und etwaige Differenzen in einem Fach zur Sprache kommen, desto gezielter können die externen Gutachter ihrer Aufgabe nachkommen, Empfehlungen auszusprechen und Lösungswege aufzuzeigen.

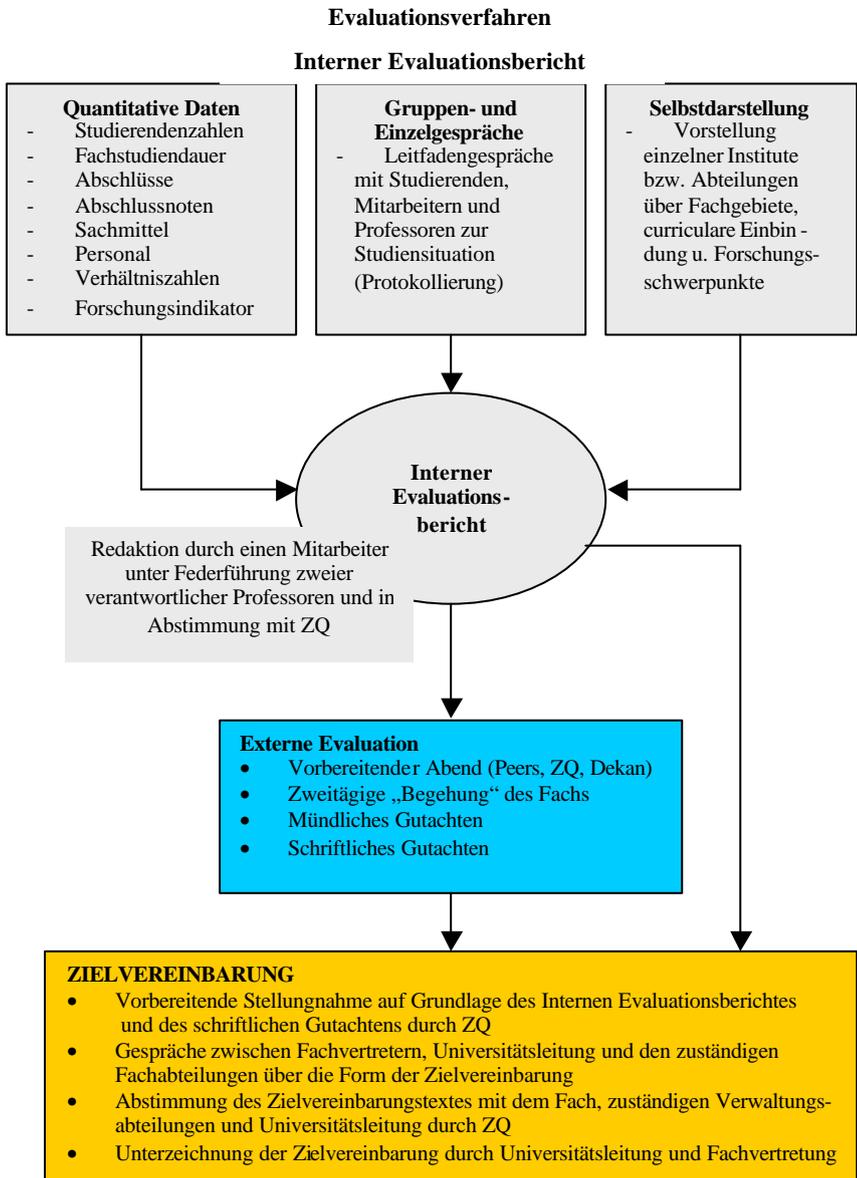
Nicht alle Fächer machen von dem Angebot Gebrauch, mit den zur Verfügung gestellten Mitteln eine Mitarbeiterstelle zu besetzen bzw. aufzustocken. Häufig wurde der interne Evaluationsbericht von einem bzw. zwei Professoren selbst verfasst und die verfügbaren Mittel wurden kompensatorisch zur Entlastung in Form von Hilfskraftmitteln verwendet. Auch dieses Modell kann sinnvoll sein. Überwiegend besteht hierbei jedoch die Gefahr, dass der tatsächliche Aufwand unterschätzt wird und die Abfassung des Berichtes im laufenden Geschäft nur mit Verzögerung erfolgt.

Auch der Modus der Berichterstellung durch Mitarbeiter kann sich nachteilig auswirken, bspw. dann, wenn nicht nur der Arbeitsvertrag sondern auch die Arbeitsaufgaben aufgestockt werden bzw. wenn sich nicht wirkliche zeitliche Freiräume sichern lassen.

Aufgrund der bisherigen Erfahrungen mit den beiden Modellen plant ZQ, versuchsweise die Erstellung des Berichtes selbst vorzunehmen, flankiert von definierten Ansprechpartnern aus allen Statusgruppen und unterschiedlichen Bereichen/Abteilungen/Schwerpunkten des Faches. Dies bietet sich in Fächern oder Fachbereichen insbesondere dann an, wenn aufgrund der internen Situation bereits die Benennung eines Mitarbeiters oder Professors dazu führen würde, dass andere Fachmitglieder sich durch diese Wahl nicht mehr neutral vertreten sehen würden.

Nach Erstellung, Diskussion und ggf. Modifikation des internen Evaluationsberichtes erfolgt die Verabschiedung durch das zuständige Gremium, i.d.R. der Fachbereichsrat bzw. die Institutsversammlung. Mit der Verabschiedung ist die Phase der internen Evaluation abgeschlossen.

Die folgende Abbildung fasst die beschriebenen und noch zu beschreibenden Bausteine des Verfahrens zusammen:



Parallel zur Diskussion und evtl. Modifikation des Berichts wird die zweite Phase, die **externe Evaluation** vorbereitet. Die Auswahl der Gutachter wird durch das Fach vorgenommen. Je nach Größe und inhaltlicher Differenzierung werden vier bis sechs Gutachter ausgewählt, die durch den Präsidenten der Hochschule eingeladen werden.

Die Benennung der Fachgutachter sollte insbesondere die fachliche Ausgewiesenheit bzw. das Renommee der Wissenschaftler berücksichtigen. Häufig werden daher DFG-Gutachter benannt, die zusätzlich - dies erweist sich im externen Verfahren als nicht unerheblich - über Kenntnisse und Routinen des Gutachtens verfügen. Fraglose fachliche Kompetenz seitens der Gutachter ist schon allein deshalb vonnöten, damit den nicht zwangsläufig angenehmen und für alle Beteiligten vorteilhaften Vorschläge der Gutachter ein gewisses Gewicht und Objektivität zugemessen werden kann.

Nicht für alle Fächer, darüber wird unter dem Punkt Zielvereinbarungen noch zu sprechen sein, verbindet sich mit der Evaluation neben dem immateriellen Nutzen auch ein Gewinn personeller oder finanzieller Art. Um so wichtiger ist es, dass die anstehenden Entscheidungen in einem Fach, bspw. über dessen zukünftige Struktur, mit Empfehlungen seitens der Gutachter begleitet werden können, denen nicht nur die Befürworter Plausibilität und Sinn einräumen können.

Auch unter diesem Aspekt empfiehlt ZQ, sich im Fach rechtzeitig über die Frage der Gutachter zu einigen und zu versuchen, hochkarätige Wissenschaftler zu gewinnen. Die Gutachtertätigkeit kann seitens der Hochschule nur mit einer symbolischen Anerkennung in Form einer Aufwandsentschädigung von DM 1000,00 (zuzgl. Unterbringung, Fahrtkosten, Verköstigung) erfolgen. Dennoch, so die bisherigen Erfahrungen der Fächer, kann es gelingen, ausgewiesene Fachwissenschaftler für diese Arbeit zu motivieren.

Das bisweilen im Zusammenhang mit der Auswahl der Gutachter gehörte Argument, man könne es den Fächern nicht überlassen, selbst Gutachter zu benennen, da dies lediglich dazu führe, ehemalige oder befreundete Kollegen aufzufordern, diese Tätigkeit zu übernehmen, trifft nicht zu. Selbst in den Fällen, in denen sich ein Fach aus unterschiedlichen Gründen bspw. dafür entschieden hat, einen ehemaligen Kollegen zu bitten, war festzustellen, dass dies auch von Vorteil sein kann, insofern als dieser von seit langem thematisierten, aber ungelösten Fragen und Problemen wusste.

In den meisten Gutachtergruppen ist eine positiv zu beurteilende Gruppendynamik derart entstanden, dass die Gutachter, gleich ob bereits mit der Hochschule vertraut oder nicht, ihrer Gutachterrolle gern und dort wo notwendig auch mit dem erforderlichen Nachdruck nachgekommen sind sowie ihren gutachterlichen Standpunkt mit Entschiedenheit vertreten haben. Die Balance zwischen fachlicher Nähe und anwaltschaftlicher Fürsprache und erforderlicher Distanz gegenüber bekannten oder gar vertrauten Fachkollegen hat sich in der Regel eingestellt.

Es hat auch Fälle gegeben, in denen sich einzelne Gutachter den an sie ergangenen Auftrag nicht zu eigen gemacht haben. Bedauerlich, und für das Verfahren nicht folgenlos, war dies insbesondere dann, wenn weitere Faktoren, bspw. die fachinhaltliche Aufteilung der Gutachtergruppe, dazu geführt haben, dass eine Kompensation des Ausfalls nicht möglich war.

Der Ablauf der externen Begehung erfolgt an zwei Tagen, vorzugsweise Montag und Dienstag, mit einem Auftaktgespräch am Sonntagabend. Diese Festlegung erscheint in doppelter Hinsicht sinnvoll. Zum einen sollte eine externe Begehung an Tagen stattfinden, an denen aufgrund des Vorlesungsbetriebes auch zahlreiche Mitglieder aller Statusgruppen anwesend sind. Außerdem gestattet die sonntägliche Anreise mit einem Auftaktgespräch am Vorabend eine entspanntere Heransgehensweise. Das Auftaktgespräch am Sonntagabend hat neben der Begrüßung durch Fachvertreter (i.d.R. Dekan und Geschäftsführender Leiter) auch die Funktion der Vorstellung des von ZQ entwickelten Verfahrens, der Besprechung des Ablaufs der externen Begehung, der Verabredung einer Aufgabenverteilung unter den Gutachtern (Moderation der Gespräche, Koordination des schriftlichen Gutachtens, Vorstellung des mündlichen Gutachtens etc.) sowie der Möglichkeit des Austauschs der Gutachter hinsichtlich einer Einschätzung und Bewertung des Fachs auf der Basis des mit dem im internen Evaluationsbericht vorgelegten Materials.

Der Ablauf der externen Begehung verläuft in den Grundzügen nach einem bereits bewährten Plan, wobei Veränderungen auf Wunsch des Faches bzw. der Gutachter möglich sind. Es ist vorgesehen, dass neben terminierten Gruppengesprächen mit den unterschiedlichen Statusgruppen Gelegenheit zu Einzelgesprächen mit den Lehrenden oder in kleineren Gruppen besteht, um die sich auf Grundlage des Berichtes ergebenden Fragen und Situationsbeschreibungen zu pointieren oder zu erweitern. Ferner findet eine Begehung der Räumlichkeiten statt und es wird die erforderliche Zeit für die Gutachtergruppe eingeräumt, sich intern über das Vorgefundene zu verständigen und zu beraten sowie die Vorstellung

des mündlichen Gutachtens und das Procedere für die Gutachtenerstellung vorzubereiten.

Auf Wunsch der Gutachter hat sich ergeben, dass ZQ auch an den internen Beratungen der Gutachtergruppe teilnimmt. Dies war anfänglich nicht der Fall. In den drei Fällen, in denen die Anwesenheit nachgefragt wurde, hat sich gezeigt, dass die Stellung von ZQ allgemein als beratende neutrale Instanz jedoch mit den durch das vorausgegangene interne Verfahren hilfreichen Kenntnissen über die Fachsituation auch an dieser Stelle für das Verfahren sinnvolle Unterstützung leisten kann.

Die externe Begehung vor Ort endet mit der mündlichen Präsentation des Gutachtens durch die Gutachtergruppe. Daran schließt sich die Erstellung des Gutachtens an, für die ca. acht Wochen anberaumt werden. Bewährt hat sich auch hier, dass das Gutachten vorab an ZQ sowie das Fach zugeschickt wird. Diesem Vorgehen liegt nicht die Intention einer kontrollierenden Eingriffsmöglichkeit und Vorabsprache zugrunde; es soll vielmehr dem Fach und den Gutachtern im Vorfeld eine zusätzliche Möglichkeit bieten, evtl. entstandene Mißverständnisse auszuräumen, Informationen nachzuliefern oder offen gebliebene Fragen zu stellen.

Mit dem Eintreffen des schriftlichen Gutachtens beginnt die dritte Phase des Evaluationsverfahrens, der Prozess der **Zielvereinbarung**. Hierzu wird von ZQ eine Empfehlung an das Fach in Form einer zusammenfassenden Stellungnahme zur internen und externen Evaluation mit Vorschlägen zur Umsetzung der Evaluationsergebnisse verfasst. Diese Stellungnahme wird fachintern und ggf. unter Hinzuziehung von ZQ diskutiert.

Ziel dieser Diskussionen ist es, verbindlich zu definieren, welche strukturellen und inhaltlichen Festlegungen ein Fach bzw. Studiengang oder Fachbereich für welchen Zeitraum und mit welchen Ressourcen treffen möchte. Verbindlichkeit wird u.a. dadurch hergestellt, dass diese Vereinbarung mit der Hochschulleitung geschlossen wird und eine Überprüfung der vereinbarten Maßnahmen nach einem geeigneten Zeitraum ebenfalls Gegenstand der Vereinbarung ist.

Zeitlich versetzt zu der fachinternen Diskussion findet eine Diskussion des Entwurfs in der Hochschulleitung und den involvierten Abteilungen der Verwaltung (i.d.R. Personal-, Haushalts- und Planungsabteilung) statt. Nach entsprechender Vorklärung und ggf. Modifikation der Vereinbarungsinhalte findet ein abschließendes gemeinsames Gespräch zwischen den beiden Vertragspartnern statt, daran schließt sich die Unterschrift der Zielvereinbarung an.

Die Rolle der Hochschulleitung im Rahmen einer Zielvereinbarung besteht insbesondere darin, universitäre Ziele und Sichtweisen zu formulieren und gegenüber dem Fach deutlich zu machen. Diese Aufgabe bringt, neben Ermutigung und Unterstützung - nicht zuletzt in Form von finanziellen Anreizen - bisweilen auch für die Fächer unangenehme Auflagen mit sich. Dennoch ist es in bisher fünf Fächern gelungen, eine Zielvereinbarung abzuschließen.

Die bislang geschlossenen Zielvereinbarungen unterscheiden sich nicht nur infolge unterschiedlicher Zielsetzungen, die das Evaluationsverfahren selbst geprägt haben. Auch die spezifische Situation und der Reformwille eines Faches sind zu berücksichtigen. Neben der inhaltlichen Festlegung, ob Evaluation neben Fragen zu Studium und Lehre auch die Bereiche Forschung und Organisation des Faches beinhalten soll, ergeben sich auch Differenzen in der Reichweite der möglichen Festlegungen bspw. im Hinblick auf die im Fach möglichen Spielräume, seien sie personeller oder sachlicher Art.

Auch der Zeitraum, der für ein Verfahren, von Beginn der internen Evaluation bis zum Abschluss der Zielvereinbarungen zu veranschlagen ist, gestaltet sich unterschiedlich lang. Insbesondere in Fällen, in denen über weitreichende und einschneidende Strukturfragen diskutiert und entschieden werden muss, wird der veranschlagte Zeitraum von ca. einem Jahr für das gesamte Verfahren nicht immer ausreichen.

Diskutiert wird daher im ZQ, inwieweit man in solchen Fällen nicht Strukturfragen aus dem Verfahren der Zielvereinbarung ausgliedert und eine gesonderte Vereinbarung zum Zeitpunkt der Klärung trifft. Vorteilhaft daran wäre, dass die übrigen Vereinbarungen bereits umgesetzt werden könnten, nachteilig daran erscheint jedoch die Zergliederung des Verfahrens.

Aufgrund der Unterschiedlichkeit der vereinbarten Maßnahmen und Ziele, die naturgemäß auch in der Umsetzung zeitlich differieren, wird eine Reevaluation der verabredeten Punkte nach einem Zeitraum von fünf Jahren avisiert. Erfahrungen mit diesem Verfahrensschritt stehen jedoch noch aus.

5. Funktion und Zielsetzung des Evaluationsverfahrens

Obleich die Klärung dessen, was Zielsetzung von Evaluation sein kann bzw. soll, auch die Berücksichtigung hochschulpolitischer Interessen und Implikationen erforderlich macht, soll dies hier nicht weiter verfolgt

werden.⁵ Bevor der Frage nachgegangen wird, welche Funktion einer Evaluation aufgrund der bislang gewonnenen Erfahrungen aus der Perspektive der evaluierenden Einheit zukommt, sollen kurz mögliche Interessen der evaluierten Einheit skizziert werden.

Für einige Fächer stellt Evaluation ein probates Mittel der Organisationsentwicklung dar, d.h. es werden Fakten und Informationen über die eigene Situation auf den Prüfstand gestellt und Überlegungen konkretisiert, welchen Aspekten und Überlegungen im Ringen um eine Festlegung der zukünftigen Entwicklung eines Faches der Vorzug zu geben ist.

Für andere scheint Evaluation mehr ein dem Zeitgeist entsprechendes Instrument zu sein, das bestenfalls dazu dient, Informationen über die Studiensituation oder die Zufriedenheit von Mitarbeiterinnen und Mitarbeitern zu gewinnen. Die Erwartungen an eine Evaluation lassen sich in diesen Fällen nur sehr vage umschreiben; einerseits soll sie Fakten liefern und vorfindbare Situationen beschreiben, andererseits wird dem Instrument Evaluation keineswegs hohe Kompetenz zuerkannt. Letztere Einschätzung hat, so ist zu vermuten, in einigen Fächern überhaupt erst den erforderlichen Beschluss zur Evaluation ermöglicht. Vertraulich teilte uns ein Dekan nach Abschluss eines Verfahrens mit, dass er sicherlich niemals die Zustimmung der Kollegen hätte bekommen können, wenn von vornherein klar gewesen wäre, wie umfassend die Studien-, Lehr- und Fachsituation beleuchtet werden würde.

Vereinfachend kann man sagen, dass sich die Fächer im Bemühen um eine angemessene Positionierung im Wettstreit um knappe Ressourcen zur Evaluation bereit erklärt haben, wenn die Nachfrage bei bereits evaluierten Fächern eine günstige Beurteilung des Verfahrens ergab und fachinterne Zweifel und Widerstände überstimmt werden konnten.

Welche Aufgabe, welche Funktion kann und soll Evaluation, gemessen an den bisherigen Erfahrungen, erfüllen? Im hier verstandenen Sinne ist Evaluation zu beschreiben als ein Prozess, der eine mittelfristige Entwicklung und Sicherung von teils generalisierbaren, teils spezifischen Qualitätskriterien in jedem Verfahren aufs neue ermöglichen muss. Zielsetzung ist, ein Fach in der Anstrengung zu unterstützen, die genuinen Aufgabenbereiche Forschung, Lehre und Administration/Organisation in der alltäglichen Praxis zu beschreiben und zu überdenken, Zielsetzungen

⁵ vgl. hierzu u.a. Katja AHLSTICH und Tanja URICH-NEITZERT sowie Ulrich DRUWE in diesem Band.

zu definieren, Veränderungsbedarf zu artikulieren und für bislang nicht angegangene Probleme Lösungen zu entwickeln.

Das Verfahren der Evaluation soll die Aufgabe der Mobilisierung eines Faches erfüllen bei gleichzeitiger Unterstützung und Beratung. Dies klingt wenig spektakulär und nicht selten ist zu hören, dass die zusammengetragenen Daten, Fakten, Materialien und Aussagen bereits bestens bekannt waren, und insofern keine Neuigkeit darstellen. Neu ist jedoch, dass all diese bereits intern in unterschiedlichen Zirkeln kursierenden Einschätzungen, mit Unterstützung von fachexternen Beratern systematisch zusammengetragen, in Form eines Berichtes dokumentiert und dem Urteil externer Fachgutachter übergeben werden. Hinter diesen Stand der Diskussion kann nun nicht mehr zurückgegangen werden.

Gerade die Aussage, Vieles sei bereits bekannt, zeigt, dass Katalysatoren in Form externer Berater und Gutachter, nicht unmittelbar Involvierter, benötigt werden, um das Stadium der immer gleichen Argumente und Gegenargumente, warum etwas nicht funktionieren könne oder immer schon so war, zu überwinden. Die in einigen Fächern virulente Frage, wie sich die Ausrichtung des Faches inhaltlich und damit auch personell verändern solle oder müsse, ist meist so komplex und heikel, dass die vermittelnde Unterstützung eines neutralen Dritten erforderlich erscheint.

ZQ nimmt eine solche, als „kuratorische Funktion“ zu beschreibende Aufgabe wahr. Es wird tätig in einem Feld, in dem die Rolle einer „intermediären Instanz“ insofern zweckmäßig erscheint, als Selbstheilungskräfte in den meisten Fällen am besten mobilisiert werden können durch externe Impulse.⁶

Resümierend ist zu sagen, dass es in vielen Fächern gelungen ist, manchmal schneller, manchmal etwas zäher, einen Prozess der Reflektion in Gang zu setzen. Neben der Motivation des jeweiligen Faches ist die an ZQ übertragene Aufgabe der Unterstützung und der Beobachtung hierfür nicht unwesentlich. Mit der Möglichkeit, nachfragend einen Einblick in Routinen und Abläufe zu erhalten, die wiederum auf den Prüfstand gestellt werden, eröffnet sich die Chance zur Definition dessen, was als Gut zu sichern ist und in welchen Bereichen Veränderungs- oder Entwicklungsbedarf besteht.

⁶ vgl. hierzu Manfred HENNEN in diesem Band.

Die im Verlauf jedes einzelnen Evaluationsverfahrens gewonnenen Erfahrungen erweitern die Schnittmenge von bereits diskutierten und übertragbaren Problemen und Fragen sowie angeführten Argumenten. Dies verhilft in jedem weiteren Verfahren dazu, den Fundus an erarbeiteten Empfehlungen und Lösungen zu erweitern und zugänglich zu machen.

6. Ergebnisse und Erfahrungen der Evaluationsverfahren

Die folgende Zusammenfassung der aus den Evaluationsverfahren gewonnen Ergebnisse und Erfahrungen befasst sich, neben einer resümierenden Beurteilung sowohl des Verfahrens als auch der durchgeführten Verfahren, abschließend auch mit den Grenzen einer Evaluation nach dem „Mainzer Modell“.

Eine Beurteilung des Verfahrens ist notwendigerweise abhängig von der Perspektive des Beurteilenden. Hier kann daher weder die Sicht der Evaluierten, sprich der Fächer, Studiengänge bzw. Fachbereiche, noch der Hochschulleitung wiedergegeben werden, sondern die Einschätzung der das Verfahren begleitenden und mit der Durchführung betrauten Organisation.

Die im Verfahren selbst angelegten Unterschiede nach Form und Gegenstand sind bereits unter 2.4 dargestellt, und die daraus abzuleitenden Empfehlungen erläutert worden. Das Verfahren der internen und externen Evaluation, gekoppelt mit Zielvereinbarungen, führt nach Einschätzung von ZQ zum überzeugendsten Ergebnis. Überzeugend insofern, als diese Art des Vorgehens zusätzlich zu der Eigenmotivation eines Faches eine externe Hilfestellung und Stabilisierung ermöglicht und eine wohltuende Konkretisierung der Ergebnisse und Entscheidungen in Form von demokratisch erarbeiteten – bisweilen auch erstrittenen - Zielvereinbarungen mit sich bringt, hinter die nicht mehr zurückgegangen werden kann und mit der ein Fach eine mindestens mittelfristige, durch die Hochschulleitung getragene Entwicklungsperspektive gewinnt.

Aus o.g. Einschätzung ist jedoch keine Zwangsläufigkeit derart abzuleiten, dass zukünftig ausschließlich nach dieser Form evaluiert werden sollte und inhaltlich nur dann ein gutes Ergebnis zu erzielen ist, wenn ausnahmslos Studium, Lehre, Forschung und Organisation der evaluierten Einheit berücksichtigt werden. Sicherlich wird es hier auch weiterhin Abweichungen von dieser Empfehlung geben, die sich in erster Linie an

der durch das Fach zu definierenden Zielsetzung der Evaluation orientieren muss.

Wie sind nun die begleiteten einzelnen Verfahren, notwendigerweise anonym und daher abstrakt, zu beurteilen? In drei von 17 Fachevaluationen wurde ausschließlich intern und mit Konzentration auf die Studien- und Lehrsituation evaluiert. Aus Sicht der Fächer konnte in diesen Fällen das Ziel bzw. der Zweck der Evaluation insofern erreicht werden, als dadurch an die Lehrenden die gewünschte Rückmeldung erfolgen konnte, wo insbesondere die Studierenden „der Schuh drückt“; damit wurde fachintern eine Diskussionsgrundlage geschaffen, auf deren Basis weitere Maßnahmen beschlossen und Entscheidungen gefällt werden können.

In vier Fächern hat sich an die interne Evaluation nur ein externes Verfahren angeschlossen. In einem dieser vier Fächer war zum Zeitpunkt der Evaluation der Schritt der Zielvereinbarung noch nicht eingeführt. In den anderen drei Fächern bestand zum Zeitpunkt des fachbereichsinternen Beschlusses zur Evaluation zunächst nur das Interesse an interner Evaluation. Nach Abschluss dieser Phase wurde, nach erneuter fachinterner Diskussion, ein externes Verfahren angehängt. Aufgrund der für diese Fächer dann jedoch unbefriedigenden Erfahrung mit der externen Evaluation wurde fachintern der Wunsch artikuliert, auf Zielvereinbarungen zu verzichten.

Im Falle der sechs abgeschlossenen Verfahren, die sowohl eine interne und externe Evaluation als auch Zielvereinbarungen vorsahen, kann aus unserer Sicht von erfolgreichen Verfahren gesprochen werden, wobei erfolgreich hier keineswegs mit reibungslos gleichzusetzen ist. Die Verfahren, nicht zuletzt infolge unterschiedlicher Ausgangssituation, Motivation und internem Konsens bzgl. der Notwendigkeit zu Reorganisation und Veränderung, konnten nicht immer ohne Hürden und Unwägbarkeiten vorangetrieben werden. In vier weiteren Fächern ist das Verfahren noch nicht abgeschlossen und bleibt daher unkommentiert.

Für eine präzisere Beurteilung des sen, was über den Erfolg eines Verfahrens mitentscheidet, müssen, zusätzlich zu Verfahrensgegenstand und Verfahrensform, auch Aspekte des jeweiligen Verfahren selbst und damit die unterschiedliche Situation der evaluierten Fächer sowie deren spezifische Motivation einbezogen werden.

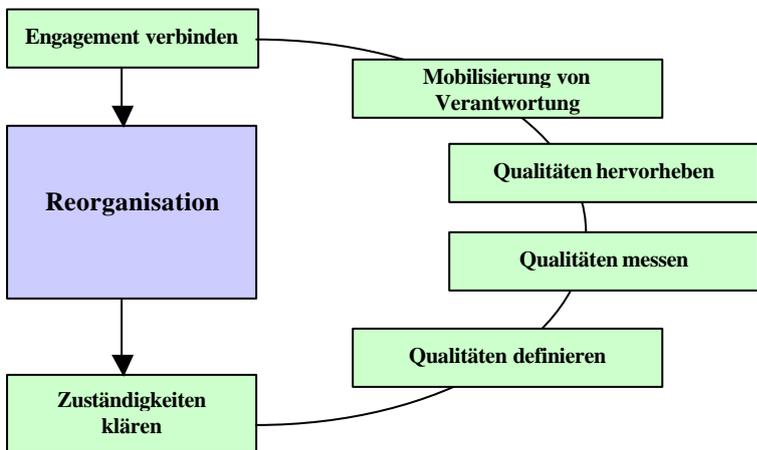
So gab es sowohl Fächer, deren Situation aufgrund bereits bekannter Eckdaten im Rahmen des maßgeblichen Personalbemessungskonzeptes ungünstig anzusehen war, als auch Fächer, die hier tendenziell mehr zu den Gewinnern zu rechnen sind. Wieder andere konnten sich in diesem

Punkt insofern unbelastet zeigen, als weder hoher Gewinn noch einschneidender Verlust als Evaluationskonsequenz zu erwarten war.

In den Fällen in denen es gelungen ist, die Fächer zu mobilisieren, haben auch die Fächer oder Bereiche, die letztlich unter dem Strich abgeben mussten, nach eigener Aussage von dem Verfahren profitieren können. Insbesondere dann, wenn der Mut zur Definition der eigenen Situation und der Wille zu notwendiger Veränderung und Reorganisation auch durch die Gutachter und die Hochschulleitung verstärkt werden konnte. Nicht von allen Fächern ist diese Chance gleichermaßen erkannt und genutzt worden.

Die Motivation⁷ der Fächer und die Möglichkeit der Mobilisierung durch das Verfahren spielen jedoch eine, wenn nicht die zentrale Rolle bei der Beantwortung der Frage, welches Potential, welche Möglichkeiten aber auch Grenzen eine Evaluation nach dem „Mainzer Modell“ beinhaltet. In Fächern, in denen der Entschluss zur Evaluation mehrheitlich dem Wunsch entspringt, nicht zu den Letzten zu gehören, die sich dem Verfahren unterziehen oder in Fächern, in denen ein vermeintlich starker Dekan die Entscheidung zur Evaluation zwar durchsetzen konnte, zu unterschiedlichen Zeitpunkten des laufenden Verfahrens aber von einflussreichen Mitgliedern des Fachs mit Ausstieg oder Widerspruch konfrontiert wird, bleibt eine das Verfahren stabilisierende Mobilisierung aus.

Zirkel institutioneller Motivation



⁷ vgl. hierzu auch Manfred HENNEN (2001): *Hat Wissenschaft Macht?*, 171.

Es wäre zu kurz gegriffen, wollte man allein an monetären Eckdaten oder personellem Zugewinn bzw. Verlust Gewinner und Verlierer eines Evaluationsverfahrens beschreiben. Die Kraft, sich der Diskussion um Fachstrukturen und -profile, Studien- und Ausbildungsziele, Leistungsanforderungen und Betreuungserfordernisse, anspruchsvolle Forschung, gute Lehre und effiziente Organisation zu stellen, ist auch von in den Fächern entstandenen Kulturen und Mentalitäten abhängig und davon, ob es gelingt, zur Übernahme von Verantwortung zu mobilisieren und institutionelle Motivation zu erzeugen.

Ausgewählte Ergebnisse der Evaluation unter systemischer Perspektive

Uwe Schmidt

1. EINFÜHRUNG.....	48
2. UNTERSCHIEDLICHE ZUGÄNGE ZUR SYSTEMATISCHEN BESCHREIBUNG DER PROBLEMBEREICHE IN HOCHSCHULEN ...	48
2.1 Anpassungskrise der Universität.....	48
2.2 Organisations- und Steuerungskrise der Universität	52
2.3 Qualitäts- und Effizienzkrise der Universität	53
3. UNIVERSITÄT ALS SYSTEM	54
3.1 Theoretische Vorüberlegungen.....	54
3.2 Systemebenen der Universität.....	60
4. EVALUATIONSERGEBNISSE UNTER DER PERSPEKTIVE VON SYSTEMPROBLEMEN DER UNIVERSITÄT.....	62
4.1 Handlungsziele in Forschung und Lehre	63
4.2 Die Dimension der Ressourcen.....	67
4.3 Die Dimension der Integration.....	73
4.4 Die Dimension des Selbstverständnisses	78
5. ZUSAMMENFASSUNG	80

1. Einführung

Die These von einer ‚Krise der Hochschulen‘ wird in der öffentlichen und hochschulpolitischen Diskussion zumeist durch Einzeldiagnosen zu unterschiedlichen Problemfeldern gestützt. Zum Teil, so lässt sich feststellen, wird hierbei auf eine detaillierte Analyse und Beschreibung der unterstellten Defizite verzichtet, werden Lösungsansätze diskutiert, ohne dass eine systematische Auswertung der jeweiligen damit verbundenen Implikationen und Problembereiche erfolgt wäre. Die Erfahrungen aus den Evaluationen nach dem Mainzer Modell bieten neben anderen Instrumenten der Hochschulforschung einen möglichen Zugang zur Erhebung von möglichen Defiziten in Forschung, Studium und Lehre.

Betrachtet man die in den vorangegangenen Jahren als zentral diskutierten ‚Krisenphänomene‘, so lassen sich diese im Wesentlichen drei Perspektiven zuordnen. Hierbei handelt es sich um die Thesen einer Anpassungskrise, einer Organisationskrise sowie einer Qualitätskrise der Universität. Die unter diesen Begriffen zu fassenden Fragestellungen sind – so ist vorweg festzuhalten – nicht in jedem Fall eindeutig zu kategorisieren und werden im Folgenden beispielhaft diskutiert.

2. Unterschiedliche Zugänge zur systematischen Beschreibung der Problembereiche in Hochschulen

2.1 Anpassungskrise der Universität

Unabhängig davon, ob die Universität als Organisation, Institution oder allgemeiner unter dem analytischen Begriff des Systems gefasst wird, ist für sie die Anpassung an sich verändernde Rahmenbedingungen konstitutiv. Von einer Anpassungskrise kann demnach erst dann gesprochen werden, wenn die notwendigen Anpassungsprozesse nicht mehr oder nur teilweise gelingen. Die Gründe hierfür sind oft struktureller Art.

In einem allgemeinen Sinne könnten demnach alle im Folgenden genannten Teilprobleme der Universität auch als Anpassungskrise beschrieben werden. In diesem Abschnitt jedoch werden nur solche exemplarischen Anpassungsprobleme beschrieben, die nicht nur temporär, sondern über einen längeren Zeitraum hinweg auftreten und an spezifische Fragen von Wissenschaft und ihre spezifische Verortung in der Universität gebunden sind.

Für die Universität ergeben sich zunächst typische Anpassungsprobleme aufgrund der fortschreitenden *Differenzierung der wissenschaftlichen*

Gegenstandsbereiche. Die Zunahme von Wissen führt zu stärkeren arbeitsteiligen Erfordernissen und zur Etablierung neuer Disziplinen und Teildisziplinen. Die damit entstehende Vielzahl an Fachgebieten und Fächern stellt die Forschung und Hochschulausbildung insgesamt, aber auch jede einzelne Universität vor erhebliche strukturelle Anpassungsprobleme. Zu nennen ist insbesondere die notwendige Synthese bzw. Reintegration des in einzelnen Disziplinen erarbeiteten Wissens in einen übergreifenden fachlichen Zusammenhang. Dies führt sowohl in der Forschung als auch im Bereich von Studium und Lehre zu umfangreichen Abstimmungs- und Integrationsaufgaben. Die immer wiederkehrende Forderung nach einer interdisziplinär orientierten Forschung oder nach Forschungsk Kooperationen insgesamt sowie die Diskussion unter dem Begriff des ‚Wissensmanagements‘¹ sind hierfür ein sinnbildlicher Beleg.

Neben dieser Differenzierung nach Gegenstandsbereichen ist zunehmend eine *funktionale Differenzierung* zu beobachten, welche die für Universitäten grundlegende Verbindung von Forschung und Lehre fragil erscheinen lässt. Die exponentielle Zunahme an Wissen und Wissensgebieten sowie die marktmäßige Organisation der Forschung haben in den vergangenen Jahren verstärkt dazu geführt, dass durch das Aufkommen außeruniversitärer Forschungsinstitute eine organisatorische Ausdifferenzierung im Bereich der Forschung stattgefunden hat. Die damit auch unter den Vorzeichen von Ressourcenknappheit entstandene Konkurrenzsituation zwischen universitärer und außeruniversitärer Forschung um wissenschaftliche Reputation und die Einweibung von Mitteln unterliegt jedoch in Anbetracht der vielfältigen Aufgaben von Hochschullehrern ungleichen Chancen. Folglich sind sowohl auf institutioneller als auch individueller Ebene dauerhaft strukturelle Nachteile für die Universität und die in ihr arbeitenden Wissenschaftler zu erwarten.

Schließlich stehen die schnelle Zunahme von Wissen und die Differenzierung von Disziplinen und Teildisziplinen mit anderen Aufgaben der Universität insofern in Konflikt, als unterschiedliche Anforderungen an die Struktur des Personals gestellt werden. Während die Organisation von Instituten sowie Verpflichtungen in Studium und Lehre Kontinuität und eine mittelfristige Personalplanung voraussetzen², ist für den raschen Wechsel von Themen und Schwerpunkten im Forschungsbereich die Möglichkeit, auch personell flexibel reagieren zu können, konstitutiv.

¹ Vgl. hierzu u.a. Helmut WILLKE (1998): *Systemisches Wissensmanagement*.

² Zu denken ist hierbei bspw. an die vielfältigen Abstimmungsnotwendigkeiten von Lehrveranstaltungsinhalten und konsistenten Leistungsanforderungen sowie an die vorzuhaltenden Informationen im Hinblick auf die Beratung von Studierenden.

Dies spielt vor allem mit Blick auf den raschen Ausbau der Universitäten in den siebziger Jahren eine bedeutende Rolle für den Entwicklungsstand der Hochschulen. Die in jener Zeit gewachsene Personalstruktur lässt in einigen Bereichen nur in Ansätzen eine Schwerpunktverlagerung in der Forschung zu, was vor allem bei Neuberufungen zum Teil zu erheblichen Anpassungsproblemen führt.

Die hierin zum Ausdruck kommende Gleichzeitigkeit divergenter Ansprüche, die zwischen notwendiger personeller Kontinuität und der – im Hinblick auf die wissenschaftliche Differenzierung – angemessenen Rekrutierung von wissenschaftlichen Spezialisten anzusiedeln sind, lässt sich als charakteristisches Kontinuum des universitären Alltags beschreiben. Hierauf wird im Verlaufe dieses Aufsatzes mit Bezug auf die konkreten Erfahrungen der Evaluation noch näher eingegangen. Ergänzend soll an dieser Stelle festgehalten werden, dass mit diesem latenten Konflikt keinesfalls eine grundsätzliche Inkompatibilität von Forschung und Lehre suggeriert werden soll. Im Gegenteil zeigen nicht zuletzt die Ergebnisse der Evaluation, dass zwar die Organisation von Forschung sowie Studium und Lehre unterschiedliche Anforderungsprofile nahe legt, die für die Universitäten konstitutive inhaltliche Verschränkung von Forschung und Lehre aber gewinnbringend ist.

Das zweite unter dem Begriff der Anpassungskrise zu fassende Moment wurde bereits mit Bezug auf die veränderten Anforderungen an die Personalstruktur angesprochen und lässt sich im Wesentlichen mit einer Deskription der Konsequenzen des Öffnungsbeschlusses erfassen. Die zu Ende der sechziger Jahre diagnostizierte Bildungskrise in den westlichen Industrieländern führte zu einer Öffnung des Schul- und Hochschulsystems, die zum Ziel hatte, die Partizipation am Bildungssystem insgesamt zu verstärken und im Besonderen von der sozialen Herkunft abzulösen. Während sich seit den siebziger Jahren die anteilige schichtabhängige Partizipation am höheren Bildungswesen nicht grundlegend verändert hat, sind die Studierendenzahlen bekanntermaßen deutlich angewachsen und führten zum Ausbau der Hochschulen, damit aber auch zu vielfachen Anpassungsproblemen. In Analogie zur Differenzierung von Wissen war an den Universitäten mit dieser Öffnung eine sukzessive Ausweitung des Lehrangebotes verbunden, die vor allem im Bereich von Studium und Lehre zu einem vermehrten Abstimmungsbedarf bspw.

aufgrund der räumlichen Trennung oder der Notwendigkeit, parallel gleiche Lehrangebote bereitzustellen, geführt hat.³

Darüber hinaus führte das Phänomen der ‚Massenuniversität‘ auf Seiten der Hochschullehrer zu veränderten Anforderungen im Hinblick auf Betreuung und Veranstaltungsgestaltung. Die von Studierenden an vielen Stellen kritisierte Anonymität des Universitätsbetriebs, die Unübersichtlichkeit und fehlende soziale Bindungen sind Ausdruck einer oft begrifflich nur schwer zu fassenden Unzufriedenheit. Nach den im Rahmen der durchgeführten Evaluationen gesammelten Erfahrungen lässt sich diese diffuse Unzufriedenheit mit dem Fehlen eines akademischen bzw. universitären Milieus zusammenfassen, das von persönlichem Austausch und gegenseitiger Verpflichtung geprägt ist. Die Universitäten bemühten sich in den vergangenen Jahren diesen Ambivalenzen, die aus der notwendigen Ausbildung vieler Studierender bei gleichzeitiger Einsicht in die Relevanz persönlicher Betreuung resultieren, Rechnung zu tragen. Gleichwohl – so zeigen die Ergebnisse aus vielen Bereichen der Hochschulforschung – handelt es sich hierbei um ein Problem, das dauerhafte Anpassungsleistungen verlangt und bislang nur partiell zufriedenstellend zu lösen war.

Als dritter Bereich lassen sich unter dem Begriff der Anpassungskrise Bemühungen um eine stärkere internationale Standardisierung des Hochschulstudiums zusammenfassen. Ausgangspunkt ist hierbei die Annahme, dass das deutsche Hochschulsystem im Vergleich zum europäischen Ausland sowie zu nordamerikanischen Universitäten weniger leistungsfähig ist. Auf die damit verbundene Frage der Qualität und Effizienz deutscher Hochschulen wird im Folgenden noch näher eingegangen.

Die zur Zeit diskutierten Konsequenzen sehen eine sukzessive Anpassung an internationale Studienstrukturen vor, um insbesondere die Studienabschlüsse zu standardisieren. Ausdruck hierfür ist die anhaltende Diskussion um die Einführung von Bachelor und Masterstudiengängen. Unabhängig davon, wie diese Maßnahmen im Einzelnen zu bewerten sind und ob sie die gewünschten Effekte einer Verkürzung der Studiendauer und praxisnahen Ausbildung erzielen, zeigt sich insgesamt, dass die Universitäten sich diesen Fragen in Anbetracht der veränderten politischen Rahmenbedingungen nicht entziehen können. Solange es nicht gelingt,

³ Vgl. hierzu ausführlicher Bernhard EINIG, Uwe SCHMIDT (1999): *Zur Situation von Studium und Lehre an der Johannes Gutenberg -Universität Mainz*, 8 ff. sowie Rainer A. MÜLLER (1996): *Geschichte der Universität*, 105 ff.

adäquate Antworten auf die damit gestellten Anforderungen zu finden, ist im Hinblick auf die Anbindung an andere europäische und außereuropäische Bildungssysteme in der Tat von einer Anpassungskrise zu sprechen.

2.2 Organisations- und Steuerungskrise der Universität

Die Diagnose der Organisationskrise impliziert – so lässt sich zusammenfassen –, dass die Universitäten den aufgrund ihres Ausbaus veränderten Strukturen mit den traditionellen Steuerungsformen nicht mehr gerecht werden können. An der bisherigen Form der Steuerung lassen sich als zentrale Kritikpunkte festhalten,

- dass die Steuerung über Gremien wenig Flexibilität zulasse und lange Entscheidungswege verursache,
- Hochschulen in vielen Bereichen zunehmend den Charakter von Unternehmen trügen und entsprechend eine ‚Managementstruktur‘ implementiert werden müsse, womit verbunden wird,
- dass der Leitung von Hochschulen mehr Entscheidungsbefugnisse zugestanden und insgesamt auf eine stärkere Hierarchisierung der Entscheidungswege hingewirkt werden müsse und damit einhergehend
- der ‚föderale Charakter‘ der Hochschulsteuerung an Gewicht verlieren und dementsprechend wieder zu größeren Organisationseinheiten zurückgekehrt werden solle, welche die gegenwärtige Fachbereichsstrukturen abzulösen hätten.

Die an einzelnen Hochschulen erprobten unterschiedlichen Versuche der Neustrukturierung führten bislang zu partiellen Veränderungen, ohne dass sich ein Modell herauskristallisiert hätte, das sich in breiter Form als tragfähig erwiesen hat. Offenkundig aber scheint, dass strukturelle Reformen dauerhaft nur gelingen können, wenn sie die Besonderheiten der Wissenschaftsorganisation widerspiegeln. Entscheidungswege zu verkürzen, darf in diesem Sinne nicht zu einer Einschränkung von wissenschaftlichen Freiräumen und damit auch intrinsischer Motivation führen.

Wenn auch zumindest auf der Organisationsebene der eingeleitete Wandel in den Hochschulen an vielen Stellen mit Augenmaß betrieben und

eine allmähliche Anpassung an die veränderten Strukturanforderungen angestrebt wird, ist zunehmend zu beobachten, dass nicht zuletzt die unterstellte ‚Managementkrise‘ der Universitäten in Zusammenwirken mit der Verknappung der öffentlichen Haushalte im Verlaufe der neunziger Jahre zu einer deutlichen Akzentuierung ökonomischer Steuerungsformen geführt hat. Unabhängig davon, ob die damit einhergehenden ökonomischen Anreizsysteme auf die kollektive (wie bspw. durch die Budgetierung von Fachbereichen und Instituten sowie durch die Einführung von Kosten-Leistungs-Rechnung) oder die individuelle Ebene (wie durch die Dienstrechtsreform) abzielen, ist mit der Ökonomisierung als Steuerungsform die Vorstellung verbunden, dass die internen Steuerungsformen im Wissenschaftssystem nicht ausreichen, um die Hochschulen zu guten Leistungen zu führen. So resümiert Kornhuber bereits Ende der achtziger Jahre, dass „je weiter das System sich entwickelt, desto lauter [...] der Ruf nach Außensteuerung, nach politischer Setzung von Prioritäten“⁴ werde. Mit anderen Worten wird auf kollektiver wie individueller Ebene eine ‚Motivationskrise‘ und damit eine notwendige Verlagerung von internen zu (system-)externen Steuerungsformen unterstellt.

2.3 *Qualitäts- und Effizienzkrise der Universität*

Als weiteres ‚Krisensymptom‘ wird in Kontrastierung zu ausländischen Bildungssystemen vor allem den deutschen Universitäten eine Qualitäts- und Effizienzkrise zugeschrieben. Hierfür werden unterschiedliche Indikatoren angeführt. Zu nennen sind zum einen im Bereich von Studium und Lehre die Fachstudiendauer, das im internationalen Vergleich hohe Alter bei Abschluss der Hochschulausbildung sowie die fehlende Praxisanbindung des Studiums. Zum anderen werden – vielfach in summarischer Form – die Forschungsleistungen an deutschen Hochschulen als mittelmäßig bewertet. Dies drücke sich unter anderem darin aus, dass es nicht gelinge, den wissenschaftlichen Nachwuchs mittel- und langfristig zu binden, da die Universitäten aufgrund ihrer Forschungsstruktur nicht mit Spitzenuniversitäten des Auslandes konkurrieren könnten. Damit aber sei dauerhaft die Wettbewerbsfähigkeit deutscher Hochschulen eingeschränkt, was bspw. unter der Perspektive rückläufiger Studierendenzahlen in den Naturwissenschaften zu einem sukzessiven Qualitätsverlust führen werde.

⁴ Hans-Helmut KORNHUBER (1988): *Mehr Forschungseffizienz durch objektivere Beurteilung von Forschungsleistungen*, 361.

Ohne an dieser Stelle detaillierter auf die hochschulpolitische Diskussion um Qualität und Effizienz der Hochschulausbildung und -forschung eingehen zu können, ist festzuhalten, dass die zu Grunde gelegten Kriterien zur Bewertung von Leistungen und Qualität der Universitäten Konjunkturren unterworfen sind, welche selten auf einem konsistenten Maßstab beruhen und sich häufig widersprechen.⁵ Dies verweist auf einen fehlenden Orientierungsrahmen bzw. auf ein fehlendes Leitbild. Hiermit soll allerdings nicht auf die einsetzende Leitbildentwicklung abgestellt werden, durch die für die jeweiligen Hochschulen ein Bezugssystem im Sinne einer Definition von Zielen und gewünschten Profilen erarbeitet werden soll. Vielmehr geht es – möchte man Kriterien für die Leistungsbewertung nicht nur lokal entwickeln – um das grundlegende Verhältnis zwischen Hochschulbildung, Hochschulforschung und Gesellschaft. Erst die Klärung der Aufgaben und Funktionen von Hochschulen in der Gesellschaft bzw. gegenüber anderen gesellschaftlichen Teilbereichen lässt eine angemessene Beschreibung von Qualitätsstandards und dementsprechend von Defiziten der Universität zu. In diesem Sinne kann die Qualitäts- und Effizienzkrisen der Universitäten zunächst auch als ‚Legitimationskrisen‘ verstanden werden, die aufgrund der fehlenden expliziten Erörterung bzw. des fehlenden Konsenses über die Ziele von Bildung insgesamt und der Universitäten im Besonderen notwendigerweise unspezifisch bleibt, damit aber auch per se Widersprüche provoziert.⁶

3. Universität als System

3.1 Theoretische Vorüberlegungen

Ein mögliches Erklärungsmodell, das die angedeutete Komplexität und die Gleichzeitigkeit der genannten Problemstellungen abzubilden vermag, lässt sich aus systemtheoretischen bzw. strukturfunktionalistischen Ansätzen entwickeln.

Die Betrachtung der Universität als System bietet sich aus mehreren Gründen an. Zunächst ermöglicht ein solcher Zugang die Beschreibung der (relativen) Stellung der Universität in der Gesellschaft. In der Sprache systemtheoretischer Ansätze lassen sich hiermit Systemgrenzen zu ande-

⁵ Vgl. dazu ausführlicher Uwe SCHMIDT, Simon HÄUSER (2001): *Problembereiche und Reformansätze in Studium und Lehre*, 1ff.

⁶ Zur Leitbildentwicklung vgl. in diesem Band den Aufsatz von Manfred HENNEN.

ren gesellschaftlichen Teilsystemen beschreiben und Aufgabenfelder der Universität bzw. des Bildungssystems insgesamt ableiten.

Hierunter fallen im Hinblick auf den Bildungsbereich insgesamt Fragestellungen, die nicht zuletzt auf die Öffnung des Bildungssystems im Verlaufe der sechziger und siebziger Jahre zurückzuführen sind. So wäre bspw. im Zusammenhang mit der Forderung nach der Praxisrelevanz des Studiums zu fragen, ob und wie nahe das Studium an spezifischen Berufsfeldern orientiert sein sollte, ohne gleichzeitig den Anspruch einer breiten wissenschaftlichen Ausbildung aufzugeben. Für den Bereich der Schulen ergeben sich zunehmend Probleme der Definition von Systemgrenzen, wenn mit Blick auf die heterogene Zusammensetzung der Schülerschaft der sozialpädagogische Anspruch den Bildungsaspekt überwiegt. In Universitäten wiederum ergeben sich Probleme aufgrund der parallelen Anforderung, sowohl viele Studierende zu betreuen als auch wissenschaftliche Spitzenleistungen zu produzieren. Auch hier sind Systemgrenzen nicht eindeutig definiert, bleibt offen bzw. wird nicht expliziert, welche primären Aufgaben an das Bildungssystem gestellt werden.

Der Zugang über eine Systemperspektive liegt auch aus einem anderen Grund nahe: In Anbetracht der in den vergangenen Jahren breiten hochschulpolitischen Diskussion kristallisierten sich immer wieder Schwerpunkte heraus, die punktuell Problembereiche in den Blick nahmen, ohne dass die jeweiligen Fragestellungen in den Kontext einer übergreifenden Analyse und Situationsbeschreibung von Hochschulen eingebettet gewesen wären. Zu denken ist in diesem Zusammenhang bspw. an die jeweils unabhängig voneinander geführten Diskussionen um die Fachstudien-dauer, die Kritik am ‚Hochschulmanagement‘ oder an der fehlenden Praxisanbindung des Universitätsstudiums. Der Systemzugang vermag grundsätzlich diese Lücke zu schließen, da er den Fokus auf die Betrachtung von Interdependenzen legt und die Gleichzeitigkeit unterschiedlicher, zum Teil gegensätzlicher Anforderungen und Funktionen abzubilden vermag.

Wie lässt sich nun Universität unter systemischer Perspektive fassen? Die Systemgrenzen der Universität bemessen sich unter einer Perspektive fortschreitender funktionaler Differenzierung und Spezialisierung an den gesellschaftlich relevanten Funktionen, die der Universität zugesprochen werden. Wurden in der Vergangenheit zum Teil noch weite Teile des heutigen Bildungswesens privat organisiert und getragen, so hat das gewachsene gesellschaftliche Interesse an dem öffentlichen Gut Bildung zu einer stetigen Institutionalisierung von Bildungseinrichtungen unter-

schiedlicher Art geführt, die in sich der weiteren funktionalen Differenzierung unterliegen.

Als grundlegende Funktionen, denen die Differenzierung von Systemen folgt, lassen sich nach Parsons im Sinne des sogenannten Vierfunktionen- oder AGIL-Schemas jene der Anpassung oder Ressourcengewinnung (**A**daptation), der Zielverwirklichung (**G**oal-Attainment), der Systemintegration (**I**ntegration) sowie der Erhaltung latenter Strukturen (**L**atent pattern maintenance) nennen.

Die Funktion der Adaptation beschreibt die Notwendigkeit der situationsgerechten Anpassung des Systems an seine Umwelt, was insbesondere die Fähigkeit zur Generalisierung von Anpassungsformen und der Übertragung auf neue, aber vergleichbare Problemstellungen voraussetzt. Anpassung an die Systemumwelt bedeutet vor allem auch die Gewinnung von Umweltressourcen. Insofern erscheint es angemessen, im Folgenden die Anpassungsfunktion im Wesentlichen unter der Perspektive der Gewinnung von Ressourcen zu betrachten, die für die spezifischen Aufgaben des Systems von Nutzen bzw. erforderlich sind.

Die Funktion der Zielverwirklichung ist ebenfalls an der Grenze zwischen System und Umwelt angesiedelt, indem sie spezifische System-Umwelt-Beziehungen herstellt, die zum einen im Hinblick auf die Systemumwelt die ‚Handlungsbereiche‘ des Systems festschreiben, zum anderen eine systeminterne Ausrichtung einnehmen, indem sie Handeln motivieren. Mit anderen Worten ist es wesentlich, dass spezifische Aufgabenfelder notwendig sind, um Handeln zu initiieren. Diffuse Fragestellungen, die an das System herangetragen werden, hingegen sind nicht dazu geeignet, Handlungsprogramme zu entwickeln.

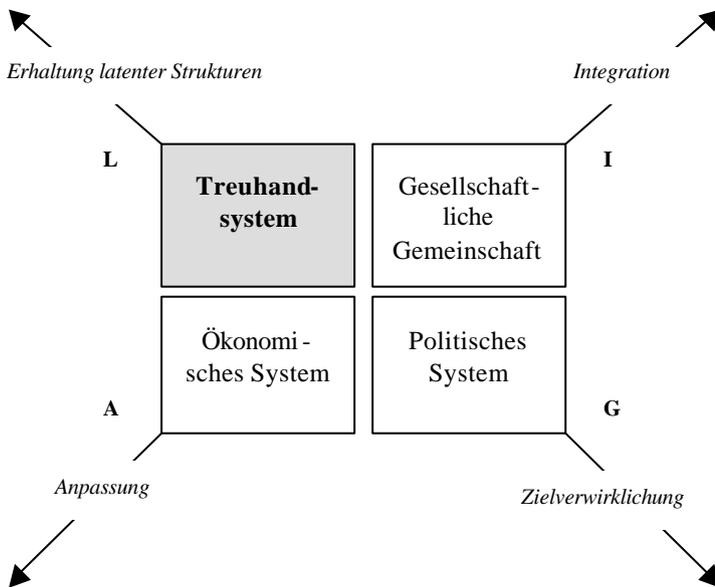
Integration ist funktional primär systemintern ausgerichtet. Dies bedeutet, dass zumal dann, wenn das System sich weiter funktional differenziert, die Funktion der Integration erfüllt sein muss, um eine systemimmanente Bindung zu erreichen und bspw. ein gemeinsames Wissen über Systeme und gelingende Organisationsformen herzustellen.

Die Erhaltung latenter Strukturen schließlich beschreibt übergreifende Systemziele und dient quasi der Identifizierung des Systems. In diesem Sinne kann diese Funktion auch als Programmbildung oder Selbstverständnis des Systems verstanden werden. Die grundlegenden Strukturen des Systems müssen hierbei hinreichend allgemein und damit flexibel

sein, damit das System insgesamt die notwendige Akzeptanz zu seinen zentralen Leitideen erhält.⁷

Betrachtet man unter dieser Perspektive zunächst die gesellschaftliche Differenzierung insgesamt, so lassen sich für die genannten Funktionen unterschiedliche Systemzuschreibungen festlegen:

Abbildung 1: Gesellschaftliche Differenzierung und Systembildung im Sozialsystem



An dieser Stelle können die einzelnen Subsysteme und ihre jeweiligen Funktionen für das soziale System nicht ausführlich erörtert werden. Hinzuweisen ist auf die Zuordnung des Bildungsbereiches zum sogenannten Treuhandsystem. Das Treuhandsystem ist insgesamt zu verstehen als ein Komplex, der gesellschaftliche Werte und Wissen zusammenfasst und

⁷ Vgl. u.a. Talcott PARSONS (1996): *Das System moderner Gesellschaften*, 12ff.; Klaus MÜLLER (1996): *Allgemeine Systemtheorie*, 282ff.; Richard MÜNCH (1982): *Theorie des Handelns*, 81ff.

treuhänderisch für die Gesellschaft verwaltet, d.h. zum einen diesbezügliche Ressourcen ‚aufbewahrt‘, um sie andererseits gleichzeitig generationenübergreifend zu tradieren und weiterzuentwickeln.⁸

Das Spezifikum des Bildungssystems lässt sich im Vergleich bspw. zum allgemeinen Wertesystem oder zum religiösen Komplex durch die Dominanz des kognitiven Aspekts erfassen. Mit anderen Worten basiert es in erster Linie nicht auf traditionellen Werten und Glaubenssätzen, sondern auf der rationalen Durchdringung und Verarbeitung von Fragestellungen und der Herstellung kognitiv fundierten und unter rationalen Erwägungen widerspruchsfreien Wissens. Vielleicht ist dies als zentrale Kategorie der Moderne insgesamt und der Ausdifferenzierung des Bildungssystems im Besonderen zu erachten: Unabhängig davon, ob unter dem Begriff der ‚Wahrheit‘¹¹ oder bspw. unter dem diskurstheoretisch fundierten Begriff des ‚Geltungsanspruchs‘¹² zusammengefasst, bilden der Aspekt der kognitiven Widerspruchsfreiheit sowie der nicht primär kumulativen, sondern systematischen Wissensgenerierung den eigentlichen Kern rationaler Systeme – und damit auch der Universität.

Bevor im Folgenden die vorgestellten Funktionen und Systemebenen auf den Hochschulbereich übertragen werden, soll auf zentrale Gedanken systemtheoretischer Ansätze hingewiesen werden, die das Verhältnis von Systemen und Subsystemen untereinander beschreiben.

Systeme, so wurde zu Beginn ausgeführt, neigen zur Differenzierung. Systemdifferenzierung ist dann zu beobachten, wenn Aufgabenfelder innerhalb des Systems nicht mehr zu erfüllen sind und die Systemintegrität bei Aufgabenzuwachs in Frage steht. Folglich entsteht Differenzierung dort, wo Systemgrenzen aufgrund wachsender Komplexität nicht mehr aufrechtzuerhalten sind und Teilsysteme spezifische Aufgaben und Funktionen übernehmen (funktionale Spezialisierung), um Komplexität zu reduzieren. Mit Bezug auf das System der Wissenschaft folgert Luhmann hieraus:

„Die Wissenschaft übernimmt [...] Teilfunktionen des gesellschaftlichen Zusammenlebens in die Eigenregie eines besonderen Sozialsystems, weil sie sich auf der Ebene der Gesellschaft selbst,

⁸ Talcott PARSONS.; Gerald M. PLATT (1990): *Die amerikanische Universität*, 33, 176ff.

¹¹ Vgl. u.a. Niklas LUHMANN (1991): *Selbststeuerung der Wissenschaft*, 234 ff.

¹² Vgl. Jürgen HABERMAS, Jürgen (1992): *Erläuterungen zur Diskursethik*, 119 ff.

gebunden an eine einheitliche, funktional-diffuse Rollenstruktur, nicht zu höherer Komplexität entfalten lassen.“¹³

Mit der Differenzierung verbunden ist die Konzentration auf Themen, die im Sinne einer ‚thematischen Reinigung‘ weitgehend exklusiv von Subsystemen besetzt werden. Hierfür lassen sich vielfältige Beispiele anführen, wie die thematische Festlegung von Familien auf die Regulierung der Paarbeziehungen und das Eltern-Kind-Verhältnis sowie die Ausgliederung produktiver Funktionen. In den Hochschulen selbst können solche Prozesse der Gegenstandsbeschränkung unter anderem anhand der Trennung von Universitäten und Fachhochschulen oder der Gründung reiner Forschungsinstitute nachgezeichnet werden.

Mit der Differenzierung von Systemen ist weiterhin ein auf den jeweiligen Gegenstandsbereich fokussierter Gewinn an Autonomie verbunden. Der exklusiven Besetzung von Themen entspricht die Abnahme des Einflusses durch andere Teilsysteme. Allerdings bedeutet Autonomie keinesfalls eine gänzliche Herauslösung aus anderen Bezügen, sondern ganz im Gegenteil handelt es sich um eine relative Autonomie, die zwar feldspezifische ‚Unabhängigkeit‘ zulässt, aber gleichzeitig zu einer erhöhten Abhängigkeit in anderen Bereichen führt.¹⁴

Schließlich ist an dieser Stelle noch auf eine grundlegende Prämisse systemischer Ansätze zu verweisen, die das Innenverhältnis des Systems mit Bezug auf die zu erfüllenden Funktionen beschreibt. Die Beziehung der funktional differenzierten Subsysteme untereinander zeichnet sich insbesondere dadurch aus, dass Systeme nur dann zumindest temporär stabil sind, wenn sie der Gleichzeitigkeit funktionaler Anforderungen Rechnung tragen. Systeme sind somit auf Ausgleich und die Herstellung einer Balance zwischen den unterschiedlichen, an spezifischen Funktionen orientierten Systemebenen angewiesen.¹⁵

¹³ Niklas LUHMANN (1991): *Selbststeuerung der Wissenschaft*, 247.

¹⁴ Vgl. zusammenfassend Hartmann TYRELL (1976): *Probleme einer Theorie der gesellschaftlichen Ausdifferenzierung*, 396f. sowie Rudolf STICHWEH, (1988): *Differenzierung des Wissenschaftssystems*, 86ff.

¹⁵ Vgl. hierzu u.a. Michael SCHMID (1998): *Soziales Handeln und strukturelle Selektion*, 248ff.

3.2 Systemebenen der Universität

Wie lassen sich nun diese Vorüberlegungen auf die Betrachtung der Universität und die im Rahmen von Evaluationen gefundenen Ergebnisse übertragen? Zunächst soll hierzu eine Übersetzung der genannten Funktionen auf Bereiche der Universität versucht werden.

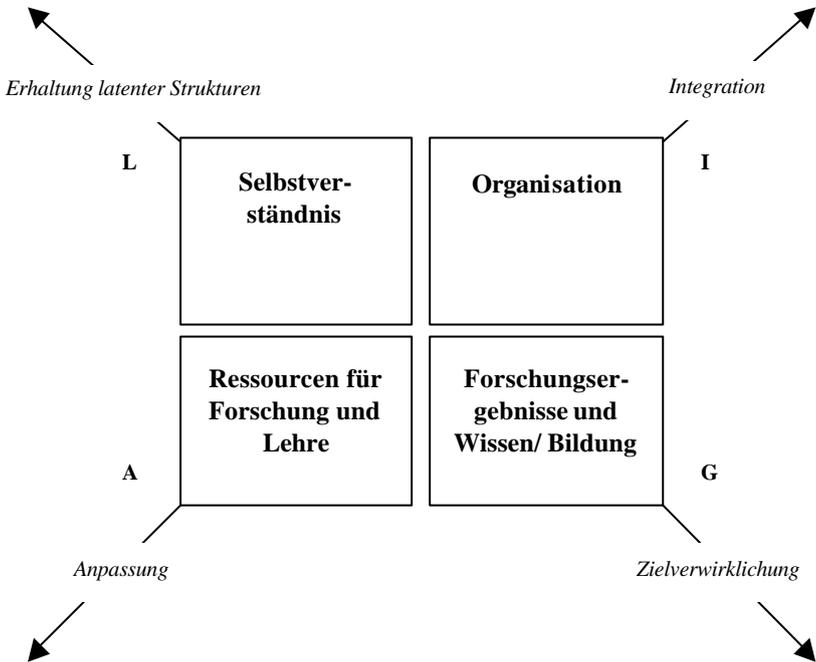
Auf der Ebene der Zielerreichung lassen sich als zentrale Handlungsziele der Universität zum einen die Weitergabe von anspruchsvollen Wissensbeständen an nachwachsende Generationen, zum anderen die Generierung neuen Wissens nennen. Hiermit sind die beiden genuinen Aufgabenbereiche von Universitäten, nämlich Forschung und Lehre, beschrieben.

Als Ressourcen sind mit Blick auf die ‚Anpassungsfunktion‘ in der Universität sowohl Sachmittel und Personal als auch Wissensbestände zu verstehen. Im Sinne einer Bedingungs- und Steuerungshierarchie¹⁶ wäre zu fragen, welche Mittel notwendig sind, um die genannten Ziele in Forschung und Lehre zu erfüllen. Hierunter fallen mithin in erster Linie Personen, die forschen und lehren, Studierende, materielle Güter, aber auch zu erneuerndes und zu adaptierendes Wissen. Die Universität muss sich entsprechend dem ökonomischen System, aber auch anderen Bildungs- und Forschungsbereichen öffnen.

Die Integrationsfunktion lässt sich für die Universität mit der internen Organisation von Forschung, Lehre und Verwaltung, aber auch mit der Gestaltung externer Kooperationen beschreiben.

Die Erhaltung latenter Strukturen schließlich berührt das Selbstverständnis der Universität und thematisiert damit das Wissen über ihre Stellung in Beziehung zu anderen Systemen sowie über systemimmanente Prozesse und Abläufe.

¹⁶ Nach Parsons lassen sich die vier Funktionen – beginnend bei der Anpassungs- bis hin zur Funktion der Erhaltung latenter Strukturen bzw. umgekehrt von der Erhaltung latenter Strukturen zur Anpassung und Öffnung zu anderen Systemen – nach einer Bedingungs- und Steuerungshierarchie unterscheiden.

Abbildung 2: Universität als System

Diesen Systembereichen lassen sich spezifische Indikatoren zuordnen, die in der hochschulpolitischen Dimension in den vergangenen Jahren unterschiedlich stark in den Blickpunkt gerieten und mit den zu Beginn dargestellten ‚Krisenmodellen‘ korrespondieren. So kann das ‚Systemproblem‘ der Zielverwirklichung im Sinne von Forschungsergebnissen und der Transformation von Wissen und Bildung weitgehend gleichgesetzt werden mit der Frage nach einer ‚Qualitäts- und Effizienzkrise‘, findet sich die Managementkrise auf der Ebene des Integrationsproblems wieder und ist die Anpassungskrise nicht zuletzt eine solche der Angemessenheit der zur Verfügung stehenden materiellen und immateriellen Ressourcen.

Ohne an dieser Stelle eine erschöpfende Darstellung denkbarer Indikatoren für die Funktionsbereiche der Universität geben zu können, sollen in

Abbildung 3 zentrale Dimensionen genannt werden, die auch für Evaluationen als Bezugsrahmen dienen können.

Abbildung 3: Universitäre Dimensionen und Indikatoren in den vier Funktionsbereichen

Selbstverständnis	<ul style="list-style-type: none"> • Leitbild (Verständnis über gesellschaftliche Funktionen und übergreifende Ziele) • Programme, Profil (z.B. Schwerpunktsetzung) • Selbstbeschreibung (Evaluation, Lehrberichte, Reflexion) • Wissenschaftliche Standards (Definition guter Forschung und Lehre, Definition von Leistungsanforderungen) • Wissenschaftsethik • Universitäre Identifikation 	<ul style="list-style-type: none"> • Forschungsorganisation (z.B. Drittmittelverwaltung, interne Kooperationen) • Lehrorganisation (Abstimmung von Veranstaltungen und Leistungsstandards) • Verwaltung • Externe Kooperationen (Forschungsk Kooperationen, Partnerschaften, Austausch von Wissenschaftlern, Auslandsaufenthalte von Studierenden) 	Organisation, Kooperation
Ressourcen	<ul style="list-style-type: none"> • Personal (Qualifikation, Betreuungsrelation, Lehrdeputate) • Studierende, wiss. Nachwuchs (Studienmotivation, Leistungsfähigkeit, Ausbildungsgrad) • Sachmittel (Grundzuweisung, Drittmittel, Geräteausstattung, Bibliothek) • Wissensbestände (z.B. Rezeption von Forschungsergebnissen, Lehrbücher) 	<ul style="list-style-type: none"> • Forschungsergebnisse (Wissen, Publikationen, Vorträge, Reputation) • Wissenstransformation, Bildung (Absolventen, Prüfungsleistungen, Praxiseinmündung, Fachstudierendauer) 	Forschungsergebnisse, Wissen

4. Evaluationsergebnisse unter der Perspektive von Systemproblemen der Universität

Entlang dieser Dimensionen können im Sinne einer Operationalisierung der universitären System- bzw. Funktionsbereiche Indikatoren gebildet werden, mit denen im Wesentlichen die auch im Rahmen der durchge-

fürten Evaluationen eruierten Stärken und Schwächen der Fächer und Fachbereiche bzw. der Universität insgesamt analysiert werden können.

In diesem Abschnitt soll daher die Relevanz möglicher Indikatoren und die gegebenenfalls fachspezifische Ausprägung einzelner Fragestellungen auf Grundlage der vorliegenden Evaluationsergebnisse beschrieben werden. Neben den Evaluationen an der Universität Mainz finden hierbei auch veröffentlichte Ergebnisse anderer Evaluationsagenturen Berücksichtigung.

4.1 Handlungsziele in Forschung und Lehre

Die als Ziele der Universität genannten Forschungsleistungen und Aufgaben der Wissenstransformation und Bildung lassen sich an den aus Forschung und Lehre hervorgegangenen Ergebnissen bzw. Produkten messen. Allerdings zeigt sich, dass die Bewertungsgrundlagen für Forschungs- und Lehrleistungen zum Teil deutlich zwischen einzelnen Fächern und Fächergruppen differieren. Zu bestimmen, was gute Forschung und gute Lehre kennzeichnen, bedeutet, Qualitätsparameter zu bilden. Diese hängen wiederum im Sinne eines relationalen Qualitätsbegriffes¹⁷ von der Definition der fachspezifischen Ziele von Forschung und Lehre sowie der besonderen Rahmenbedingungen ab. So zeigt sich, dass bspw. die Anzahl von Publikationen nicht in jedem Fach ein zwingender Beleg für gute Forschung und eine kurze Fachstudiendauer nicht gleichbedeutend mit einem qualitativ hochwertigen Studium oder einer angemessenen Praxiseinmündung ist.

Die Evaluationsergebnisse deuten darauf hin, dass im Bereich der Forschung die quantitativen Wissenschaftsindikatoren in den Geistes- und Sozialwissenschaften eine geringere Rolle spielen. Anerkannte Publikations- und Zitationsindizes sind – anders als in den Naturwissenschaften – nicht oder nur zum Teil vorhanden. Befragt nach zentralen Publikationsorganen, fällt es den Geistes-, Sozial- und Kulturwissenschaften bedeutend schwerer, Zeitschriften zu nennen, die weithin anerkannt sind. Vielfach ist in einzelnen Fächern die wissenschaftliche Differenzierung so weit vorangeschritten, dass jeweils für einzelne Teilgebiete als bedeutend erachtete Zeitschriften oder Verlage genannt werden.

¹⁷ Vgl. zum Qualitätsbegriff in diesem Band Katja AHLSTICH und Tanja URICH-NEITZERT.

In den Naturwissenschaften mögen ebenfalls abweichende Einschätzungen über die Qualität einzelner Publikationsorgane vorhanden sein, doch die gleichsam institutionelle Bedeutung der vorhandenen „Science Citation Indices“¹⁸ enthebt die Anwendung solcher Verfahren zur Messung von Forschungsleistungen der fortgesetzten Legitimierungspflicht. Und in der Tat zeigen die an der Universität Mainz durchgeführten Evaluationen, dass diese Parameter im Hinblick auf die *herausragenden* Leistungen in der Forschung durchaus einen guten Anhaltspunkt bieten. So stimmte die gegenseitige Einschätzung der befragten Wissenschaftler mit den verwendeten Indikatoren, wie bspw. der Veröffentlichungsfrequenz sowie der Häufigkeit der Zitationen, weitgehend überein. Gleichzeitig aber zeigte sich bei der Betrachtung mittlerer Platzierungen, dass dieses Instrument zur Identifikation von Forschungsleistungen kaum noch trennscharf ist.

Die im Rahmen der Evaluationen ermittelten Forschungsleistungen korrespondieren bei den auf Grundlage der genannten Indikatoren herausragenden Wissenschaftlern mit ihrer Reputation in der Lehre. Dieses Ergebnis stimmt mit anderen Befunden der Hochschulforschung überein und charakterisiert ein durchgehend hohes Engagement dieses Personenkreises in Forschung *und* Lehre. Offenbar ist damit die oft zitierte Konkurrenz zwischen Forschung und Lehre nicht zwingend. Vielmehr wirken beide Bereiche – zumindest bei einem Teil der Wissenschaftler – gegenseitig stimulierend.

Gleichzeitig trifft diese Diagnose wiederum nur bei der isolierten Betrachtung universitärer Forschungsleistungen zu. Betrachtet man die Konkurrenzsituation zu außeruniversitären Forschungsinstituten, so deutet vieles auf eine zunehmende Arbeitsteilung hin; denn die hohe Lehrbelastung sowie die vielfältige Einbindung in die Selbstverwaltung in den Universitäten haben ungleiche Voraussetzungen für das Engagement in der Forschung geschaffen, so dass außeruniversitärer Forschungsinstitute zunehmend im Vorteil gegenüber universitären Forschungseinrichtungen sind.

Doch nicht nur diese partielle Verlagerung der Forschungsförderung und damit der entsprechenden Ressourcen setzt die universitäre Forschung zunehmend unter Druck. Auch die Mittelbindung an Forschungsleistun-

¹⁸ Hier ist vor allem die Datenbank SCISEARCH zu nennen. Einschränkend ist festzuhalten, dass die Publikations- und Zitations-Indizes auch in den Naturwissenschaften nicht unumstritten sind (vgl. hierzu u.a. Stefan HORNPOSTEL (1997): *Wissenschaftsindikatoren*, 239ff.), doch genießen sie insgesamt eine deutlich höhere Akzeptanz als in anderen Disziplinen.

gen, die Forderung nach Drittmittelwerbung sowie die damit verbundenen Konjunkturen von Themen und Förderpraktiken lassen den Spielraum der Fächer und Fachbereiche enger werden. Forschungsvorhaben, die langfristig angelegt sind, wie bspw. das Verfassen von Monographien, sind kaum noch opportun, solange sie nicht an ein Drittmittelprojekt angegliedert sind. Die ‚Produktion‘ von Wissenschaft sowie ihre zeitnahe Legitimierung innerhalb und außerhalb der Universitäten führt zu einer kaum noch überschaubaren Flut an Veröffentlichungen. Doch obwohl – unabhängig von den einzelnen Fachrichtungen – diese Praxis nicht gut geheißt wird, tragen unter anderem Kennzahlen gestützte Mittelverteilungsmodelle, vor allem aber auch die mit langen Publikationslisten verbundenen persönlichen Berufschancen auf dem Wissenschaftsmarkt offenbar dazu bei, dass dessen ungeachtet die Zahl der Veröffentlichungen weiterhin sukzessive zunimmt.

Ein damit verbundenes Problem, das zumindest für Teilbereiche der naturwissenschaftlichen Fächer von Bedeutung ist, liegt in dem Abwägen, inwieweit verstärkt auf Auftragsforschung auf Kosten der Grundlagenforschung gezielt werden soll. Im Hinblick auf die Drittmittelwerbung, so die Erfahrung in den Naturwissenschaften, erscheinen zum einen Auftragsprojekte lukrativer und sind zum anderen die Kriterien bspw. für die von der DFG geförderte Grundlagenforschung deutlich restriktiver. Es läge daher nahe, sich verstärkt im Bereich der oft an spezifischen Problemen orientierten Auftragsforschung zu engagieren. Doch auch dies birgt Gefahren in sich. So ist vor allem die an aktuelle und damit auch an vorübergehende Probleme gebundene Auftragsforschung von nicht zu kontrollierenden Konjunkturen geprägt, die bspw. im Bereich der Waldschadensforschung im Verlaufe der neunziger Jahre zunächst zu einer hohen Drittmittelrekrutierung, jedoch mit dem schwindenden gesellschaftlichen Interesse zu einer relativ unvermittelten Einstellung der Forschungsförderung führte.

Für die Wissenschaft bedeutet dieser in der Zukunft – so ist zu erwarten – noch schnellere Themenwandel kaum lösbare Anpassungsprobleme. So sind an Forschungsfelder Geräteinvestitionen und aufgrund der wissenschaftlichen Differenzierung der Fachgebiete zunehmend hoch spezialisierte wissenschaftliche und nichtwissenschaftliche Fachkräfte zum Teil längerfristig gebunden. Die ‚Bewegungsfähigkeit‘ der universitären Forschung ist damit stark eingeschränkt. Die unter dem Stichwort der Profilbildung diskutierten Prozesse innerhalb der Hochschulen bedürfen unter diesen Bedingungen oft jahrelanger Vorbereitung, bevor sie greifen. Sie

bergen selbst wiederum die Gefahr in sich, Festlegungen einzuleiten, die vergleichsweise schnell zu neuen Anpassungsproblemen führen.

Zusammenfassend lassen sich diese, auf der Ebene der Forschungsziele generierten Entwicklungen als zunehmende Ambivalenz kennzeichnen, die den Universitäten einerseits eine Bündelung ihrer Kapazitäten auf einzelne Forschungsfragen nahe legt, gleichzeitig aber ein hohes Maß an Elastizität erfordert, um den sich wandelnden gesellschaftlichen Erwartungen an Wissenschaft entsprechen zu können.

Im Hinblick auf die Ziele von Studium und Lehre sind ähnliche Ambivalenzen zu beobachten. Wenn die wesentliche Aufgabe der Universität in der Weitergabe anspruchsvoller Wissensgegenstände an die nachfolgende Generation besteht¹⁹, liegt es nahe, den wissenschaftlichen Anspruch der universitären Ausbildung hervorzuheben. Andererseits haben sich die Rahmenbedingungen, an denen sich universitäre Lehre zu orientieren hat, in den vergangenen dreißig Jahren einschneidend verändert. Die notwendige Öffnung der Hochschulen hat nicht nur zu dem Phänomen der Massenuniversität geführt, sondern daran anschließend die zuvor selbstverständliche berufliche Anschlussfähigkeit nach einem Studium relativiert. Damit aber stellen sich der Universität zum Teil gänzlich neue Aufgaben. Die ausschließliche Orientierung an der Reproduktion des wissenschaftlichen Systems durch die akademische Ausbildung des wissenschaftlichen Nachwuchses ist der Forderung nach Praxisorientierung und Praxisanbindung gewichen.

Sich dieser Frage anzunehmen, so zeigen die Ergebnisse der Evaluation, fällt den Fächern mehr oder weniger leicht. Ohne dies uneingeschränkt verallgemeinern zu können, lässt sich beobachten, dass die naturwissenschaftlichen Fächer seltener ‚Berührungspunkte‘ mit unterschiedlichen Praxisfeldern haben. In den Geistes-, Sozial- und Kulturwissenschaften hingegen wird die Frage nach einer praxisnahen Ausbildung tendenziell eher zurückhaltend beantwortet und auf das Ziel der Vermittlung allgemeiner, übergreifender Fähigkeiten verwiesen.

Worauf ist dies zurückzuführen? Eine naheliegende Erklärung ergibt sich aus den unterschiedlichen Studienabschlüssen. Der in den Geistes- und Sozial- sowie den Sprach- und Kulturwissenschaften weit verbreitete Magisterstudiengang zielt seiner Struktur nach im Vergleich zu Diplomstudiengängen stärker auf ‚Querschnittswissen‘ ab. Darüber hinaus gibt es in

¹⁹ Vgl. hierzu Manfred HENNEN (1998): *Leitbilder zur Evaluation der Lehre*.

diesen Bereichen nur wenige ‚zwingende‘ Berufsfelder nach dem Studium, die zudem einem starken Wandel unterliegen. Ein auf einen oder wenige Berufe ausgerichtetes Studium würde zwar einerseits für diese Bereiche die berufsspezifischen Fertigkeiten vermitteln und damit den Übergang vom Studium in den Beruf erleichtern. Andererseits aber bedeutete eine solche berufsorientierte Akzentuierung der universitären Ausbildung zugleich eine Minderung der Optionen nach dem Studium. Dies korrespondiert mit der Einschätzung zu den Zielen von Studium und Lehre, die im Rahmen der Evaluation von den genannten Wissenschaftsbereichen vertreten wurden. Danach wird die Vermittlung von Fertigkeiten angestrebt, die auf das Erkennen übergreifender Zusammenhänge, exemplarisches Arbeiten und Problemlösungskompetenz abzielen.

In den Naturwissenschaften ist selbstverständlich auch zu beobachten, dass fachferne Berufe ergriffen werden. So ist das ‚Abwandern‘ von Naturwissenschaftlern in informationstechnologische Berufe kein untypisches Phänomen. Doch unterscheiden sich diese Fächer von den Geistes- und Sozial- sowie Sprach- und Kulturwissenschaften dadurch, dass überhaupt dominierende Berufsbilder vorhanden sind und nur temporär ‚fachfremde‘ Tätigkeiten aufgenommen werden. Die durch die Anlage von Diplomstudiengängen stärkere Fixierung auf ein Fach bildet somit ein Äquivalent für die grundsätzlich spezifischere Orientierung an Berufsfeldern.

4.2 Die Dimension der Ressourcen

Die Funktion der Anpassung wird unter systemtheoretischer Sicht als Öffnung zu anderen Systemen oder Subsystemen verstanden. Systeme, so wurde zuvor beschrieben, werden instabil, wenn ihre Grenzen nicht mehr eindeutig sind. Gleichzeitig aber sind Systeme – so auch die Universität – nicht autark und insofern auf Ressourcen angewiesen, die sie von anderen Systemen, wie bspw. dem ökonomischen, beziehen. Mit anderen Worten sind sie auf ein ausbalanciertes Verhältnis von Systemöffnung und Systemschließung angewiesen.

Für die Universität lassen sich sehr unterschiedliche Formen von Ressourcen, die für Forschung und Lehre notwendig sind, identifizieren. Zunächst handelt es sich hierbei um finanzielle Ressourcen, die in Personalstellen und Sachmitteln ihren Ausdruck finden. Zum anderen sind aber auch bereits vorhandene Wissensbestände, die Möglichkeiten der Rekrui-

tierung des wissenschaftlichen Nachwuchses und nicht zuletzt die Studierenden und ihr jeweiliges Leistungsvermögen als Ressourcen der Wissenschaft und mithin der Universität hervorzuheben.

In den vergangenen Jahren wurden den Hochschulen insgesamt zum Teil massive finanzielle Kürzungen auferlegt. Personalabbau und die Streichung von Sachmitteln bis hin zur Schließung ganzer Institute waren nicht nur im Rahmen der deutschen Wiedervereinigung zu Beginn der neunziger Jahre die Regel. Die Erfahrungen der Evaluation zeigen, dass die Fächer Fragen der Personal- und Sachmittelausstattung keineswegs als Begründung für jedwede Art von Problemen oder Defiziten in Forschung und Lehre anführen. Andererseits zeigt sich vor allem in den sogenannten Massenfächern, wie der Pädagogik, der Politikwissenschaft, der Germanistik und den Wirtschaftswissenschaften, dass die Personalsituation denkbar schlecht ist und die Betreuung von Studierenden zum Teil nur noch selektiv wahrgenommen werden kann. Hierzu trägt bei, dass die Einwerbung von Drittmitteln in diesen Fächern im Vergleich zu naturwissenschaftlichen Fächern gering ist. Während die Naturwissenschaften bspw. in Rheinland-Pfalz die durch das sogenannte Personalbemessungskonzept geregelte und aufgrund der rückläufigen Studierendenzahlen geringere Grundzuweisung zum Teil durch die erhöhte Drittmiteleinwerbung und eine höhere Anzahl an Promovenden kompensieren können²⁰, rücken in den Sozial- und Geistes- sowie den Sprach- und Kulturwissenschaften einige Fächer nahe an die Ausstattungsgrenzen von Numerus-Clausus-Fächern. Davon abgesehen, dass sich diese Fächergruppen per se traditionell bei Einwerbung von Drittmitteln etwas schwerer tun, führt die an der Studierendenzahl bemessene knappe Grundausrüstung auch dazu, dass kaum personelle Ressourcen freigesetzt werden können, um Drittmittelanträge in größerem Umfang zu stellen.

Die Verknappung der den Fächern zur Verfügung stehenden Sachmittel geht mit Blick auf die Lehrmittel einher mit einer in den vergangenen Jahren raschen technischen Entwicklung. Zu nennen ist hier zunächst die Computerausstattung, die zwar im Verlaufe der neunziger Jahre an den Hochschulen deutlich verbessert wurde, aber nach wie vor eine dauerhafte Anpassung an neue technische Standards erfordert. Dies gilt vor allem für die sogenannten CIP-Pools, die in einigen Fächern – gemessen an der

²⁰ Nach dem Personalbemessungskonzept bemisst sich die Höhe der Grundzuweisung durch das Land zum einen nach der Anzahl der Studierenden in der Regelstudienzeit und ist hierbei an die Berechnungsgrundlagen der Kapazitätsverordnung angelehnt; zum anderen werden besondere Leistungen, wie die Höhe der Drittmiteleinwerbung, die Anzahl der Promovenden und Habilitanden, die Einrichtung von Sonderforschungsbereichen und Maßnahmen zur Förderung von Nachwuchswissenschaftlerinnen in Rechnung gestellt.

Zahl der Studierenden – schlecht ausgestattet sind.²¹ Hinzu kommt, dass die Bereitstellung von entsprechender Hard- und Software alleine nicht ausreicht, sondern auch entsprechende personelle Kapazitäten durch die Betreuung der Computer gebunden sind. Dies wirkt sich vor allem in kleineren Fächern oder Fachbereichen aus, in denen keine personellen Kapazitäten bereitgestellt werden können, durch welche die entsprechende Betreuung von CIP-Pools und anderer PC-Ausstattung wahrzunehmen wäre. Doch nicht nur die Ausstattung mit Computern, sondern auch jene mit neuen visuellen Übertragungsmedien, wie bspw. Beamern, bedarf der sukzessiven Anpassung, die von den Fächern nur partiell aus ihrem regulären Etat wahrgenommen werden kann.²²

Darüber hinaus ist für die Naturwissenschaften im Hinblick auf die Lehrmittel die notwendige und fortlaufende Reinvestition für die Geräteausstattung in den Praktika zu vermerken. Die Evaluationen zeigen, dass mit dem Ausbau der Universitäten in den siebziger und achtziger Jahren in vielen Fächern zwar zunächst eine gute Geräteausstattung gegeben war, die notwendigen Anpassungen jedoch nicht fortlaufend geleistet wurden und somit partiell in den vergangenen Jahren ein wachsender Bedarf an einer Modernisierung der Praktika entstanden ist. Dies gilt sowohl für den Bereich der Lehre als auch jenen der Forschung. Die während der Evaluationen in Einzelgesprächen zur Forschungssituation gewonnenen Ergebnisse weisen darauf hin, dass zum Teil die durch Berufungsverhandlungen vereinbarte Ausstattung zwar für die ersten Jahre eine ausreichende Geräteausstattung gewährleistet, die Erneuerung der Geräte jedoch Mühe bereitet.

Schließlich ist als in den Evaluationen hervorgehobenes Ressourcenproblem die auch in der öffentlichen Diskussion temporär diskutierte Bibliotheksausstattung zu nennen. Auch hier haben die finanziellen Engpässe der Hochschulen dazu geführt, dass Buchbestände nicht angemessen aktualisiert werden können und vor allem Fachzeitschriften zum Teil bereits abbestellt werden mussten. Die Hochschulen reagieren hierauf mit der Konzentration von Ressourcen, indem Fachbibliotheken zusammengeführt oder in größere Bibliotheken integriert werden. Diese Maßnahmen werden aufgrund der damit erwarteten Einspareffekte sowie der erhofften besseren personellen Rahmenbedingungen und längeren Öff-

²¹ So stehen bspw. in der Germanistik ca. 1.300 Hauptfachstudierenden 10 Computer-Arbeitsplätze zur Verfügung.

²² Die Universität Mainz reagierte hierauf mit einem umfangreichen Programm, das aus zentralen Mitteln die Ausstattung von Hörsälen und Seminarräumen mit modernen optischen und akustischen Übertragungsgeräten erlaubt.

nungszeiten von den Fächern sowie den Studierenden in vielen Fällen mitgetragen. Ob damit dauerhaft die gegenwärtige Situation, an der die Studierenden vor allem die begrenzten Öffnungszeiten der Bibliotheken und das zum Teil fehlende Fachpersonal kritisieren, verbessert werden kann, bleibt mit Blick auf die wachsenden Aufgabenfelder des Bibliothekspersonals dahingestellt.²³

Die bislang zusammengefassten Probleme der Sach- und der personellen Ausstattung beschreiben allerdings nicht ausreichend die Dimension der Ressourcen. Vielmehr sind hierunter auch qualitative Aspekte zu nennen, die das Ausbildungsniveau sowie die Rekrutierung von Studierenden und Wissenschaftlern erfassen.

Eine in allen an der Universität Mainz durchgeführten Evaluationen zu Tage tretende Einschätzung ist, dass die Vorbildung und damit das Leistungsniveau eines nicht unbeträchtlichen Teils der Studierenden zu gering ist. Als ein Indiz für diese Diagnose mögen die hohen ‚Dropout‘-Quoten bereits während des Grundstudiums dienen, die in allen Fachrichtungen zu finden sind.²⁴ Bezeichnender allerdings ist meines Erachtens, dass in den Evaluationsgesprächen nicht nur von den Lehrenden, sondern mit gleicher Entschiedenheit von vielen Studierenden die Auffassung vertreten wird, dass die Heterogenität des Leistungsniveaus für ihr Studium nicht zuträglich ist und eine frühe Selektion auch im Interesse schwächerer Studierender wünschenswert sei. Die Abstimmung zwischen den Erwartungen der Hochschulen an die Studienanfänger und den Ausbildungsprämissen in den Schulen – so zeigen die Erfahrungen der Evaluation – wird in Anbetracht der fortschreitenden Ausdifferenzierung wissenschaftlicher Disziplinen zunehmend zu einer zentralen Aufgabe, die bislang nicht systematisch erfasst und gelöst wurde.

Als weitere und sicherlich bedeutendste Ressource für Forschung und Lehre ist die Rekrutierung des wissenschaftlichen Personals zu nennen.

²³ Vor allem die Umstellung der Bibliotheken auf die Möglichkeit der Online-Recherche und die entsprechende Erfassung der Buchbestände bindet nach wie vor erhebliche personelle Ressourcen.

²⁴ Auf der Grundlage von Kohortenanalyse zeigt sich, dass nicht nur – wie vielfach angenommen – in den Naturwissenschaften dieser Selektionsprozess bereits zu einem frühen Zeitpunkt einsetzt, sondern in annähernd gleichem Maße auch in den geistes- und sozial- sowie den sprach- und kulturwissenschaftlichen Fächern diese Entwicklung zu beobachten ist. Die ‚Dropout‘-Quoten, die nicht zwischen Hochschul- und Fachwechslern sowie Studienabbrechern unterscheiden, liegen in vielen Fächern bei der Hälfte der Studienanfänger.

Hierunter lassen sich sowohl Modi der Berufung als auch die Förderung des wissenschaftlichen Nachwuchses summieren.

Berufungen sind ein wesentliches Instrument zur Anpassung bei der Verlagerung von Forschungsschwerpunkten, aber auch von Studieninhalten. Die jeweilige Berufungspolitik der Fächer und Fachbereiche ist in zweierlei Form Gegenstand von Evaluationen. Zum einen wird sie indirekt durch die Reorganisationsbemühungen der Fächer erfasst, indem sie vielfach als Instrument zur Profilbildung genutzt wird. Schwerpunktsetzungen im Bereich der Forschung und die im Hinblick auf die knapper werdenden Mittel zum Teil unumgängliche Bündelung von Ressourcen sind oft nur dann möglich, wenn Professorenstellen durch Emeritierung freigesetzt werden. Zum anderen aber trugen bereits erfolgte Berufungen in der überwiegenden Zahl der Fächer wesentlich dazu bei, Reorganisationen einzuleiten. Die an der Universität Mainz in der überwiegenden Zahl der Fälle auf Initiative der Fächer durchgeführten Evaluationen kamen nicht zuletzt auf Initiative von neu berufenen Professorinnen und Professoren zustande. Aufgrund ihrer Außenperspektive, die sie zu Beginn ihrer Tätigkeit noch einnehmen, sind sie daran interessiert, Neuerungen einzuführen und eigene Vorstellungen einzubringen.

Als konstantes Problemfeld konnte im Verlaufe der Evaluationen die Situation des wissenschaftlichen Nachwuchses identifiziert werden. Hierbei ist wiederum zu unterscheiden zwischen der Situation in den Naturwissenschaften und in den Geistes-, Sozial-, Sprach- und Kulturwissenschaften. In den Naturwissenschaften ist aufgrund der stärkeren Forschungsaktivitäten die Situation des wissenschaftlichen Nachwuchses durch größere Diskontinuitäten hinsichtlich der beruflichen Anbindung geprägt. Die Beschäftigung von Doktoranden bspw. erfolgt zu einem großen Teil über Drittmittelprojekte. Dies wiederum führt zu ungleichen Voraussetzungen im Hinblick auf die Ausbildungschancen. Zunächst zeigt sich, dass die Finanzierung von Stellen über Drittmittelprojekte Unsicherheiten mit sich bringt, die bei ‚regulären‘ Promotionsstellen aus Landesmitteln geringer sind. So kommt es nicht selten vor, dass während der Promotion die ‚Finanzierungsmodelle‘ gewechselt werden müssen, indem ein- oder zweijährige Mitarbeiterverträge wechseln mit Zeiten, in denen Promovenden keine Stellen bzw. solche als wissenschaftliche Hilfskräfte innehaben. Dies mag insofern – so deutet sich bereits jetzt an – in den kommenden Jahren eine geringere Rolle spielen, als in vielen naturwissenschaftlichen Bereichen Nachwuchskräfte fehlen und sich damit die Berufschancen deutlich verbessern.

Darüber hinaus entsteht ein Ungleichgewicht derart, dass Promovenden, die über Landesstellen beschäftigt sind, stärker in die Lehre eingebunden werden, was einer größeren Belastung und geringeren Zeitressourcen für eigene Forschungsaktivitäten entspricht. Gleichzeitig partizipieren Promovenden in Forschungsprojekten grundsätzlich weniger an Lehraufgaben, was zumindest dann, wenn sie eine Hochschulkarriere anstreben, zu Defiziten in der Lehrqualifikation führt. Die Evaluationen zeigen, dass insbesondere naturwissenschaftliche Fächer einen erhöhten Bedarf an hochschuldidaktischer Qualifizierung haben.

In den Geistes- und Sozial- sowie Sprach- und Kulturwissenschaften ist ebenfalls ein Nachholbedarf an didaktischer Qualifikation zu beobachten. Die Einführung in die Lehraufgaben erfolgt – und hierbei sind kaum Unterschiede zwischen den Disziplinen zu erkennen – eher zufällig. Lehre ist anders als Forschung kaum geleitet und der überwiegende Teil der Hochschullehrer hat im Hinblick auf die Aufgaben in Studium und Lehre einen Prozess autodidaktischer Aneignung durchlaufen. Die meisten der vorliegenden Zielvereinbarungen weisen in diesem Zusammenhang nicht nur auf das Fehlen didaktischer Grundfertigkeiten hin, sondern sie machen deutlich, dass auch eine Einführung in Verwaltungs- und Selbstverwaltungskenntnisse, in Aufgaben der Personalführung, der Forschungs- und Lehrorganisation dringend notwendig wird.

Für die Geistes- und Sozial- sowie Sprach- und Kulturwissenschaften ist darüber hinaus zu beobachten, dass aufgrund der geringen Einbindung in Forschungsprojekte und damit der fehlenden Kooperationsmöglichkeiten während der Abfassung von Qualifikationsarbeiten der Wunsch nach Austausch latent ist. Es lässt sich beobachten, dass vor allem solche Nachwuchswissenschaftler, die nicht über Mitarbeiterstellen an der Universität eingebunden sind, häufig den Wunsch äußern, bspw. über regelmäßige Kolloquien eine stärkere Integration in ihr Fach zu erfahren. Dies geht mit der Beobachtung konform, dass in diesen Disziplinen eine Fachkultur, die auf Forschungsk Kooperationen aufbaut, deutlich schwächer ausgeprägt ist. Hierzu trägt bei, dass viele Doktoranden beklagen, dass ihre Möglichkeiten zur Partizipation an nationalen und internationalen Kongressen aufgrund fehlender finanzieller Unterstützung stark eingeschränkt ist. Zu empfehlen wären hier zentrale Programme, die bereits auf der Ebene des wissenschaftlichen Nachwuchses Formen der internen und externen Forschungsk Kooperation fördern, um langfristig eine Fachkultur zu etablieren, in der die Zusammenarbeit in gemeinsamen und nicht regional beschränkten Projekten selbstverständlich ist. Nicht zuletzt

hieran – so meine These – fehlt es in den genannten Disziplinen, um auf Dauer ihre Forschungseffizienz zu erhöhen.

4.3 Die Dimension der Integration

Diese Überlegungen leiten über zu den Ergebnissen der Evaluation, die der Dimension der Integration zuzuordnen sind. Die Integration von Wissensbeständen und – im Hinblick auf die organisatorischen Anbindung – von Forschung, Studium und Lehre ist nach Luhmann als das zentrale Problem unter Bedingungen der Systemdifferenzierung zu sehen:

„Im Falle jeder Differenzierung wird der Zusammenhang zum Problem. Die erreichbare Trennschärfe wird mit Einheitsverlust bezahlt, und es bedarf zwar nicht einer dialektischen Synthese (denn es handelt sich nicht um ‚Widersprüche‘), aber den Beziehungen zwischen den auseinandergezogenen Formen muß besondere Aufmerksamkeit gewidmet werden.“²⁵

Die hierunter zu summierenden Fragen zielen sowohl auf die Forschungs- und Lehrorganisation ab als auch bspw. auf die Integration von Wissen innerhalb und zwischen den Disziplinen sowie auf die Verständigung der Wissenschaftler untereinander und mit Studierenden. Die Dimension der Integration kann unter dieser Perspektive auch als Implementierung und Konkretisierung von spezifischen Fach- und Hochschulkulturen verstanden werden.

Im Bereich der Forschung, so zeigen die Ergebnisse der Evaluation, gilt zumindest für die Naturwissenschaften, dass sich eine Forschungskultur etabliert hat, die nur in seltenen Fällen institutionalisierter Hilfestellungen für Forschungsk Kooperationen bedarf. Mit anderen Worten sind gemeinsame Forschungsvorhaben grundsätzlich nicht zu ‚verordnen‘, sondern sie basieren häufig auf individuellen Initiativen und sind im Bereich der Naturwissenschaften sowohl innerhalb als auch außerhalb der eigenen Universität selbstverständlicher als in anderen Disziplinen. Ungeachtet dessen können die Fächer und die Universität insgesamt Rahmenbedingungen für ‚gelingende Forschung‘ und entsprechende Kooperationen schaffen, indem die fachlichen Schwerpunkte aufeinander abgestimmt, die notwendige Grundausrüstung zur Verfügung gestellt und bspw. die An-

²⁵ Niklas LUHMANN (1994): *Perspektiven für Hochschulpolitik*, 219.

siedlung von Forschungseinrichtungen in Universitätsnähe – wie Max-Planck-Institute – unterstützt werden.

In den Geistes-, Sozial-, Sprach- und Kulturwissenschaften gestaltet es sich – wie bereits angedeutet – nach unseren Eindrücken wesentlich schwieriger, gemeinsame, gegebenenfalls interdisziplinär ausgerichtete Forschungsprojekte zu initiieren. Es herrscht eine individualisierte Forschungskultur vor, die nicht nur Kooperationen, sondern auch die Integration von Wissensbeständen behindert. Diese Disziplinen zeichnen sich in ihrer Forschungsstruktur häufig durch gegenseitige Abgrenzung aus, die kaum zur Kumulation von Wissen beitragen. Dies gilt in einigen Fächern nicht zuletzt im Hinblick auf eine im Vergleich zu naturwissenschaftlichen Forschungsgebieten schwächere Ausrichtung an internationalen Forschungsvorhaben.

Integrationsprozesse spielen jedoch nicht nur auf der Forschungsebene, sondern in gleichem Maße bei der Koordination von Lehrveranstaltungen und der Betreuung von Studierenden eine wichtige Rolle. Die Abstimmung von Kerninhalten der Lehre ist als ein fächerunabhängiger durchgehender Problembereich zu identifizieren und entsprechend Gegenstand nahezu aller Zielvereinbarungen zwischen Fächern und Hochschulleitungen.²⁶ In vielen Fällen ist zu beobachten, dass die Entwicklung von Curricula und der damit intendierte aufeinander aufbauende und verzahnte Studiengang im Verlaufe der Zeit bspw. aufgrund hoher personeller Fluktuation an Präsenz für die einzelnen Wissenschaftler verliert und einen primär formalen Status erhält, dessen Einlösung nur partiell gelingt. Eine vor allem im Hinblick auf die Gestaltung des Grundstudiums falsch verstandene ‚Freiheit der Lehre‘ führt häufig dazu, dass die Abstimmung mit anderen Veranstaltungen sowie die Verständigung zwischen den Dozenten über zentrale Lehrinhalte ausbleibt. Dies begründet zum Teil unnötige Redundanzen einzelner Stoffgebiete, die in unterschiedlichen Veranstaltungen und Veranstaltungstypen mehrmals angeboten werden, während gleichzeitig die notwendigen Grundlagen für das Hauptstudium nicht oder unzureichend gelegt werden.

²⁶ Vgl. u.a. FACHBEREICH CHEMIE, Hrsg. (2000): *Evaluation im Fachbereich Chemie an der Johannes Gutenberg-Universität* sowie die Evaluationsberichte des VERBUND NORDDEUTSCHER HOCHSCHULEN, Hg (1999): *Evaluation von Studium und Lehre im Fach Mathematik 1997/98* sowie (1999): *Evaluation von Studium und Lehre im Fach Geowissenschaften 1997/98*.

Als spezifisches Problem erweist sich in diesem Zusammenhang die Abstimmung mit Veranstaltungen, die als Zubringerleistungen aus anderen Fachbereichen oder Fächern in das Studium integriert sind. Dies trifft vor allem für Diplomstudiengänge, aber bspw. auch für die pädagogische Ausbildung von Lehramtsstudierenden zu. In den naturwissenschaftlichen Fächern werden in vielen Fällen durch die Serviceleistungen – wie bspw. Mathematik für Biologen oder Chemie für Mediziner – die Grundlagen für das weitere Studium gelegt. Eine entsprechende Abstimmung mit den darauf aufbauenden Veranstaltungen und eine Entwicklung eines Lehrkanons, der den Anforderungen des Studiengangs entspricht, fällt oft deshalb schwer, weil hier unterschiedliche Perspektiven zum Tragen kommen. Während bspw. der Biologie daran gelegen ist, für die Ausbildung ihrer Studierenden spezifische, auf biologische Stoffbereiche und Studienanforderungen ausgerichtete Inhalte von den sogenannten Servicefächern zu erhalten, ist aus der Sicht der Mathematik die Vermittlung von mathematischen Grundlagen unerlässlich, auch wenn diese in keinem direkten Bezug zu den nötigen Kenntnissen für das Biologiestudium stehen. Da zudem aufgrund der geringen Attraktivität, die das Erbringen von Serviceleistungen in den Fächern genießt, die entsprechenden Veranstaltungen zum Teil von unterschiedlichen Hochschullehrern angeboten werden, bedarf es der kontinuierlichen Abgleichung der Lehrinhalte, für die – so zeigen die Erfahrungen der Evaluation – ein Hochschullehrer im Fach verantwortlich zeichnen sollte.

Doch nicht nur die Abstimmung von Kerninhalten der Lehre, sondern auch die Definition und Angleichung von Leistungsanforderungen ist regelmäßig Gegenstand der Evaluationen. Auf die häufig von Studierenden im Rahmen der Gruppengespräche vertretene Auffassung, dass höhere Leistungsanforderungen, die durchaus zu Selektionsprozessen führen sollten, gewünscht werden, wurde bereits hingewiesen. In der Tat zeigen die Ergebnisse der Evaluationen, dass sich in einigen Fächern eine allmähliche Nivellierung der Notenskala entwickelt hat, die entweder nach oben oder nach unten keine oder kaum noch Spielräume bietet.²⁷ Mit einer solchen Praxis entziehen sich die Fächer mit Blick auf die Berufschancen ihrer Absolventen einer Verantwortung, indem sie Leistungsdivergenzierung kaum noch transparent machen.

²⁷ Verwiesen sei hierbei z.B. auf die Rechtswissenschaften, die traditionell gute und sehr gute Notenwerte fast gänzlich aus ihrer Bewertungsskala entfernt haben – bzw. umgekehrt auf die Biologie, in der eine Abschlussnote ‚gut‘ bereits den unteren Skalenwert abbildet.

Darüber hinaus trägt eine Festschreibung von Leistungsstandards, die unabhängig von den jeweiligen Dozenten in allen Veranstaltungen ihren Niederschlag finden, zur Sicherheit unter den Studierenden bei. Eine Einschätzung über den eigenen Leistungsstand, so die Studierenden, gelingt zum Teil erst im Rahmen von Zwischen- und Abschlussprüfungen und führt unter anderem dazu, dass sich aufgrund der fehlenden Gewissheit über den Wissensstand der Studienabschluss verzögert. In diesem Zusammenhang ist ein nahezu durchgehendes Ergebnis der durchgeführten Evaluationen, dass sich die Studierenden eine detaillierte und qualitative Leistungsrückmeldung wünschen, die dazu beiträgt, eigene Lücken im Umgang mit wissenschaftlichen Themen besser identifizieren zu können.

Dies wiederum verweist auf einen weiteren Aspekt der Integration, der sich mit Fragen der Studienberatung und Betreuung umschreiben lässt. Die Organisation der Lehre findet hierin ein Äquivalent in der Organisation des Lernens. Die Anforderungen an selbstorganisiertes Lernen sind in allen Fächern hoch. Während allerdings in Diplomstudiengängen das Studium zumindest im Grundstudium straffer organisiert ist, gestaltet sich die Studienplanung in Magisterstudiengängen deutlich schwieriger. Die großen Wahlmöglichkeiten im Hinblick auf die Nebenfächer bei gleichzeitig geringer Verpflichtung auf spezifische Veranstaltungen suggeriert, dass Studierende in frühen Semestern bereits einen Überblick über ihre gewählten Fächer und die möglichen sinnvollen Verzahnungen von Veranstaltungen und Schwerpunkten haben. Dies jedoch scheint eher zweifelhaft. Die Äußerungen in den Gruppengesprächen mit Studierenden, aber auch Absolventenbefragungen, die im Rahmen der Evaluationen an der Universität Mainz durchgeführt wurden, deuten darauf hin, dass eine stärkere Verpflichtung auf die Kerninhalte des Fachs sowie angemessene und transparente Leistungsanforderungen bei gleichzeitig guter Betreuung durch die Dozenten ein von vielen Studierenden präferiertes Studienmodell ist.

Diese Kombination von hohen, aber nachvollziehbaren Anforderungen und intensiver Betreuung führt, wenn sie frei ist von den Zynismen des Lehralltags, zu einer engeren Bindung an das Fach, damit aber auch zu einer erhöhten Motivation für das Studium.²⁸ Hierbei kommt vor allem der unter dem Begriff der ‚Betreuung‘ zusammengefassten sozialen Beziehung zwischen Dozenten und Studierenden eine zentrale Rolle zu. Geht man davon aus, dass soziale persönliche Anerkennung ein basales

²⁸ Vgl. zu dem Ansatz der Bindung an Handlungsprogramme Manfred HENNEN, Thomas REIN (1994): *Bindung und Motivation als implizite Annahmen der „Rational Choice“ Theorie.*

Handlungsmotiv ist, so wird nachvollziehbar, dass Betreuung sich insbesondere auf qualifizierte persönliche Rückmeldungen bezieht. Es ist bezeichnend, dass Studierende in sogenannten Massenfächern, wie bspw. der Germanistik, erstaunt sind, wenn sie von Dozenten namentlich angesprochen werden. Eine solche, nicht zuletzt aufgrund knapper personeller Ressourcen gewachsene Fachkultur wird der Tatsache nicht gerecht, dass es sich bei Universitäten in erster Linie um einen sozialen Raum handelt. Schon alleine die unaufgeforderte Nachfrage an fortgeschrittene Studierende, wann sie ihren Abschluss planen, oder eine persönliche Einladung zu einem Examenskolloquium kann unserer Erfahrung nach oftmals tiefer greifen als strukturelle und curriculare Maßnahmen zur Verkürzung der Studienzeit.

Die Kultur eines Faches im Sinne einer akademischen Gemeinschaft zeigt sich jedoch nicht nur an der Art und Weise des Austauschs zwischen Dozenten und Studierenden, sondern in gleichem Maße an der Verständigung unter den Studierenden selbst. Auch hier lassen sich wiederum deutliche Unterschiede zwischen naturwissenschaftlichen und den übrigen Fächern feststellen. Während die stärkere ‚Verschulung‘ des Grundstudiums sowie die Organisation in Arbeits- und Lerngruppen in den Naturwissenschaften den Austausch zwischen Studierenden fördert, gilt es unter Studierenden in vielen geistes-, sozial-, wirtschafts- und rechtswissenschaftlichen sowie in sprach- und kulturwissenschaftlichen Fächern aufgrund der hohen Zahl an Studierenden, aber auch aufgrund der Struktur der Magisterstudiengänge als eher unwahrscheinlich, dass Studierende sich während ihres gemeinsamen Studiums häufiger begegnen. Somit findet die zuvor beschriebene vergleichsweise geringere Forschungsoperation in den genannten Fachgebieten bereits während des Studiums eine Entsprechung: An Stelle der bei starker Differenzierung von Fachwissen und der zunehmenden Ausbildung von Teildisziplinen notwendigen Zusammenarbeit ist nach wie vor eine Kultur des ‚Einzelkämpfertums‘ zu beobachten, die durch die bestehenden Fachstrukturen gestärkt wird.

Schließlich, so lässt sich bereits aus der Diagnose der nicht ausreichenden Abstimmung von Lehrinhalten und Leistungsanforderungen folgern, weisen die Evaluationen darauf hin, dass auch die Kommunikation nicht nur zwischen den Lehrenden, sondern ebenso zwischen den wissenschaftlichen und nichtwissenschaftlichen Mitarbeiterinnen und Mitarbeitern insgesamt verbesserungswürdig erscheint. Hierbei zeigen die Gruppendiskussionen mit wissenschaftlichen und nichtwissenschaftlichen Mitarbeite-

rinnen und Mitarbeitern sowie mit Lehrbeauftragten, dass es weniger um eine grundsätzliche Infragestellung von gewachsenen Hierarchien geht, als vielmehr um die Transparenz von relevanten Informationen und Entscheidungen. Zum Teil ist zu beobachten, dass bspw. Initiativen von Mitarbeitern zur Reorganisation der Lehre in den Instituten keine Rückmeldung erfahren und quasi ‚lautlos‘ verloren gehen. Hierbei kommt zum Tragen, dass Fächer kaum interne Informationsorgane und -wege kennen, die jenseits der informellen Weitergabe von Informationen zu einer höheren Transparenz und der Herstellung gemeinsamen Organisationswissens beitragen. Die in der Organisationsforschung unter dem Begriff des Wissensmanagements diskutierten Problemlagen haben die Universitätspraxis offenbar bislang kaum erreicht. Gerade aber die hohe personelle Fluktuation an Universitäten macht den bewussten Umgang mit und die strukturierte Weitergabe von internem Organisationswissen notwendig.

Dies ist im Übrigen auch hinsichtlich des Kontaktes zwischen wissenschaftlichen Einrichtungen und der Hochschulverwaltung festzustellen. Die zum Teil für die Fächer und Fachbereiche komplizierte Materie des Verwaltungshandelns eignen sich Generationen von Wissenschaftlern stets neu an, ohne auf bereits vorhandenes Wissen zurückgreifen zu können. Die *vorübergehende* Partizipation an der universitären Selbstverwaltung legt es jedoch geradezu nahe, vorhandenes Wissen zu sichern und weiterzugeben. So ist bspw. bei einer durchschnittlichen Verweildauer von zwei Jahren im Amt eines Dekans und gleichzeitig schlechter Personalausstattung in vielen Dekanaten das Vorhandensein entsprechender Wissensbestände oft entscheidend für die angemessene Wahrnehmung der Fachbereichsinteressen.

4.4 Die Dimension des Selbstverständnisses

Unter dem Selbstverständnis von Fächern und Fachbereichen sowie der Universität insgesamt sind solche Prozesse zu verstehen, die zur Selbstvergewisserung über die langfristigen Zielsetzungen und die Reflexion über die Organisationseinheiten beitragen. In diesem Sinne kann die Evaluation selbst, wenn sie in einem umfassenderen Sinn nicht lediglich als Erhebung von Leistungsmerkmalen, sondern als Mobilisierung zur Definition und Umsetzung von Qualitätsstandards verstanden wird, dieser Dimension zugeordnet werden.

Neben der Evaluation wird zur Zeit mit unterschiedlichen Zugängen der Rechenschaftslegung oder der Leitbildentwicklung sowie der Festschreibung einer ‚corporate identity‘ ein Prozess an den Universitäten eingeleitet, der darauf ausgerichtet ist, die Aufgaben und Systemgrenzen der Hochschulen zu beschreiben bzw. zu definieren. Vor allem im Hinblick auf die Entwicklung von Leitbildern ist hierbei grundsätzlich zu differenzieren zwischen Bemühungen, die in Form einer Profilbildung das jeweils spezifische Bild einer Universität zu er- und vermitteln versuchen, und auf umfassendere Zusammenhänge zielenden Fragen nach der Stellung und (funktionalen) Einbindung der Hochschulen in der Gesellschaft.

Wie bereits angeführt, soll an dieser Stelle nicht weiter auf die Diskussion um mögliche Leitbilder eingegangen werden, die bereits in der Vergangenheit unter dem Begriff der ‚Idee der Universität‘ geführt wurde und mithin keinesfalls neu ist. Vielmehr interessiert, inwieweit in den evaluierten Fächern und Fachbereichen ein Bewusstsein für solche Fragen und entsprechende Orientierungen vorhanden ist. Hierbei zeigt sich, dass viele Fächer sich äußerst schwer damit tun, ihre eigenen Stärken und besonderen Leistungsmerkmale zu beschreiben. In vielen Fällen wurde deutlich, dass ein Gelingen der Darstellung des eigenen Fachs gegenüber Fachleuten, aber auch gegenüber den eigenen Studierenden nicht selbstverständlich ist. Einige Evaluationen zeigten, dass die Fächer nahezu konsterniert davon waren, wie sie von ihren Studierenden wahrgenommen wurden.

Damit einher geht, dass Kriterien für gute Forschung und Lehre für viele Fächer äußerst schwierig zu bestimmen sind. Nach möglichen Indikatoren für die Qualität von Forschung und Lehre befragt, konnten Hochschullehrer oftmals zwar darlegen, welche Kollegen gute Forschung betreiben und dass es grundsätzlich einfach sei, gute von mittelmäßigen oder leistungsschwachen Studierenden zu unterscheiden. Doch fiel es nahezu allen Befragten schwer, die für diese Einschätzungen zu Grunde gelegten Kriterien zu benennen.²⁹

Ein entsprechendes Bewusstsein über die Ziele und das Selbstverständnis des Faches ist meines Erachtens keinesfalls marginal und ist bestenfalls nur indirekt über die Produktion von Hochglanzbroschüren und der damit erhofften Außendarstellung zu lösen. Wichtiger ist vielmehr, dass bei Studierenden und Lehrenden ein gemeinsames Verständnis über grundle-

²⁹ Einschränkend muss angemerkt werden, dass die Definition von Qualitätsmerkmalen wiederum in den Naturwissenschaften offensichtlich aufgrund der größeren Anerkennung von führenden Zeitschriften (wie bspw. *Nature*) und den darauf aufbauenden Indizes leichter fällt; allerdings sind die sogenannten ‚impact-Faktoren‘ keineswegs unumstritten.

gende Zugänge und Erwartungen vorhanden ist, um daraus Anforderungen und Leistungsdefinitionen des Fachs ableiten zu können. Der Entwurf eines so verstandenen ‚Leitbildes‘, das im Sinne einer funktionalen gesellschaftlichen Einbindung von Hochschulen beschreibt, welche Aufgaben Universitäten und mit ihnen einzelnen Fächern zufallen, ist kein einfaches Unterfangen. Das Zusammenfügen von ‚ad hoc-Thesen‘ ohne Rückgriff auf wissenschaftshistorische und gesellschaftswissenschaftliche Vorüberlegungen greift hier mit Sicherheit zu kurz. Die Anstrengung einer solchen Definition der Aufgaben von Universitäten zu unterlassen, bedeutet allerdings, die gegenwärtige Ohnmacht der Universitäten bzw. der Hochschulen insgesamt in der hochschulpolitischen Diskussion festzuschreiben, denn den vielfach als Zumutungen empfundenen und häufig wechselnden öffentlichen Erwartungen können die Hochschulen in einer angemessenen Weise nur dann begegnen, wenn sie in einem doppelten Sinne Selbstbewusstsein entwickeln, das auf der begründeten Positionierung in der Gesellschaft beruht.

5. Zusammenfassung

Ausgehend von unterschiedlichen Diagnosen zur Krise der Hochschulen konnte gezeigt werden, dass die damit verbundenen Beschreibungen nur unzureichend die Interdependenzen abbilden können, denen Hochschulen als komplexe Organisationen unterliegen. Der häufige, zum Teil dem Zufall entsprungene Themenwechsel in der hochschulpolitischen Diskussion ist – so die hier vertretene These – nicht dazu geeignet, Hochschulreformen rational zu steuern.

Der Zugang über die Betrachtung der Universität als System soll in diesem Sinne dazu beitragen, ein *mögliches* Ordnungsschema zu entwerfen, das dazu dient, die Gleichzeitigkeit und die Notwendigkeit zur Ausbalancierung divergenter Ansprüche an Forschung und Lehre aufzuzeigen. Gleichzeitig kann ein solcher Zugang bei sinnvoller Operationalisierung die Fragen und Themen von Evaluationen leiten und ein Bezugssystem für die Hochschulentwicklung und entsprechende Maßnahmen herstellen.

Die hier dargestellten Ergebnisse der Evaluationen legen eine solche Herangehensweise nahe, zeigen sie doch die Ambivalenzen, denen Hochschulforschung und -lehre ausgeliefert sind. Antworten darauf zu finden,

wie bspw. die Praxisrelevanz des Studiums eingelöst werden kann bei gleichzeitiger Erfüllung der Forderung nach ‚exzellenter‘ Wissenschaft, welches die primären Ziele der Universitäten sind und ob z.B. die Ökonomisierung der Hochschulausbildung nicht bereits dazu geführt hat, dass die Einwerbung von Drittmitteln nicht mehr als Ressource für Forschung und Lehre, sondern bereits als Selbstzweck zu verstehen ist, scheint unumgänglich, um die hochschulpolitische Diskussion zu versachlichen und damit partiell auch zu entpolitisieren. Denn nicht nur aus systemischer Perspektive kann die Logik der Wissenschaft nicht äquivalent mit der politischen Logik sein.

Die Zukunft der Evaluation

Manfred Hennen

1. EINLEITUNG	83
2. BEDEUTUNG UND SELBSTVERSTÄNDNIS DER EVALUATION.....	85
2.1 Analyse und Aktivierung	85
2.2 Wissen zweiter Ordnung	86
2.3 Steuerungshilfen	87
2.4 Die Rolle der Theorie	91
2.5 Handeln und Motivation.....	94
3. FESTLEGUNGEN UND GRENZEN	100
3.1 Evaluationsbegriff	100
3.2 Qualitätsbegriff	100
3.3 Grenzen der Standardisierung	103
3.4 Partnerschaft	105
3.5 Aufgabe und Grenzen der Experten	107
4. EVALUATION ALS BERATUNG.....	109
4.1 Intermediäre Instanzen	109
4.2 Selbständigkeit und Weiterungen.....	112
4.3 Sprache und Loyalität	116
4.4 Logik des Dialogs.....	118
4.5 Der Mechanismus der Defektion	128
5. LEITBILDER	130
5.1 Externer und interner Werbefeldzug	130
5.2 Idee und Leitidee	133
5.3 Psychische und soziale Mechanismen	137
5.4 Organisation und Institution.....	139
6. ZUSAMMENFASSUNG.....	142

1. EINLEITUNG

Die folgenden Erörterungen beruhen auf den praktischen und theoretischen Erfahrungen im Umgang mit Evaluation, wie sie in den vorangegangenen Beiträgen beschrieben worden ist. Das Verfahren der Evaluation, wie es in den Fächern und Fachbereichen der Johannes Gutenberg-Universität Mainz durchgeführt worden ist, hat sich im Laufe der Zeit in seinen Schwerpunkten in dem Maße verändert, wie neue Erkenntnisse nicht nur aus den Evaluationen selbst gewonnen worden sind, sondern auch aus den anderen Tätigkeitsbereichen im Umgang mit Hochschule und deren Fragestellungen sowie anderen Einrichtungen. Ich nehme diese Erfahrungen und Erkenntnisse zum Anlass, einen Ausblick auf wahrscheinliche Entwicklungen und neue Fragestellungen zu wagen. Die verwendeten Beispiele beziehen sich überwiegend auf eigene Erfahrungsbereiche, weil ich damit authentische Auskünfte geben zu können glaube.

Kritische und selbstkritische Überlegungen über die Zukunft eines Verfahrens mit beispielloser Karriere vom Schreckgespenst über eine Plauderveranstaltung zum umfassenden Reformkonzept sollten gerade diejenigen hilfreich und sinnvoll erscheinen, die sich beruflich mit Bestandsaufnahmen und Reorganisationen in Organisationen befassen. Wer für die eigene Arbeit den Namen Evaluation in Anspruch nimmt, sollte sich nicht scheuen, sich selbst mit den Fragen zu beschäftigen, welche die eigene Arbeit betreffen und dort nach möglichen Fehleinschätzungen zu suchen und Überlegungen darüber anzustellen, welche Entwicklungen absehbar erscheinen und welche Neuorientierungen notwendig werden könnten.

Für solche Überlegungen war in meinen Augen außerordentlich hilfreich, dass uns im Umgang mit Hochschule und Bildungseinrichtungen unterschiedliche Perspektiven zur Verfügung gestanden haben, die verglichen und zusammengeführt werden konnten. Insbesondere waren es Evaluationsverfahren in anderen Bildungseinrichtungen außerhalb der Mauern der eigenen Hochschule, Beobachtungsverfahren unterschiedlicher Art, Maßnahmen der Organisationsentwicklung, eigene empirische Projekte und andere Formen der Hochschulforschung, die zu einer Zusammenschau herausfordern. Dazu gehören auch Maßnahmen und Untersuchungen über die Entwicklung in Schulen, z.B. über Wahlen und Abwahlen naturwissenschaftlicher Fächer. Hervorheben möchte ich auch Einblicke, die mir der Leiter des Projektes zur Förderung des wissenschaftlichen Nachwuch-

ses bzw. der Habilitandenförderung¹ in für mich wertvollen Gesprächen gewähren konnte. In seiner Arbeit sind nicht nur die Nachwuchswissenschaftler und Arbeitsgruppenleiter in den Naturwissenschaften zu Wort gekommen, sondern auch deren Vorgesetzte und sonstige Professoren mit amtlichen Zuständigkeiten.²

Meine eigenen, in unverstellter, mir mit kollegialem Entgegenkommen gewährten Sachgespräche gehören für mich zu den Höhepunkten der Einsichtnahme in die wirklichen Probleme und ungelösten Fragen. Sie geben aber auch Anlass zur Hoffnung auf Chancen für institutionelle Reorganisation. Unter Bedingungen offener und ernsthafter Verständigung fließen viele unterschiedene Einschätzungen zusammen, für die in der hochschulpolitischen Öffentlichkeit wenig Platz ist. Diese Erkenntnisse lassen sich bündeln und machen die Beschäftigung mit der Aufrechterhaltung von Wissen, seiner Erzeugung, Umwälzung sowie seiner Weitergabe in einer Gesellschaft unabweislich, die davon schicksalhaft abhängig geworden ist. Wir müssen Wege finden, durch dieses Insiderwissen die Öffentlichkeit zu aktivieren und zu nutzen. Das Licht der Öffentlichkeit sollte nicht gescheut werden, sondern zur Verfügung stehen und geschätzt werden.

Zunächst werde ich mich mit der Frage auseinandersetzen, ob Evaluation als ein Verfahren in theoretischer oder praktischer Absicht anzusehen ist und was zur Klärung des Status einer theoretischen Fundierung beizutragen ist. Hier wird insbesondere dem Problem nachzugehen sein, was unter Motivation zu verstehen sei. Dabei spielt auch die praktische Seite der Aktivierung von Programmverantwortung und Programmentwicklung bei den Organisationsmitgliedern eine wichtige Rolle. Ich wende mich dann der Bedeutung von Evaluation zu. Wie ist die Karriere eines Begriffs und eines Verfahrens zu erklären? Kann Evaluation als eine Ersatzsteuerung verstanden werden?

Im dritten Abschnitt geht es um Definitions- und Abgrenzungsbemühungen; der Evaluations- und der Qualitätsbegriff stehen hier im Vordergrund. Evaluation im hier verstandenen Sinne ist keine hoheitliche Maßnahme politischer Steuerung oder zur Sanktion von Organisationsmitglie-

¹ Die Koordinationsstelle zur Förderung der Lehrkompetenz des wissenschaftlichen Nachwuchses wird geleitet von PD Dr. Manfred Herzer; zum Programm vgl.: <http://www.verwaltung.uni-mainz.de/ZQ/wn-ueberblick.htm>

² Die im Folgenden verwendete männliche Form für Berufsbezeichnungen gilt gleichermaßen für beide Geschlechter.

dern. Sie ist ein Verfahren zur Unterstützung von Reorganisationsprozessen. Diese sind unverzichtbar angewiesen auf die wenigstens latente Bereitschaft der überwiegenden Zahl von Verantwortlichen in Organisationen Organisationsdefizite als Problem zu behandeln. Dazu sind Partnerschaften herzustellen. Nur auf dieser Basis ist Mobilisierung von Verantwortung möglich. Die Expertenbegutachtung gilt als selbstverständlich. Ist das wirklich gesichert?

Ich komme zu dem Ergebnis, dass Evaluation in eine Form der Beratung übergehen muss. Dazu kommt der Rechtsform der beratenden Instanzen, die über eigene Autorität und die Fähigkeit der Neutralität verfügen müssen, eine besondere Bedeutung zu. Ich bezeichne sie als intermediär. Ihnen kommt als eine zentrale Aufgabe zu, Maßstäbe zu entwickeln und festzulegen, vermittels derer Qualität gemessen werden kann. Meine besondere Aufmerksamkeit gilt der Überwindung von Kooperationsblockaden in Organisationen sowie der Offenlegung der Verweigerung von Verantwortung gegenüber Reorganisationsprozessen. Dazu müssen sprachlich verfasste Einschätzungen von Organisationsmitgliedern aus Dialogen gewonnen und zu synthetisierten Dialogen zusammengeführt werden.

Schließlich befaße ich mich mit dem Begriff des Leitbildes für Organisationen. Was kann darunter sinnvoller Weise verstanden werden und welche motivierende Kraft kann davon ausgehen? Dazu ist es notwendig, auf synonyme Begriffe zurückzugreifen.

2. Bedeutung und Selbstverständnis der Evaluation

2.1 Analyse und Aktivierung

Evaluationen sind gefragt und scheinen es auf lange Sicht zu sein. Mit Selbstverständlichkeit wird der Begriff in Gesetzestexte und Reformprogramme aufgenommen. Überall wird mit Nachdruck davon gesprochen, dass Evaluation wichtige *Steuerungsfunktionen* zu übernehmen habe. In einer solchen Atmosphäre der Evaluationseuphorie und in dieser Zeit der Konjunktur und der Akzeptanz eines Verfahrens zur Analyse und Bewertung von Einrichtungen, Programmen, Innovationen und Reformen unter dem Aspekt ihrer Wirksamkeit und Qualität wären Evaluationseinrichtungen schlecht beraten, wenn sie von der Fraglosigkeit eingespielter Methoden und Abläufe ausgingen und sich auf dem Erfolg des scheinbar

Selbstverständlichen ausrichteten. Wer Situations- und Problemdiagnose zum Zweck der Mobilisierung von Reformbewusstsein und Aktivität mit Augenmaß betreibt, sollte sich um eine *professionelle Skepsis* bemühen, die für sich abzeichnende neue Fragen sensibel ist.

2.2 Wissen zweiter Ordnung

Fraglos sind alle Gesellschaften und Kulturen abhängig von Wissen, aber noch nie hat die Akkumulation in Verbindung mit der Anwendung von Wissen so danach verlangt, Wissen wissentlich zu organisieren, wie in den gegenwärtigen Industrie- und Dienstleistungsgesellschaften. ‚*Wissensmanagement*‘ ist dieser Prozess einer neuen Ebene von Wissen getauft worden. Dieser Begriff sollte jedoch gerade nicht zur Überschätzung der Steuerungsmöglichkeiten Anlass geben, dazu gibt es keinen Grund. Hier geht es mir um eine Auflistung und Systematisierung von einigen wichtigen Aspekten zum Komplex der Evaluation, um die Abfassung einer Standortbestimmung.

Es geht ja keineswegs nur um Wissen, das personen- und ortsungebunden wäre; nicht einmal organisationsinternes Wissen steht den Verantwortlichen wie selbstverständlich zur Verfügung. Immer mehr kommt es auf die Zusammenführung unterschiedlichen Wissens an. Die ungeheuren Differenzierungsprozesse unserer Gesellschaften haben auch zu einer Partikularisierung, einer *Einhegung des Wissens* geführt. Wir schwärmen davon, dass die Globalisierung und die elektronischen Medien – allen voran das Internet – Wissen in bisher ungeahnter Weise zugänglich machen, aber das ist nur die Vision von Möglichkeiten, jedoch damit noch keine Garantie für die Verschränkung und damit Verfügbarkeit der entscheidenden Wissensbestände. Gerade in der Evaluationsarbeit kann man sich ein Bild von der unglaublichen Schwierigkeit machen, unter der selbst in einer mittleren Organisation getrennte Wissenswelten zusammenzuführen und kompatibel zu machen sind.

Ausdruck dieser skeptischen Mahnung ist die Erfindung eines Instrumentes des Wissensmanagements, das ‚*Mikroart*‘ genannt wird und in Anlehnung an die Praxis wissenschaftlicher Artikel für Mikroartikel steht. Organisatoren aller Art sollen sich nach diesem Modus in der Praxis wie Wissenschaftler verhalten, die einen Aufsatz schreiben um Rechenschaft über den Stand ihres Handwerks abzulegen. Auch als neu angepriesene

Gerichte können nach alten Rezepturen gekocht werden. Ein solcher Artikel sei verstanden als ein Beitrag der „Revision von Wissen, weil seine Nutzung Anlass zu Kritik, Erweiterung, Nachfragen und neuen Anregungen“ verlangt.³ Wie man sieht, ist offenbar sogar den Beteiligten selbst nicht bewusst, was sie wissen und sie müssen sich erst Rechenschaft darüber ablegen. Wie viel schwerer muss das erst beim Aufspüren von Wissensressourcen und bei deren Zusammenführung in größeren Systemen sein. Wir benötigen ein *Wissen zweiter Ordnung*, eine Kunst und eine Systematisierung für das Synthetisieren von Wissen. Die Evaluation wird darauf stets aufs Neue verwiesen.

2.3 Steuerungshilfen

Neben diesem Befund gibt es weitere Gründe, welche es sinnvoll machen, die Frage zu stellen, wie es kommen konnte, dass Evaluation so populär geworden ist. Für welche Defizite vermag sie Abhilfe zu versprechen? Ich glaube, dass dafür die durchgehende Erfahrung von *Steuerungsresistenzen* in unternehmerischen, administrativen und politischen Anordnungs-, Leitungs- und Managementformen verantwortlich ist. Das sind keine Entwicklungen, die sich von heute auf morgen herausstellen. Sie zeichnen sich langsam ab, werden erst nach und nach manifest um schließlich zu einem durchgehenden Erfahrungs- und Kritikprogramm zu werden.

Im Anschluss an das oben Gesagte kann man allgemein formulieren: Evaluationen werden dort vorrangig notwendig, wo ein hoher Grad an *Autonomie von Wissen* unterschiedlichster Art, dessen intelligente Verwaltung und Koordination und die Mobilisierung von verantwortungsvoller Rechenschaftsablegung zusammenkommen müssen. Diese Entwicklungen in Organisationen innerhalb und außerhalb des Hochschulwesens haben systematisch die Evaluation als ein Instrument eines ergänzenden, genauer *subsidiären Steuerungsverfahrens* erscheinen lassen, als eine Hilfestellung oder einen Behelf für ansonsten missglückende Normalsteuerung. Evaluation ist auf diese Weise zu einem Hoffnungsträger für die Gestaltung steuerungsresistenter Funktionsbereiche avanciert.

Es ist aufschlussreich, dass sich die Ratlosigkeit in Steuerungs- und Mobilisierungsfragen weniger auf die i.e.S. technischen und administrativen

³ Helmut WILLKE (1998): *Systemisches Wissensmanagement*, 100ff.

Vorgänge erstreckt, sondern die damit verbundenen *sozialen Transaktionsprozesse*. Einige Beispiele aus dem Hochschulalltag sollen das verdeutlichen, sie lassen sich aber ohne Weiteres auf andere Organisationen und auch auf Unternehmen übertragen. Nicht die Tatsache eines bis zur Überfälligkeit unangepassten Curriculums ist das eigentliche Problem eines Faches oder Fachbereichs, sondern die ausbleibende Wahrnehmung dieses Missstandes, die ausbleibende Korrektur, die ungeklärte oder geleugnete Zuständigkeit.

Es ist nicht selten der Fall, dass diese Mängel bereits verschiedentlich angesprochen, aber immer wieder verdrängt worden sind. Das liegt besonders am Mangel einer übergreifenden organisatorischen Zuständigkeit und Verantwortung. Auch Geschäftsführende Leiter, Dekane oder sogar Hochschulleitungen müssen wegen einer in anderer Hinsicht bewährten Kollegialstruktur befürchten, sich Mund oder Finger zu verbrennen. Wer ärgerliche Tatsachen der mangelnden Koordination, Absprachen und Verständigung zur Sprache bringt, muss damit rechnen, als Störenfried oder Nörgler gebrandmarkt zu werden.

Will man dies revidieren, benötigt man zunächst einmal die Perspektive des Außenstehenden, des Dritten, des Unparteiischen, der überhaupt erst auf die unabweisbare Notwendigkeit von Kommunikationen, Verhandlungen, Entscheidungen, Kooperationen, Kontrolle und Verantwortung aufmerksam macht. Die Ökonomen bezeichnen diesen Aufwand als *Transaktionskosten*. Transaktionen stellen Aushandlungs- und Übertragungsleistungen dar, die dort, wo sie nicht als selbstverständlich eingeführt und gesichert sind, knapp werden und als Zumutung angesehen werden. Transaktionen sind Leistungen von Person zu Person, von Kollektiv zu Kollektiv und soziale Leistungen. Das erklärt, warum sie leicht als nachgeordnet oder sogar inferior angesehen werden. Leistungen technischer und produktiver Art gelten als eigentliche Produktionsschwerpunkte. Transaktionen werden als etwas Selbstverständliches, etwas Natürliches oder gar Triviales angesehen, über das man sich nicht unnötig den Kopf zerbrechen sollte.

Ein weiteres Beispiel für die soziale Qualität der Defizite, die uns häufig begegnen: Nicht die Tatsache, dass bürokratisch falsche Entscheidungen sich innerhalb der Hochschulverwaltung ereignen ist das eigentliche Problem in der Organisation, sondern deren Hinnahme und das Ausbleiben einer notwendigen Neuordnung. In diesen Fällen wird deutlich, dass die technisch-administrativen Fehlleistungen für sich genommen schon vor Rätsel stellen können, aber dass deren Problematik erst durch die

Sichtbarmachung der Dimension eines *sozialen Transaktionsdefizits* erscheint und den eigentlichen Handlungsbedarf offen legt. Die Grenzen der Administration liegen eindeutig im Sozialmanagement. Es fehlt an Kooperationsleistungen, Entscheidungsleistungen, Überzeugungsleistungen und vor allen Dingen an Motivationsleistungen.

Evaluation ist keineswegs ausschließlich, aber sehr oft mit dieser sozialen Dimension von Prozessen befasst. Ich erinnere mich an mehrere Fälle, bei denen in Fächern und Fachbereichen heikle Fragen notwendiger Reorganisation von Forschungs-, Studien- und Organisationsstrukturen anstanden, von denen niemand aus dem Kreis der Verantwortlichen oder der Experten hätte sagen könnten, dass sie im Handstreich zu lösen und leicht zu entscheiden gewesen wären. Und doch ist in diesen sachlichen Fragen nicht der Kern des Problems zu suchen. Entscheidungen können schwierig, aufwendig und sogar schmerzhaft sein, aber sie müssen vor allen Dingen angegangen werden.

Das Aussitzen-Wollen hat meistens nur die politische Dimension, nicht angreifbar zu sein, selten besteht es im planvollen und sachlich begründeten Verschieben des Zeitpunktes der Entscheidung. Die Unterlassung einer unverzichtbaren Transaktionsleistung unter Fachkollegen oder zwischen Leitungspersonen auf verschiedenen Ebenen stellt das eigentliche Manko dar. Der Widerstand gegen diese Diagnose ist fast ausnahmslos ganz erheblich. Das macht das Leben der Evaluatoren nicht einfach. Evaluation hat es mit einem zentralen Transaktionsproblem zu tun, oft mit einem Kampf der Argumentationen gegen abwehrende Verdächtigungen aller Art. Nicht selten sind Transaktionsprobleme gerade dort von einer latenten Wirksamkeit, wo die Einheitlichkeit der Interessen als gegeben unterstellt wird. Das ist beispielsweise oft in Leitungsgremien der Fall. Was als gebündeltes Führungsinteresse auf unterschiedlichen Ebenen unterstellt wird, entpuppt sich oft als ein System schnell wechselnder Koalitionen. Das Problem wird nicht kleiner dadurch, dass gerade dies auch unter den Beteiligten geleugnet werden muss. Selbstverständnis, Präsentation und Wirklichkeit der Kooperation klaffen erheblich auseinander.

Die hier beschriebene Steuerungsresistenz in Organisationen im Sinne einer Blockade der sozialen Dimension der Transaktion hat neben der erwähnten Autonomie, die in der Kompetenz der Personen und Organisationseinheiten liegt, noch andere Ursachen, über die zu reden ist. Das Universitätsleben gibt wie der öffentliche Sektor allgemein viele Beispiele dafür, dass die verfassungsrechtlich garantierte Selbständigkeit von

Personen im Rahmen der *Wissenschafts- und Lehrfreiheit* zu Potentialen der Verweigerung von Transaktionsleistungen und als Folge davon zu Resignation führen. Diese grundlegenden Freiheiten sind hohe Rechtsgüter, deren große Bedeutung stets gerade in ihrer Gefährdung sichtbar und sinnfällig geworden ist. Diese Gefährdung hat sich besonders in der Form des politischen Übergriffs gezeigt. Dies ist eine Gefahr, die besonders gerade deshalb nie übersehen werden darf, weil diese Formen der Einflussnahme dem Wandel unterliegen können und damit die Gefahr verdecken.

Aber selbst Hochwertkonstellationen sind nicht dagegen gefeit, für Minderzwecke in Anspruch genommen zu werden. Wo im Sinne des Wissenschaftszwecks notwendige Transaktionen mit dem Hinweis auf Wissenschafts- und Lehrfreiheit zurückgewiesen werden, findet eine *Ausbeute eines kostbaren Rechtsgutes* statt, dessen Gefährdung auf diese Weise ganz andere Formen annimmt.

Schutzgarantien haben nicht selten den unbeabsichtigten und paradoxen Effekt, auch dort für Steuerungsresistenz zu sorgen, wo das nicht durch die primäre Schutzidee gedeckt ist. Daraus entstehen unterschiedlichste Abweisbarkeitschancen. Unsere Erfahrungen haben sehr viele solcher Unzugänglichkeiten ans Tageslicht gebracht. Die Gefahr ist groß, dass sich in diesen Fällen Zynismus auf der einen Seite, Resignation auf der anderen breit macht. Solche Barrieren sind aber nicht eindimensional als Verweigerung zu charakterisieren, das wäre unangebracht.

Wo Verweigerungen strukturell möglich sind, wird Selbstzuschreibung von Verantwortung wichtig. Damit entstehen Formen der Verantwortungs-, Zuschreibungs- und Selbststeuerungsabhängigkeit, die deshalb prekär sind, weil sie eine *moralische Sonderleistung* darstellen, von der zunächst einmal nicht auszugehen ist. Vielleicht ist das eine ganz zentrale Beschreibung der Abhängigkeit moderner Gesellschaften. Wir sind auf den Betriebsstoff *nicht-standardisierbarer Motivationen* angewiesen zur Überbrückung der Steuerungsdefizite und –resistenzen. Das erfordert eine besondere Form der Aktivität, wie Etzioni früh erkannt hat, die weder innerhalb noch außerhalb von Organisationen einfach programmierbar ist. Die angemessene Beschreibung von Situationen und Problemlagen auch in Organisationen erfordert *Authentizität*, die nur durch fortgesetzte Aktivierung erzielt werden kann.⁴

⁴ Amitai ETZIONI(1975): *Die aktive Gesellschaft*, 659.

Aktivität ist aber per se rar und nicht ohne weiteres einzuholen. Je mehr wir auf Umdenken, fortgesetztes Lernen und andere Dinge pochen, desto schwerer wird es sein, solche exzeptionellen Leistungen einzuklagen. Vielleicht liegt hier das Geheimnis eines Verfahrens mit dem Namen Evaluation darin, ein Prozess der außerordentlichen, außerplanmäßigen extern veranlassten Introspektion der Qualität von Leistungen zu sein. Evaluation soll und kann nicht trivial kontrollieren, aber sie kann es nach meinem Verständnis, wenn sie professionell gemacht ist, in einer mobilisierenden und Interesse und Verantwortung auslösenden Weise als *nicht-triviale Aktivierung* tun. Das Interesse sich einer Evaluation zu unterziehen ist in den meisten Fällen sowohl mit dem Bewusstsein eines Defizits an Selbststeuerung verbunden als auch mit der Hoffnung sich freiwillig zu exponieren um mit Hilfe einer neutralen Instanz diesem Defizit entgegenzuwirken.

2.4 Die Rolle der Theorie

An dieser Stelle muss aber nicht nur vor einer Überschätzung der Planbarkeit von Wissensprozessen gewarnt werden, sondern auch vor derjenigen einer umfassenden Verwissenschaftlichung von Evaluation. Es ist zu unterscheiden zwischen Evaluation als Qualitätsbefund und Qualitätsentwicklung einerseits und Evaluationsforschung andererseits. Das erstere besteht in der Gewinnung eines Urteils über Handlungsvorschriften oder Handlungsanleitungen, die allgemein als *Programme* bezeichnet werden; das letztere entspricht einer quasi-experimentellen Forschungslogik, mit der kausalen Zuschreibung von Effekten explizit gemachter Interventionsprogramme zur Lösung eines definierten gesellschaftlichen Problems anhand einer Theorie. Nur in diesem Fall handelte es sich um ‚*Programmwirkungstheorien*‘.⁵

Ich habe großen Respekt vor der disziplinierenden Kraft theoriegeleiteter Konzepte und verkenne keineswegs die Tatsache, dass Evaluationsfelder mit Größen zu tun haben, die typisch sind für sozialwissenschaftliche Theoriebildung. Wir haben es in Evaluationen stets mit Akteuren zu tun,

⁵ Vgl. den ehrgeizigen Versuch, ein deutsches Lehrbuch theoriegeleiteter Evaluationsforschung zu verfassen, das auf diesen Unterschied durchgehend hinweist: Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT: *Rational Choice und theoriegeleitete Evaluationsforschung*, 11.

mit Handlungen, mit Programmen, mit Intentionen, Zielen und nicht zuletzt mit stets überraschenden Ergebnissen und Wirkungen. Aber es wäre vermessen zu behaupten, wir könnten in der Praxis auch nur annähernd pure kausal-experimentelle Evaluationsforschung betreiben. Wir haben uns stattdessen ständig mit einer Gemengelage zu befassen, in der die genannten Größen alle vorkommen, aber nur begrenzt in formalisierte Forschungskonzepte überführt werden können. Ich bevorzuge eine Sichtweise, wie sie im Konzept der ‚*wissenschaftlich gestützten Evaluation*‘ zum Ausdruck kommt⁶, ohne mich der Sache nach von der Vorstellung zu entfernen, dass Evaluation auf Theorie beruhen muss.⁷

Die Anknüpfungspunkte an eine Rational-Choice-basierte Motivations- theorie sind, wie ich zu zeigen hoffe, überzeugend. Akteure handeln nur scheinbar nach expliziten Programmen, tatsächlich folgen sie *latenten impliziten Handlungsschemata*, die ihnen selbst unvertraut sind oder die sie vor sich selbst verdunkeln. Sie geben vor, Ziele zu verfolgen, die sie gleichzeitig im Alltag ständig verletzen. Sie hegen geheime Intentionen, die nicht zu den Zielen und Programmen der Leitbilder in Hochglanzpapier passen und sie erzeugen gemeinsam im Verbund des Handelns Folgen, welche ihnen als unbeabsichtigt, fremd und paradox begegnen. In einem solchen Feld könnte es im Grenzfall dem Ziel der Evaluation gänzlich widersprechen, auf Hinweise und Eingriffe während der Erhebung zu verzichten, um die Analyse der Wirkungsmechanismen nicht zu gefährden. Das Interesse am Nachweis von programmatischen Fehlleistungen ex post zur Abgabe von anschließenden Empfehlungen könnte von den Auftraggebern geradezu als zynisch bezeichnet werden.

Die gängige Unterscheidung von *summativer und formativer Evaluation* hat ihre Berechtigung. Die formative Evaluation bezeichnet das begleitende, modifizierende und durch Interventionen prägende Verfahren qualitativer Art. Es handelt sich hier um angewandte Forschung im Feld und nicht in einem Quasi-Labor. Es unterstellt sozusagen eine Ratgeberpflicht im Verfahrensverlauf mit diplomatischem Geschick. Die summativ-evaluation urteilt über wissenschaftlich ermittelte und durch Einwirkungen im Verlauf ungetrübte Effekte am Ende der Beobachtungsperi-

⁶ Heinrich WOTTAWA/ Heike THIERAU (1998): *Lehrbuch Evaluation*, 12.

⁷ Ich kritisiere nicht die Grundposition von Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*. Ihre theoretischen Formalisierungsvorstellungen überzeugen mich nicht. Wenn sie schlicht sagen könnten, dass Rational-Choice eine Art Leitwährung oder Abfrageschema in Evaluationsverfahren sei, könnte ich mich damit anfreunden.

ode.⁸ Für unsere Auftraggeber war die Verfahrensbegleitung im Sinne der *formativen Evaluation* stets das gewünschte Verfahren. Wir waren immer ohnehin einem Ereignisstrom mit vielen internen und externen Einwirkungen ausgesetzt, der eine kontrollierte Isolation von Einflussgrößen als unmöglich erscheinen ließ.

Hinzu kommt, dass die erwähnten impliziten Vorstellungen keineswegs bei allen Akteuren die gleichen sind. Ein Großteil der Arbeit innerhalb unserer Evaluationsverfahren besteht darin, dass diese unterschiedlichen Vorstellungen überhaupt ins Bewusstsein gehoben werden, damit dann unter den Beteiligten eine Verständigung herbeigeführt werden kann. Evaluation ist zu einem großen Teil Forschung in Verbindung mit aktiver Veränderung. Das ist auch als *Aktionsforschung* bezeichnet worden, freilich mit dem Anspruch eines besonderen theoretischen Hintergrundes. Das ist hier nicht der Fall. Evaluation ist angewandte Theorie, nämlich Handlungs- und Motivationstheorie im praktischen Interventionsvollzug. Demgegenüber sollte die Bezeichnung ‚Evaluationsforschung‘ auf den Bereich exakter puristisch-experimenteller Wirkungstheorie beschränkt bleiben, auf reine theoretische Erkenntnis, welche allenfalls danach für praktische Entscheidungen hinzugezogen werden könnte.

Die Bemühungen, die erwähnten impliziten Vorstellungen der sogenannten Programmverantwortlichen und Programmbeteiligten als Ursache-Wirkungsvorstellungen über Zustände in einem Evaluationsfeld theoretisch zu integrieren, erscheinen mir als Modell für die praktische Arbeit so selbstverständlich wie als Theorie suspekt. Dazu ein Beispiel: Wenn in Hochschulevaluationen die Auffassungen von Professoren, akademischen Mitarbeitern und Studierenden über die Ursachen schlechter Studienqualität protokolliert und analysiert werden, erhält man Kausalattributionen von den Prozessbeteiligten, die auch unter den Namen ‚*subjektive Theorien*‘ oder ‚*implizite Theorien*‘ oder ‚*Alltagstheorien*‘ bekannt sind. Diese Theorien von Nichttheoretikern sollen nun mit wissenschaftlichen Theorien verbunden werden, bzw. mit deren Hilfe in eine wissenschaftliche Form gebracht werden.⁹

Theoretische Rekonstruktionen sind aber nach meinem Verständnis etwas anderes als Theorien. Warum aber nur dieser Anspruch an der falschen

⁸ Jürgen BORTZ, Nicola DÖRING (2002): *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*, 61, 103, 105, 107, 112f.

⁹ Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 233ff.

Stelle? So kann man allenfalls zu *Hybrid-Theorien* kommen, nicht aber zu Theorien in reiner Form. Die Alltagstheorien können höchstens eine Art Theorien zweiter Klasse sein; klarer wäre es von Gegenständen der Theorien zu sprechen. Dass die Bearbeitung des Alltagsmaterials mit theoretischem Wissen angegangen wird, sollte eine Selbstverständlichkeit sein.

Wenn ein sozialwissenschaftlich denkender Evaluator weiß, welche typischen Mechanismen dafür verantwortlich sind, dass Organisationsziele pervertieren, wenn er ferner ein analytisches Menschenbild mit empirischer Bewährung besitzt und weiß, wie aus *engen Grundmotiven* anspruchsvolle *Motivationen* in Verbindung mit *Programmen* werden, wenn ihm vertraut ist, wie diese Motivationen stabilisiert werden können oder sich auflösen, wenn ihm Grundlagen der Mehrebenenanalyse geläufig sind und wenn er das alles in seine Betrachtungen mit einbezieht, dann können wir von einer theoriegestützten Evaluation sprechen, aber damit liegt eben noch keine Theorie im strengen Sinn eines analytisch-empirisch getesteten Wirkungszusammenhangs vor.

Nach den Äußerungen dieser Skepsis möchte ich mich mit einigen theoretischen Überlegungen zurückmelden, die befreit sind vom Zwang der unmittelbaren Umsetzung von Theorie in Praxis, aber gleichzeitig von der Notwendigkeit theoretischer Fundierung zu überzeugen vermögen. Das Herz der Evaluation, sagt Patton dem Sinn nach, liege darin, die Wirklichkeit daraufhin zu testen, wo offizielle und tatsächliche Handlungsprogramme dissonant seien.¹⁰ Handlungsprogramme machen Sinn und bewegen Menschen. Wenn wir sie als Evaluationspraktiker in theoretisch fundiertes Rahmenwerk übersetzen, haben wir viel für die Praxis und viel für die Theorie getan.

2.5 Handeln und Motivation

Die *Diskrepanz des Handelns in Programmen offizieller und inoffizieller Art* ist ein Kernstück sozialwissenschaftlichen und auch sozialpolitischen Denkens. Alle Vorbilder oder Entwürfe planvoller, geordneter oder aufeinander abgestimmter Formen sozialen Lebens haben das Merkmal von Programmen des Handelns. Handeln ist nur als zweiseitig organisierter Prozess denkbar. Es ist individuell motiviert und gleichzeitig programm-

¹⁰ Michael Quinn PATTON (1996): *Utilization-Focused Evaluation*, 222.

bestimmt. So wenig es Privatsprachen geben kann, so ausgeschlossen sind *Privathandlungen*.¹¹ Handlungen weisen stets sozial vorgefertigte Elemente und Bausteine auf. Das ist für die Analyse wie für die Gestaltung oder Umgestaltung von Systemen des Kooperierens und des Wissens von großer Bedeutung. Die Präsenz dieses zentralen Gedankens zur Beurteilung von Praxisfeldern ist bei den *Programmverantwortlichen* und *Programmbeteiligten* äußerst rar.

Ich komme auf die Zweiseitigkeit der Motivation zu sprechen, die für Evaluationen ebenfalls eine sehr große Rolle spielt. Ich gehe davon aus, dass die Kopplung von selbstinteressierten allgemeinen menschlichen Strebungen und kultureller überindividueller Programmatik ein zentrales Modell nicht nur der Motivationsforschung ist, sondern auch die praktischen Fragestellungen der Evaluation bedient. Laien sind häufig der Auffassung, diese Perspektive leugne den Einfallsreichtum oder unterschätze die normative Steuerbarkeit der Menschen. Tatsächlich sollen beide Perspektiven auf diese Weise in ihrer Einseitigkeit vermieden werden. Eine Vorstellung von *Handlungsregulation* betont, dass für die Entstehung von Mustern des Handelns soziale und kulturelle Evolutionen verantwortlich sind. Sie betont gleichzeitig, dass Individuen diese Muster in der Reproduktion permanent verändern. Handlung ist immer reproduktiv und produktiv in diesem Sinn.¹²

Der Handlungsbegriff leidet beinahe überall in der Sozialtheorie daran, dass er entweder mit privaten Entwürfen, mit Entscheidungen oder mit Kalkulationen identifiziert wird oder aber mit Strukturmustern gleichgesetzt wird. In einem Fall haben wir es mit einer autonomistischen Schiefelage zu tun, im anderen krankt die Analyse an einem strukturalistischen Bias.

Es gibt zahlreiche Begriffe, welche die Handlungsregulation kenntlich machen sollen. Mit dem Wort ‚Norm‘ beispielsweise wird die Sanktion überbetont. Begriffe wie ‚Habits‘, ‚Frames‘, ‚Scripts‘¹³ treffen den hier angepeilten Sachverhalt schon besser, sie sind aber noch zu statisch um die Wandelbarkeit von Programmen zu erfassen. Es kommt ja gerade unter dem Gesichtspunkt der Bewertung auf die Darstellung der ge-

¹¹ Vittorio HÖSLE (1997): *Moral und Politik. Grundlagen einer Politischen Ethik für das 21. Jahrhundert*, 311ff.

¹² Neuerlich sehr klar: Jürgen RITSERT (2001): *Soziologie des Individuums*. Darmstadt, bes. 15.

¹³ Hierzu: Manfred HENNEN (1994): *Motivation als Konstrukt einer Sozialtheorie*, 166ff.

schickten Abänderung und der Ausbeutung von Programmen an. Handlungsregulationen stellen selbst Systeme von Wahlmöglichkeiten und Erfolgsaussichten dar.¹⁴ Die Kreativität des Handelns ist mit der Handlungsregulation eng verbunden, beide Seiten bedingen sich.¹⁵ Es bleibt noch der Begriff des ‚Schemas‘. Auch hier fehlt aber die Vorstellung, dass es sich im Grunde um ein System von Schemata handelt. Hans Lenk hat deshalb die Bezeichnung ‚Schemaspiele‘ vorgeschlagen. Aber dazu gehören auch konsequenterweise ‚Schemainterpretationsspiele‘.¹⁶ Das wird dann begrifflich zu unattraktiv.

Ich schlage deshalb den Begriff des *Handlungsprogramms* bzw. kurz des *Programms* vor. Er schließt an das Konzept des ‚Verhaltensprogramms‘ an, wie es in den Bio- und Neurowissenschaften üblich ist. Die hier zum Ausdruck kommende Interaktion von Natur und Kultur ist für eine Mehrebenentheorie des Sozialen außerordentlich wichtig. Der Programmbegriff hat zudem in der Evaluationstheorie eine große Bedeutung. Er kann m.E. mit Gewinn auf Handlungen übertragen werden ohne dabei als deterministisch zu gelten; die neueren Entwicklungen auf dem Gebiet der elektronischen Datenverarbeitung halten brauchbare Analogien bereit. Die Programmsoftware ist ja gerade dadurch mehr und mehr gekennzeichnet, dass eine unüberschaubare *Repertoirevielfalt* existiert, welche vom Nutzer kaum durchschaut werden kann, die aber Hintergrundmöglichkeiten bereithält. Die strukturierende Qualität der Nutzungsmöglichkeiten verdeutlicht die relative Handlungsentkopplung. Die *Programmtiefe*, auch wenn sie unbewusst bleibt, existiert auch unabhängig von der Kompetenz der auf der Basis des Programms Handelnden. Programme haben insofern ein Eigenleben.

Die Perspektive ist hier zweiseitig. Handeln wird durch Programme plastisch gesteuert, aber dem Handelnden eröffnen sich auch Wahlmöglichkeiten innerhalb der Programme und zwischen ihnen. Damit entscheidet sich aber das kulturelle und soziale Schicksal der Programme, die ohne ausreichende Akzeptanz ihre kulturelle Existenz verlieren. Gerade hierin offenbart sich der enge *Zusammenhang von Evaluation und Motivation*. Ich gehe soweit zu behaupten, dass alle Formen der Evaluation sich mit

¹⁴ Volker KUNZ (1997): *Theorie rationalen Verhaltens*, 234 und passim.

¹⁵ Manfred HENNEN, Thomas REIN (1994): *Bindung und Motivation als implizite Annahmen der ‚Rational Choice‘-Theorie*, 221ff.

¹⁶ Hans LENK (1995): *Schemaspiele*, 249.

einer doppelten Frage zu befassen haben, die sich ihrerseits wieder zweifach untergliedern:¹⁷

- 1a. Welche Detaillierung haben Interventions- oder Standardprogramme als Programme des Handelns?
- 1b. Welchen Umgestaltungen und Routinen unterliegen diese Programme im täglichen Betrieb? Max Weber hat diesen Vorgang bekanntlich ‚Veralltäglicung‘ genannt.
- 2a. An welche Motive des Selbstinteresses und des Behagens appellieren die offiziellen Programme?
- 2b. Welche Vorteilsmotive des kleineren Karos werden dabei freigesetzt und von inoffiziellen Praktiken genutzt?

Patton's witziger Kipling-gestützter Evaluationsreim ‚Who – what – why – when – where – how‘¹⁸ kommt nicht umhin, immer wieder an die handlungstheoretische Unterscheidung von zwei Ebenen anzuknüpfen. Welche Nutzen-Kalkulationen und Affekte verbinden sich mit welchen Programmen oder induzieren welche Programmmodifikationen? Das Leben und Eigenleben der Programme ist überall im sozialen Leben und auch in Organisationen an Handlungsbereitschaften geknüpft, sie kommen nur als Motivationen vor und als Kette von *Motivationsinduktionen*.

„Jede extensive und intensive Benutzung eines Handlungsprogramms stellt grundsätzlich eine Art Provokation dar. Die Motivationen anderer werden berührt, und es bilden sich Anreize, diese zu verändern. In diesem Sinne ist jedes akute Phänomen, das Gegenstand gesellschaftspolitischer Auseinandersetzungen wird, als Prozeß einer Vielzahl von Motivationsinduktionen zu interpretieren, besonders dann, wenn sehr verschiedene Motivationen sich auf ein kompliziertes Zusammenspiel einlassen müssen.“¹⁹

¹⁷ Vgl. Peter H. ROSSI, Howard E. FREEMAN (1993): *Evaluation. A Systematic Approach*, 105ff.; Peter H. ROSSI, Howard E. FREEMAN, Mark W. LIPSEY (1999): *Evaluation. A Systematic Approach*, 69, 102ff.; Rossi und die jeweiligen Koautoren legen allerdings fast ausschließlich den Akzent auf mehrstufige beabsichtigte Effekte der Evaluation, unintendierte, überraschende und geradezu paradoxe Wirkungen während der Programmverfolgung werden so gut wie nicht berücksichtigt. Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 42ff.

¹⁸ Michael Quinn PATTON (1996): *Utilization-Focused Evaluation*, 298.

¹⁹ Manfred HENNEN (2000): *Soziale Motivation und paradoxe Handlungsfolgen*, 138.

Die Evaluation von Programmen verweist aus diesem Grund auf deren Wirklichkeit in Handlungen und Handlungsbereitschaften. Theorie-basierte Evaluation muss einen engen Bezug zu einer Theorie der Motivation aufweisen.²⁰ Das Grundmodell der Motivation kann hier nur geringfügig vertieft werden. Es beruht auf der Vorstellung einer besonderen Form von *Bindung*. Wie bereits erwähnt, ist zunächst für das Verständnis von Motivationsvorgängen von einer allgemein-menschlichen unspezifischen Antriebsenergie auszugehen, welche im Wesentlichen aus den Bedürfnissen nach psycho-physischem Wohlergehen, nach sozialer Anerkennung, Verlustminimierung und Erfolg besteht. Ich hatte das unter den Begriffen ‚Nutzen‘ und ‚Affekt‘ zusammengefasst. Die zweiseitige Organisation zeigt sich darin, dass den allgemein-menschlichen Antrieben Erwartungen von außen gegenüber stehen, die insgesamt als Handlungsprogramme bezeichnet worden sind. Sie erscheinen in der Form von Zumutungen und Anstrengungen, sei es in Bildungssystemen, im Sport, in moralischen/rechtlichen Institutionen oder in Organisationen. Programme sind Fernziele, die an die zuvor beschriebenen Nahziele gekoppelt werden müssen. Motivation stellt immer *Belohnungsaufschub* dar und ist insofern chronisch knapp und riskiert. Die Kopplung interner Antriebsenergie und externer Zumutung muss als Bindung verstanden werden, als Akt *innerer Verpflichtung*.²¹

Die Sicherung dieser Bindung bedarf der Sicherung in mehreren Dimensionen, um stabil zu sein. Diese sind:

Persönlich, d.h. Programme müssen zur eigenen *Identität* gerechnet werden können.

Zeitlich, d.h. Programme müssen zur Überbrückung von Jetzt und Später eine Durststrecke durchlaufen können. Sie fordern *Moral* ein.

Sachlich, d.h. Programme müssen in der Sache überzeugen können, sie müssen ihre *Rationalität* haben.

Sozial, d.h. Programme müssen dadurch Bindung finden, dass sie bereits bei anderen *Akzeptanz* gefunden haben.

²⁰ Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 97.

²¹ Volker KUNZ (1997): *Theorie rationalen Handelns* 45, 234.

Alle vier Dimensionen spielen im Alltagsleben der Organisationen eine große Rolle; es kommt darauf an, sie zusammen zu führen. Mit einer einzigen Dimension allein kann dauerhaft keine Motivation stabilisiert werden, obwohl das oft erhofft wird. So wird die sachliche Dimension oft überschätzt. Programme erhalten ihre Attraktivität keineswegs ausschließlich kraft der ihnen eigenen Rationalität. Das wird auch im Dilemma von individueller und kollektiver Motivation sichtbar.²² Kollektive Motivation kann zunächst nur in den entworfenen Programmen liegen; diese werden aber nur dann real, wenn sie die Bindung von hinreichend vielen Personen in den vier Dimensionen finden können. Die üblich gewordene Rede vom kollektiven *Lernen in Organisationen* kann mit dem Konzept von Programm und motivationaler Bindung anschaulich gemacht werden. Was Organisationsmitgliedern und Mitarbeitern zugemutet werden muss, ist Bindung; der Anspruch an Programme lautet *Bindungsqualität*.²³

Belohnungsaufschub ist eng an den Faktor Zeit gebunden. Rossi und andere unterscheiden i.d.S. sehr deutlich zwischen unmittelbaren oder ‚proximalen‘ Programmsergebnissen einerseits und den Ziel- oder ‚distalen‘ Programmsergebnissen andererseits.²⁴ Wenn der Umgang mit Zeit durch unangemessene programmatische Verpflichtungen nicht mehr angemessen im Motivhaushalt verankert werden kann, sind die Demotivation und der Verlust von Moral die Folge. Im öffentlichen Sektor zeigt sich das leider oft, in Bildungseinrichtungen dann, wenn „mehr und mehr Zeit in Ausschusssitzungen als im Klassenraum mit dem Unterricht, mit der Beratung der Studenten oder mit der Forschung“ verbracht wird.²⁵

²² Näheres hierzu Manfred HENNEN (1994): *Ist Verhalten beeinflussbar?* 147ff.; Manfred HENNEN (1994): *Egoismus und Altruismus in der Sozialtheorie*, 311ff.; Manfred HENNEN (1994): *Motivation als Konstrukt einer Sozialtheorie*, 153ff.

²³ Manfred HENNEN (2001): *Hat Wissenschaft Macht?* , 164f.

²⁴ Sebastian BAMBERG, Harald GUMBL, Peter SCHMIDT (2000): *Rational Choice und theoriegeleitete Evaluationsforschung*, 44ff.; Peter H. ROSSI, Howard E. FREEMAN, Mark W. LIPSEY (1999): *Evaluation. A Systematic Approach*, 102ff.

²⁵ Peter F. DRUCKER(2000): *Die Kunst des Managements*, 208.

3. Festlegungen und Grenzen

3.1 Evaluationsbegriff

Dem zuvor Gesagten soll jetzt eine exaktere *Definition* von Evaluation folgen.

Evaluation besteht in der Analyse und Bewertung von Einrichtungen, Programmen, Innovationen unter dem Aspekt ihrer Wirksamkeit, Akzeptanz und Qualität mit dem Ziel der Situations- und Problemdiagnose sowie der Mobilisierung von Aktivität. Evaluation erfolgt unter der Zugrundelegung von Maßstäben, die aus dem Evaluationsfeld oder von außen stammen können.

In dieser Definition sind die Schlüsselbegriffe *Qualität*, *Mobilisierung* und *Maßstab* enthalten. Es erscheint zunächst sinnvoll, sich dem *Qualitätsbegriff* zuzuwenden, weil sich mit dessen Konzeption viele wichtige Entscheidungen über die Art der Evaluation und die daraus sich ergebenden Chancen für die Zukunft ableiten.

3.2 Qualitätsbegriff

Qualitätskriterien lassen sich auf empirische Weise und durch Operationalisierungen gewinnen, aber sie sind ihrem Charakter nach zunächst nicht quantitativer, sondern qualitativer Natur. Sie können natürlich durch Zahlen oder Indizes eingelöst oder ergänzt werden. Die von uns entwickelten Qualitätskriterien sind eher als *Hypothesen* zu verstehen, als Beschreibungen von typischen und generalisierbaren Problemzonen in Organisationen und Kooperationen, aus deren Kritik sich positive Maßstäbe wie von selbst ergeben. Als hilfreich bieten sich beispielhaft gelungene Leistungen und Standards für Qualitätsvorstellungen des Organisationshandelns in *Vergleichseinrichtungen* an. Hier eine Definition von *Qualität* in Anlehnung an unsere Erfahrungen:

Qualität soll für unsere Zwecke heißen: Ein zugeschriebener skaliertes leistungsbezogener Wert von Gegenständen, Einrichtungen, Ereignissen, Eigenschaften und Prozessen unter Zugrundelegung eines Maßstabes. Qualität besteht in Übereinstimmung bzw. Abweichung zwischen definierten Zielen und einer messbaren oder beschreibbaren Wirklichkeit. Qualität ist ein Maß der Kompatibilität von Zielen und Teilzielen mit ihrer Verwirklichung.

Die Parallelen zur Evaluationsdefinition sind augenfällig und nicht zufällig. Kriterien von Qualität lassen sich in kontrollierte Beobachtungen, in systematische Fragenkomplexe umsetzen aber weitaus seltener in konzentrierte Quantifizierungen. Im Rahmen der Hochschulevaluation sind bibliographische Methoden der Zitation, Statistiken über Drittmittel, Examens- und Doktorarbeiten hilfreich aber nicht ohne zusätzliche Einschätzungen verwendbar. Selbsteinschätzungen in Ergänzung zu den quantitativen Beschreibungen, die wir schriftlich dokumentieren, bilden eine wesentliche Voraussetzung für die Bewertung, weil sie im Rahmen einer externen Begutachtung durch angesehene Fachkollegen eine höchst kompetente Kontrolle erfahren, welche die Selbsturteile durch entsprechende Antizipationen der Beurteilung der externen Gutachter wahrhaftig macht. Wichtig ist auch, dass Ziele selbst kollidieren können. Mittelfristige Ziele können erfüllt sein während längerfristige verletzt werden.

Der Maßstab kann auf verschiedene Weise gewonnen werden. Immer handelt es sich, wie oben formuliert, um *Vergleiche*. Es kann gesicherte Belege geben wie schnellere Entscheidungen, mehr Publikationen. Das erscheint noch als unproblematisch. Aber ist die Gesamtqualität damit beschrieben? Wenn Einspruch erhoben wird, müssen weitere Maßstäbe hinzugezogen werden. Man kann nach der Einschätzung von Fachkollegen fragen oder die Zitation prüfen. Weitere Einreden kann es geben, Einschätzungen von anderen Qualitäten sind möglich. Mir kommt es darauf an, dass das *Ringens um Maßstäbe* im äußersten Fall selbst zu einem kommunikativen Verfahren werden kann, das selbst bei divergierenden Urteilen Konsens- oder relative Verständigungsbereitschaft hervorbringt. Das haben wir sehr oft erlebt. Die völlige Übereinstimmung ist gar nicht nötig, aber die aus Distanz, Anonymität und Nichtbefassung geborene Fremd- und Selbsttäuschung wird unmöglich gemacht.²⁶

Trotz dieser unleugbaren Grenzen umfassender oder gar quantitativ objektiverer Verfahren und Erhebungen setzen viele Hochschulleitungen, Verantwortliche in der Wissenschaftsbürokratie und diejenigen, die mit der Erstellung von Rankings beauftragt sind, auf den Zahlenvergleich. Sie tun das über den Unterschied zwischen Äpfeln und Birnen hinweg, und es geschieht nicht selten wider besseres Wissen. Die gleichen Personen, die sich über populäre Rankings und deren Methodik mokieren, gefallen sich darin, ihr gutes Abschneiden in den kritisierten Verfahren

²⁶ Zu ‚Vergleichsaussagen‘: Helmut KROMREY (2000): *Qualität und Evaluation im System Hochschule*, 234, 233 – 258.

augenzwinkernd und genüsslich hervorzuheben. Die sachlichen Maßstäbe sind längst durch die medialen korrumpiert. Die wichtigeren Orientierungsmaßstäbe und Leitlinien sind nach meiner Einschätzung abhanden gekommen und einem *populären Qualitätsverständnis* gewichen, das auch dort quantifizierenden Maßstäben gehorcht, wo sie sachlich nicht gedeckt sind. Es scheint so zu sein, dass populistische Reputation dominiert. Vergleichsmaßstäbe sind eher noch schwieriger geworden. *Maßstabsarbeit* ist hier eine zentrale Aufgabe.

Ist die Mathematik an einer Hochschule besser als deren Volkskunde? Sind die 1000 Studierenden eines Modefaches höher zu bewerten als die Studiengemeinschaft von 12 ernsthaften Immatrikulierten eines Orchideenfaches? Soll eine Teildisziplin eines Faches aufgegeben werden, weil sie an anderen Universitäten ebenfalls geschlossen wird oder ist das ein Grund, sie gerade deshalb für alle Fälle lebendig zu halten? Sind die Forschungsleistungen eines Hochschullehrers doppelt oder halb so viel wert wie Präsenz und Betreuungsleistungen eines in der Forschung weniger hervorgetretenen Kollegen? Sind Drittmittelwerbungen höher anzuschlagen als die aufopferungsvolle Abfassung eines längst überfälligen Lehrbuches? Fragen über Fragen und wenig systematische Selbstvergewisserungen. Hier ist Evaluation ein Instrument der Rückkehr zu den Basics der Selbstbesinnung und des Vergleichens.

Wir haben im Rahmen unserer Evaluationen in vielen vertraulichen kollegialen Gesprächen nach Antworten auf Fragen dieser Art gesucht. Viele ungeschminkte Antworten sind uns hier gegeben worden. Das Ringen um die Art der angemessenen Fragestellung und erst recht die Bemühung um angemessene Antworten ist überall zu spüren. Ernsthaftigkeit dokumentiert sich in allen Gesprächen. Aber die Entschlossenheit des öffentlichen Fragens ist äußerst verhalten. Hochschulpolitik im Makro-, aber auch im Mikrobereich ist einer *grundsätzlichen Verdächtigung* z.B. der Professoren anheim gefallen. Aber auch sie sind voller *Resignation*. Die Dinge seien politisch gewollt, beschlossen, unabänderlich heißt es immer wieder. Hochschulangehörige, auch Studierende sind nicht selten resignativ und kleinlaut geworden. Das gibt zu denken. Nicht zuletzt Naturwissenschaftler, denen der Umgang mit Zahlen und Rechenmodellen geläufig ist, haben uns geholfen, sich dem Ansinnen zu entziehen, unbegründete Zahlenwerke wären besser als angemessene Überlegungen, die noch nicht quantifiziert werden können.

Maßstabssuche führt wie beschrieben zum Vergleich, auch zum *Vergleich der Maßstäbe*. Es kann auch die offene Fragestellung nach unterschiedli-

chen institutspolitischen oder fachpolitischen Wegen auf unterschiedlichen Maßstäben beruhen, andere Universitäten und deren Erfolg, andere Länder können als Standard herangezogen werden. Handlungsalternativen können scheinbar unversöhnlich nebeneinander stehen. Aber darin sind – so meine Erfahrung – Vorteile zu sehen: ein unerhörter Gewinn an Klarheit, eine Chance für Kompromisse und vor allen Dingen komprimierte Fragestellungen an externe Gutachter oder andere Berater. Ich rate stets, den schweren Weg bis zur Formulierung von Alternativen zu Ende zu gehen, weil diese Klarheit, die sich dadurch einstellt, einen Gewinn gegenüber polemischer Abwehr in der Form dialogisierender Einlassung bedeutet.

3.3 Grenzen der Standardisierung

Die Institutionalisierung der Evaluation tut sich, wie beschrieben, mit dem häufig erhobenen Anspruch nach umfassender Quantifizierung schwer. Das gilt aber auch mit einem vergleichbaren Anspruch für *Standardisierung von Verfahren*. Wie kann es angehen, dass es an einer hinreichend komplexen Diagnose der Hochschulsituation über weite Strecken fehlt und gleichzeitig auf eine Art der *Technisierung der Hochschulsteuerung* gesetzt wird? Mehr noch: die Evaluation selbst erliegt der trügerischen Hoffnung, eine Art Standardverfahren werden zu können, mit dem überall und unmittelbar brauchbare Ergebnisse zu erzielen wären.

Die intensive Nachfrage nach Evaluation bei gleichzeitiger Unsicherheit hinsichtlich der Bestimmung ihrer Aufgabe bilden hier eine ungünstige Allianz. Die Evaluation ist als Verfahren nur halbherzig etabliert. Es ist oft hastig implementiert worden um einem Bedürfnis des ‚Wir-auch‘ nachzukommen. Helmut Kromrey spricht in diesem Zusammenhang von ‚Evaluationsritualen‘.²⁷ Die Evaluation kommt in diesen Fällen über eine *legitimierende Funktion* nicht hinaus. Damit haben wir auch zu kämpfen gehabt und haben dort weniger gute Ergebnisse erzielt, wo man uns das hinter vorgehaltener Hand gesagt hatte ohne dass wir uns deutlich distanziert hätten. Aber wir haben das Verfahren geändert und das wird akzeptiert. Unter diesen veränderten Umständen ist die Akzeptanz wesentlich gesteigert worden. Uns ist auch schon begegnet, dass ein Evaluationsver-

²⁷ Helmut Kromrey (2000): *Qualität und Evaluation im System Hochschule*, 238.

fahren einstimmig beschlossen, sein Beginn aber faktisch behindert worden ist. Das ist in sich bereits ein Symptom für die Notwendigkeit der Reorganisation.

Standardfragebogen mit der Erhebung von eher allgemeinen Einschätzungen zur Lehre oder gar nur zur Präsentation von Lehrinhalten oder die fragebogengestützte Selbsteinschätzung werden in Zukunft nicht allein die Grundlage der Evaluation liefern können, selbst dann nicht, wenn die Gegenstände der Erhebung erfolgreich erweitert werden sollten. Diese zunächst valide wirkenden Erhebungen verpuffen wegen wiederholter methodischer Einwände, wegen einer weitgehend folgenlosen Zur-Kennntnisnahme der Ergebnisse, vor allen Dingen aber wegen der Unmöglichkeit, das oben beschriebene Hauptproblem der Steuerungs- und Verantwortungsdefizite angemessen zur Sprache und zum Bewusstsein zu bringen. Mit der nach meiner Einschätzung nur vorläufigen und provisorischen Standardisierung solcher Erhebungsverfahren wird keine Überzeugungskraft für wirklich notwendige Veränderungen gewonnen. Was als methodisch gesicherte Diagnose erscheint, macht dem Namen Evaluation wenig Ehre. Das Verfahren wird konsumiert ohne wirklich akzeptiert zu sein. Das aber kann für eine noch ungeklärte Innovation kein Markenzeichen sein. Wir haben es allenfalls mit einer ‚Objektivierung durch Verfahren‘ zu tun.²⁸

In einem anspruchsvollen Feld wie es die Hochschulen darstellen, aber auch in vergleichbaren Organisationen mit großer Kompetenz der Fachspezialisten, sind Standarderhebungen formalisierter Natur vor allen Dingen als Mittel der Mobilisierung wenig überzeugend. Sie fordern eher Widerspruch oder die Übergehung heraus und tragen sehr wenig zu dem bei, was wir als zentrale Aufgabe der Evaluation ansehen, nämlich dort tätig zu werden, wo die *Initiativ- und Verantwortungsbrache* herrscht. Gerade die sehr erfolgreichen Evaluationen haben für unsere Arbeit in allen Fällen auch eine Art *freundlichen Kampfspiels* um die Ernsthaftigkeit des Engagements und der Selbstdefinition derjenigen, die sich der Evaluation unterziehen, mit sich gebracht. In einem Feld des teilweisen *Steuerungsvakuum*s, das völlig legitim erscheinen muss, kann kein empirisches Verfahren die Arbeit der Überzeugung und Mobilisierung allein übernehmen.

²⁸ Helmut Kromrey (2000): *Qualität und Evaluation im System Hochschule*, 246ff.

Erhebungstechniken sind für sich genommen noch keine Evaluation. Damit sind sie völlig überfordert und dafür sind sie nicht erfunden worden. Es sei daran erinnert, dass Datenerhebung selbst im Geschäft der Wissenschaft für die Erkenntnisgewinnung ohne die Formulierung von Fragestellungen und ohne die Vermutung über Zusammenhänge in Form von Theorien kraftlos ist. In einer Organisation geht es um *zielgeführte Kooperationen*, um Verantwortungen und damit verbunden um *Entscheidungen*.²⁹ Für diese Entscheidungen wird eine Fragestellung benötigt. Wenn diese Fragen sicher platziert sein sollen, braucht man dazu Entschlossenheit und Mut zur Kommunikation. Um dies herbeizuführen wird etwas anderes benötigt als die Mechanik eines All-round-Verfahrens zur Produktion von Einblick.

Man stelle sich vor, ein Wissenschaftler erhält einen ersten Ruf, findet an der neuen Universität Kollegen, Klima und Verfahrenspraktiken des überindividuellen Wissenschaftsvollzuges vor, die ihm wenig Anreiz und Herausforderung bieten. Die eigene wissenschaftliche Arbeit und die eigenen Darstellungen in der Lehre besetzen das Leben der ersten Semester mit hoher Arbeitsintensität. Ein Engagement gegen bestehende Strukturen kosten noch mehr Kraft. Woher sollte eine Entschlossenheit zur Umgestaltung des Institutswesens kommen? Woher die Kraft, wenn es keinerlei zur Autorität in diesen Fragen berufene Ansprechpersonen gibt, sondern nur Signale, das Bestehende nicht durch Übereifer gefährden zu sollen. Das aber ist eine Wirklichkeit in Organisationen aller Art, nicht nur in Hochschulen. In dieser Situation helfen standardisierte Verfahren der empirischen Erhebung so gut wie nichts, sie machen sich verdächtig, werden u.U. ausgenutzt und sogar ironisiert. Gerade *frisch berufene Professoren* haben durch ihre Nachfrage Evaluationen eingeleitet. Unsere Aufgabe war es, deren Anliegen durch die Darstellung vor Fachbereichsräten oder Institutsvertretungen sowie den Evaluationsprozess auch gegenüber denjenigen zu vertreten, die weniger oder überhaupt nicht von dieser Notwendigkeit überzeugt waren.

3.4 Partnerschaft

Wenn die Ernsthaftigkeit für Reformen und Innovationen hergestellt werden soll, ist es von entscheidender Bedeutung, dass Chancen autonomer Verfahrensreflexion miteingebaut werden können. Das aber ist nicht

²⁹ Vgl.: Niklas LUHMANN (2000): *Organisation und Entscheidung*, 71, 123, 261.

immer der Fall. Das Ringen um Darstellung der Lage und um Zuspitzung der offenen Fragen und ungelösten Probleme erfordert *Partnerschaft* als Voraussetzung für Akzeptanz und Respekt. Partnerschaft zumindest bei einigen zentralen Prozessbeteiligten ist unverzichtbar. Viele Evaluations-einrichtungen im Bereich der Hochschule verfügen nur über einen äußerst geringen Personalbestand, der entsprechende Interaktionen herstellen könnte. Zudem können diese Stellen durch die Art der Anbindung wenig Selbständigkeit garantieren. Einsame BAT-IIa-Stellen sind, nicht selten sogar als Halbtagsstellen, mit Akademischen Mitarbeitern und Mitarbeiterinnen besetzt und Präsidenten, Vizepräsidenten, Rektoren oder Kanzlern unterstellt oder einer nur formal aber nicht sachlich zuständigen Professur.

Diese Abhängigkeit oder zumindest *relative Unselbständigkeit* kann nur dann die Unangreifbarkeit sichern, wenn die Person sich der Leitungsfunktion unterwirft oder sich in die besprochene Methodenformalisierung flüchtet. Beides kommt einer selbstbewussten und ernstgenommenen Evaluationsinstanz, die vor allen Dingen Neutralität garantieren kann, wenig nahe. Man denke nur an die starke Stellung von *Studiendekanen* im bayrischen Hochschulgesetz, um mit diesem Vergleich Aufschluss über die notwendige Unabhängigkeit zu gewinnen. Unselbständigkeit und fehlende Erfahrung sind denkbar schlechte Voraussetzungen, um Bewertungsmaßstäbe entwickeln und in Verhandlungen über die Offenlegung von Perspektiven eintreten zu können. Eine *immunisierende Mechanisierung des Verfahrens* hat geringe Chancen für die notwendige Überzeugungsarbeit und erst recht für notwendige Interventionen.³⁰

An der Universität Mainz ist ein anderer Weg beschritten worden. Es wurde nämlich eine selbständige und wissenschaftliche Einrichtung geschaffen, um eine Garantie für *Handlungsfreiheit und Neutralität* zu bieten und ein personelles Volumen zu schaffen, das den Ausbau für die Akquisition externer Projekte ermöglicht. Die für Evaluationen erforderliche Unabhängigkeit wächst natürlich mit einer solchen Ausdehnung der Erfahrungen in unterschiedlichen Feldern.

Das zeigt sich in mehrerer Hinsicht und erscheint uns für die Zukunft der Evaluation richtungsweisend. Wir selbst sind im Laufe unserer Evaluationserfahrungen nach und nach dazu übergegangen, bereits zu Beginn die Verantwortlichen, die sich Evaluationen unterziehen wollten, danach zu

³⁰ Helmut KROMREY (2000): *Qualität und Evaluation im System Hochschule*, 254f.

fragen, was sie denn wissen wollten, welche Ziele sie mit der Evaluation verbinden, welche ungelösten Problemstellungen sie zu behandeln gedächten und welche Vermutungen über eventuelle Defizite sie bewegten. Diese Vorgehensweise ist äußerst aufschlussreich. Es zeigt sich nämlich sehr schnell, ob die innere Bereitschaft, wirklich zentrale Fragen zu stellen, vorhanden ist oder ob stattdessen eher ein Bedürfnis vorliegt, sich einem unspezifischen Prozess der Produktion von unbedrohlichen Daten zu unterziehen. Es ist ein offenes Geheimnis, dass Evaluationen in einer weichen Form z.B. der Meinungsmessungen durch Hörerbefragungen mit der Bildung von zusammenfassenden Mittelwerten als Legitimation angesehen werden nach dem Motto: Wir haben uns dem Verfahren unterzogen, sind gar nicht so schlecht weggekommen und nehmen dieses Ergebnis als ein Gütesiegel. Erst wenn Evaluationsinstanzen sich einer solcher Vorgehensweise verweigern, erfahren sie, ob ihnen ein ehrliches Interesse entgegengebracht wird, mehr als einen flüchtigen Blick in die Küche werfen zu lassen.

Zu erfahren ist auf diese Weise auch, ob sie in einer Rolle akzeptiert werden, der Kompetenz zugeschrieben wird. Damit ist der Übergang zu einem Geschäft beschrieben, das *Beratung* genannt werden kann. Neuerdings haben wir in geheimen Abstimmungen über den Wunsch nach Evaluationen nach Voranfrage und nach unserer Selbstpräsentation und Befragung über heikle Themen des Verfahrens abstimmen lassen, auch das hat sich bewährt, besonders wenn die Zustimmung hoch ist. Eine solchermaßen auf Ernsthaftigkeit und Entschiedenheit verpflichtete Ausgangslage erlaubt Partnerschaft und schafft zudem die Grundlage für entschiedene Interventionen. Damit verändert sich aber die Beziehung zu den Auftraggebern als Organisationen entscheidend, wir bewegen uns im Feld der Beratung.

3.5 Aufgabe und Grenzen der Experten

Wenn die Zukunft der Evaluation und ihre Professionalisierung eher im Stil der *Beratung* gesehen werden muss, so hat das auch Konsequenzen für die Stellung von Experten, die in der Hochschulevaluation unter dem Namen ‚Peers‘ bekannt sind. Um in der Lage zu sein, ein Verfahren der Evaluation partnerschaftlich und beratend durchzuführen und angemessene Maßstäbe der Beurteilung und des Vergleichs entwickeln zu können bedarf es der Nähe zu den evaluierten Einrichtungen ebenso wie der Kontinuität der Ansprache und Ansprechbarkeit. In unserem Zentrum haben

wir das erst durch die an uns gerichteten Erwartungen an eine intensive und über das eigentliche Evaluationsverfahren i.e.S. hinausgehende Betreuung lernen müssen. Das gibt Anlass über das dreigliedrige Verfahren der Hochschulevaluationen nachzudenken.³¹ Es besteht aus der internen, der externen Evaluation und der Zielvereinbarung. Das interne Verfahren pflegt auch als Selbstbericht oder Selbstreport bezeichnet zu werden. Hier möchte ich auf das Verhältnis von interner und externer Evaluation näher eingehen.

M.E. ist mit einer stärkeren Akzentuierung der Beratung auch die etablierte *Beziehung von interner und externer Evaluation* in Frage gestellt. Viele Dinge sind hier ungeklärt. Was ist eine interne Evaluation, was ein interner Bericht? Die meisten Hochschulevaluationen in Deutschland und in den ausländischen Vorbild-Ländern stützen sich auf *Selbstevaluationen*, die auf Beschreibungen und Bewertungen anhand von Fragebögen nicht nur durch Studierende, sondern auch durch Mitglieder der Statusgruppen der zu evaluierenden Fächer erfolgen. Neben den Grunddaten geht es hierbei um Einschätzungen von Stärken und Schwächen der Fächer in verschiedenen Feldern. Hiermit ist der Anspruch der objektiven Messbarkeit weitgehend zugunsten einer Selbstbeschreibung aufgegeben. Objektivität wird der externen Evaluation zugeschrieben, in der Fachgutachter auf der Basis von Begehungen, Befragungen und Beobachtungen Urteile treffen und diese in einem Bericht wiedergeben. Meistens sind dafür weniger als zwei vollständige Besuchstage vorgesehen.

Der Blick von externen Gutachtern kann aber nach unserer Auffassung und Erfahrung nur so unbestechlich und ihr Urteil treffsicher sein, wie die Grundlagen vorausgegangener Problembeschreibungen und Fragestellungen es hergeben. Wissenschaftliche Gutachter sind Spezialisten ihres Faches kennen sich normalerweise in Organisationsfragen ihrer Disziplin, oft auch in Fachgutachten aus, aber sie sind keine Spezialisten der Erhebung, der Beobachtung und der Befragung. Es wäre auch unsinnig, das von ihnen zu erwarten. Ihre Wahrnehmungen und ihre Urteile können kaum aufmerksamer, schärfer und präziser sein als die zuvor an sie gerichteten Fragestellungen. Es sind ihre Urteile, die sie so wertvoll machen, aber diese hängen von präzisierten Fragen ab. Wer Gutachter ohne sehr genaue *Fixierung der anstehenden offenen Fragen oder Beschreibungen* ins Haus holt mit der diffusen Bitte, sich einmal ein Bild von der Situation einer Einrichtung und ihrer Programme zu machen, gibt das

³¹ Vgl. den Beitrag von Elisabeth SPRINGER in diesem Band.

Anliegen aus der Hand. Das dürfen Professionelle der Evaluation nicht zulassen.

Stattdessen muss im Bericht bereits eine sehr weit und tief geführte *Vordiagnose* erkennbar sein, ebenso eine bis dahin erbrachte tiefgehende Artikulation der Organisationsverantwortlichen. Die Vorlagen für die externen Experten sollten den Status von wissenschaftlichen Hypothesen haben, um die Sachkenntnis der Experten und ihr Urteilsvermögen in besonders wirksamer Weise herauszufordern. Zugespißt formuliert könnte man sagen, die Experten seien nur so gut wie die Fragen, welche an sie gerichtet werden. Auch externe Experten arbeiten mit der Zuschreibung von *Qualität durch Vergleich*. Dabei werden Qualitätsdefinitionen offen gehalten ohne der Versuchung zu erliegen, Qualität eindimensional festzuschreiben zu wollen. Qualität ist vielschichtig und mehrwertig. Der Reichtum in der Qualitätsbegutachtung durch Externe liegt in der Vorbereitung durch die Evaluatoren.

Wir selbst haben durch unsere Erfahrungen auf dem Weg von anfänglichen eher weichen Beschreibungen mit dem Bedürfnis, niemanden zu verprellen bis zu sehr ausführlichen Diskussionen um die Kernprobleme gemacht. Der deutliche Bericht lohnt die Streitgespräche und Auseinandersetzungen. Rational geführte, sachlich begründete Auseinandersetzungen können durchaus verbinden. Die Externen haben die daraus hervorgehende Deutlichkeit der Analyse stets für ihre Arbeit gewürdigt und die Kollegen vor Ort dafür gelobt. Die Zukunft des Evaluationsverfahrens wird nicht von der Besessenheit ausschließlicher Quantifizierung abhängen, sondern von der Sicherung von Vielfalt und Differenziertheit des Vergleichens sowie einer Methode des kontrollierten Selbstbeurteilens und Fremdbeurteilens. Diese Voraussetzung lässt Beratungsqualitäten zentral werden.

4. Evaluation als Beratung

4.1 Intermediäre Instanzen

Wenn die bisherigen Evaluationen in den deutschen Hochschulen eher unspezifisch entworfen worden sind, dann hat das sicherlich auch politische Gründe. Der Begriff Evaluation taucht zunächst einmal als ein hoheitliches Instrument des Vollzuges von feststehenden Entscheidungen hinsichtlich der Schließung von Institutionen auf. Dies ist unter dem Beg-

riff der sogenannten *Abwicklung* der früheren DDR-Forschungseinrichtungen bekannt geworden. Bis heute gibt es vergleichende Bewertungen von wissenschaftlichen und Bildungseinrichtungen mit der Maßgabe der Schließung eines Teils von ihnen. Natürlich bringt das viel Unruhe bei den potentiell Betroffenen mit sich. Ein Evaluationsverfahren, das über diesen Verdacht des totalen Eingriffs erhaben ist, erscheint als Abwendung einer Drohung. Ein bissiger Hund bleibt an der Kette. Die Zukunft der Evaluation liegt nach meiner Einschätzung zwischen der Gestalt dieses gefährlichen Kettenhundes und einem nur kläffenden Schoßhündchen, das des Pläsiers wegen gehalten wird.

Das ist aber eine Gefahr; denn Evaluation wird zu *aufwendig und teuer* sein, wenn sich herausstellen sollte, dass sie dort wenig bewirkt, wo das Organisations- und Leistungsgefüge Spuren der Diffusion von Kompetenz und Verantwortung zeigt. Aus den beschriebenen Gründen wird es nicht möglich sein, subtile Leistungsprozesse, die von Motivationen und Verantwortung abhängig sind, hierarchisch zu steuern. Insofern scheiden rein hoheitliche Steuerungsvorhaben aus. Und doch müssen diese Steuerungsresistenzen überwunden werden. Kommt das einer Quadratur des Kreises gleich? Ich glaube das nicht. Das Feld der zukünftigen Evaluation heißt Beratung.

Beratung ist zunächst mehr als Diagnose. Sie verspricht Therapie, die Diagnose voraussetzt und zu erneuter Diagnose führt. Beide Elemente vermögen erst in ihrer Verbindung zu überzeugen.

Beratung besteht in der zweiseitigen Überzeugung, die Nachfrage nach Beurteilungswissen könne Hilfen und Gewinne schaffen. Sie beruht auf der Bereitschaft, sich auf neue Perspektiven und die Mobilisierung eigener Kräfte einzulassen. Das ist der teilweise Verzicht auf eigene Entscheidungsgewalt, was diese Vorentscheidungen betrifft bei gleichzeitigem Fortbestand grundsätzlicher Entscheidungsfreiheit.

Das klingt vielleicht kompliziert, wird aber sofort verständlich, wenn man an ein *Arzt-Patient-Verhältnis* denkt als dem Paradigma der Beratung. Die Zuständigkeiten für die ureigenen Angelegenheiten von Leben, Körper und Geist sind selbstverständlich unangetastet und gleichzeitig werden die Entscheidungen darüber aufgeschoben und bis zu einem Kompetenzzuwachs ausgesetzt. Beratung benötigt besondere Kompetenzen, über die in einer Vertrauensbeziehung zur Erweiterung der eigenen Entscheidungsvoraussetzungen vom Beratenen verfügt werden kann.

Wenn sich in unseren Evaluationsverfahren solche Beratungssituationen ergeben haben, handelte es sich um sehr intensive, dichte und ernsthafte Momente des Austausches, um Formen des Ringens um eine neue Perspektive, die eine Lösung erlaubte, der Nachfrage und Rückfrage, der Abwägung und wechselseitige Bemühung um angemessene Situationsdefinitionen. Auch wenn es bis dahin Reibung, Skepsis, manchmal Streitgespräche gegeben hatte, die intensivsten Momente der Beratung bestanden in Formen der *Transaktion*, der Übertragung von konfliktreichen Anschauungen.

Ich erinnere mich beispielsweise an offene Fragen der Umstrukturierung eines Studiengangs und es wurde klar, dass davon buchstäblich das Schicksal der beteiligten Fächer abhing, weil sie sich in heftiger Konkurrenz befanden. Nachdem völlig offen lag, worum es ging, waren die *Gespräche mit wechselnden Gruppen und Personen* von einer angespannten Nähe und voller Vertrauen darauf, dass jetzt nur mit intensiver Ernsthaftigkeit auf allen Seiten weiter zu kommen war. Die Erzeugung angemessener Vielfalt von Entscheidungsperspektiven unter Einschluss von Qualitätsvorstellungen aller Art ist vielleicht das Geheimnis der Beratung. Beratung ist die glaubhafte Herstellung der Transparenz von Handlungsfolgen, weil sich dahinter Skepsis, Angst und Hoffnungen gleichermaßen verbergen.

Evaluationen solchermaßen in Nähe der *Organisationsberatung* gerückt, können – so wird jetzt wohl deutlich – nur zum Teil auf Routineverfahren beruhen. Evaluationen haben auch immer etwas Außeralltägliches an sich, das sollen sie auch und darin ist eine Art Markenzeichen des Gelingens zu sehen. Der Alltag mit seinen eingefahrenen Routinen bringt ja gerade die schleichenden Qualitätsverluste mit sich. Die *unausgestandenen Entscheidungen* über einen neuen Studiengang oder gemeinsame Forschungsschwerpunkte können die Blockaden ausmachen. Kann es eine besondere Evaluationskompetenz für die Diagnostik der alltäglichen Abläufe in außeralltäglichen Situationen geben? Hier nur soviel: es handelt sich um eine Mischung aus Kenntnis vergleichbarer Prozesse in Organisationen und Institutionen, Hintergrundwissen über Kernprozesse wie Forschen und Lehren innerhalb der Universität und sozialwissenschaftliche Kenntnisse über das, was hier wiederholt Transaktionen genannt worden ist. Ich werde später auf die Kompetenzen im Detail eingehen. Alles andere ist als Kunst des Zuhörens, Fragens, Verstehens und Argumentierens wenig lernbar und bleibt deshalb immer eine sehr vorläufige und vage Kompetenz.

Evaluation als eine Form der Organisationsberatung wird damit in die Nähe anderer Begriffe von Beratung gerückt. Der Beratungsbegriff hat Konjunktur mit vielen Synonymen. Wir hören und lesen von Moderatoren, Mediatoren, Schlichtern, Spezialberatern aller Art, Vermittlern, Kuratoren, Treuhändern, Gutachern und Ombudsleuten, die sogar im Rechtssystem mit seinen Präferenzen für formalisierte Prozesse ihren festen Platz gefunden haben.³² All diesen Funktionen ist gemeinsam, dass sie außenstehende, primär unbeteiligte Instanzen ohne Eigeninteresse in der Sache, d.h. am Beratungsgegenstand darstellen und eine Vermittlungsaufgabe wahrnehmen. Ich verwende für diese Einrichtungen den Sammelbegriff ‚*Intermediäre Instanzen*‘.

4.2 Selbständigkeit und Weiterungen

Intermediäre Instanzen verpflichten sich einer *treuhänderischen Aufgabe* in dem Sinne, dass sie vorübergehend und in definierter Weise fremde Interessen zu ihren eigenen machen. Dieser professionelle Altruismus hat natürlich Stützen im Berufsbild. Der Erfolg ist der Erfolg für Andere und zwischen Anderen. Beratung ist in allen Fällen der Spezialisierung zu einer ergänzenden und unterstützenden Steuerung geworden. M.E. ist das ein weiterer Beleg für ein internes Steuerungsdefizit besonders in wissensbasierten und Wissen verarbeitenden Einrichtungen, die abhängig sind von *Selbststeuerungsprozessen*. In immer mehr Bereichen zur Bereitstellung anspruchsvoller Leistungen finden wir die Konstellation von vielfältigen Interaktionen und Interdependenzen vor, die erhebliche Freiräume, aber eben auch Möglichkeiten der Verweigerung mit sich bringen.

In vielen Praxisfeldern des sozialen Lebens wird deshalb *Motivation* zum Thema gemacht. Motivation wird dabei nicht als Eingriff, sondern als Selbstbindung und Selbstverpflichtung verstanden. „Motivation gilt als positive Grundeinstellung gegenüber Handlungsanforderungen.“³³ *Motivation und Beratung sind als komplementäre Konstrukte* zu deuten. Beratung kann sich nur an Motivationen wenden und muss einen solchen

³² Otto Ernst KRASNEY (2000): *Ein Ombudsmann für die Opfer von Massenunglücken?* 575 - 588.

³³ Manfred HENNEN (1994): *Egoismus und Altruismus in der Sozialtheorie*, 322.

Begriff voraussetzen. Selbstinteresse muss mit sozialen Notwendigkeiten verschmelzen können.³⁴

Das macht verständlich, dass die Gewinnung von Vertrauen eine der wichtigsten Voraussetzungen für eine gelingende Evaluation im Sinne einer Beratung darstellt. Sie ist professioneller Ausdruck des zeitlich begrenzten Sonderverhältnisses, eine Art neutraler Intimität oder *intimer Neutralität*. Dabei ist es oft das schwerste Geschäft, das Vertrauen unterschiedlicher, oft antagonistischer Personenkreise und Instanzen zu gewinnen. Sie alle verfügen über Wissen und unterschiedliche Entscheidungsbefugnisse. Entscheidungen sollten keineswegs nur bei Vorgesetzten unterschiedlicher Ebenen gesucht werden, auch nachgeordnete Entscheidungen können von großem Interesse sein. Vertrauen bei all diesen Personen zu gewinnen erfordert eine komplexe und überzeugende Darstellung von Interesse und Desinteresse. Ersteres bezieht sich auf die Entschlossenheit, Zusammenhänge, welche die Perspektive der Einzelnen übersteigen, offen zu legen, Letzteres entsteht aus dem Verzicht der Parteilung. Die Qualität der Evaluation soll neuerdings auch hierzulande durch ‚Standards‘ gesichert werden. Zu ihnen gehören die Klassen ‚Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit‘, die weiter spezifiziert werden.³⁵

Das offenbart aber eine oft übersehene Notwendigkeit, die an das zuvor Gesagte anknüpft. Der Rolle der Vorgesetzten gegenüber darf die beratende Person nicht grundsätzlich in ihrer Kompetenz unterlegen sein. Man muss nicht Chefarzt sein um Chefärzte zu beraten, aber man darf auch nicht zur Klasse der Assistenzärzte gehören. Genauer gesagt, sollte die beratende Instanz nicht ausschließlich aus Abhängigen bestehen. Es sei an den Fall der Stabsstelle beim Universitätsrektor erinnert. Wer gegenüber den Vorgesetzten im Beratungsfeld als unabhängig gelten kann, wird dies auch gegenüber den in der Hierarchie tiefer Angesiedelten leichter tun können. Für die beratende Einrichtung hat die eigene Hierarchie u.U. eine wichtige Schutzfunktion. Wir haben immer wieder erfahren, dass wissenschaftliche Mitarbeiterinnen und Mitarbeiter trotz ihrer Jugend auch von Leitungspersonen respektiert werden, wenn sie als Teammitglieder eingeführt worden waren. Die Beobachtung hat uns oft amüsiert, dass dieselben Personen weniger Akzeptanz fanden, wenn sie das erste Mal allein auftraten.

³⁴ Manfred HENNEN (1994): *Motivation als Konstrukt einer Sozialtheorie*, 171.

³⁵ Thomas WIDMER (2000): *Qualität der Evaluation*, 77 – 102.

Der Charakter der Evaluation als Beratung hat noch eine andere Seite. Die Evaluationskosten sind, wie gesagt, nicht unbeträchtlich und deshalb wird es in Zukunft wahrscheinlich wesentlich weniger flächendeckende Über-alles-Evaluationen in ganzen Einrichtungen im Stil einer empirischen Gesamterhebung geben. Was dagegen mit hoher Wahrscheinlichkeit zunehmen wird, ist ein Bewusstsein für Interventionsbedarf in *Brennpunkten* von Organisationen und Einrichtungen. In unserem Erfahrungsbereich deuten Nachfragen nach enger definierten *Problem- und Interventionsfeldern* darauf hin, dass es spezifische Beratungsaufträge geben wird, die Zukunft haben. Es handelt sich besonders um solche Handlungsfelder und Handlungsparteien, deren Funktionsdefizite und Koordinationsbedarf seit langem bekannt sind, ohne dass Interventionsmöglichkeiten für Geschäftsführungen, Dekanate, Hochschulleitungen oder das zuständige Ministerium in befriedigender Weise zur Verfügung gestanden hätten. Nicht selten sind die Parteien so zerstritten und in ihren Beschuldigungen und Forderungen so unversöhnlich, dass die Reibungsverluste sehr weit reichen.

Es versteht sich leicht, dass eine Evaluationseinrichtung, die mit einer solchen Aufgabe betraut wird, sich nicht einfach als Beweger in einem weitgehend unverbindlichen Verfahren sehen kann, in dem sie allseits mit offenen Armen aufgenommen wird. Es geht um Beratungen in einem verminten Feld, wie das von Eingeweihten bezeichnet wird. Hier kommt es darauf an, die eigene *Unabhängigkeit zu bewahren* und das heißt auch immer, notfalls nein sagen zu können. Nein zu sagen kann gleichbedeutend sein mit dem Eingeständnis des eigenen Scheiterns. Das ist in der Beratung nicht anders als überall im Sozialleben. Aber es verweist in diesem Zusammenhang erneut auf die notwendige Form der Institutionalisierung und Eigenständigkeit der Evaluation dieser Art. Wer ohne Vorlauf mit einschlägiger breiter Erfahrung und ohne Reputation ist, kann sich solchen Aufgaben nicht stellen, weil die Erpressbarkeit zu groß ist. Es klingt paradox, aber es scheint so zu sein, dass man mit eher weichen Beratungsformen beginnen muss, um entsprechende Kriterien für energischere Verfahren aufbauen zu können. Das ist jedoch eher die Beschreibung eines sich einstellenden Effektes als die eines geplanten Vorgehens.

Zum Abschluss der Erörterungen über die Evaluation als Beratung muss noch etwas gesagt werden zu dem, was ich ‚*Ausstrahlung der Evaluation*‘ nennen möchte. Damit ist zunächst gemeint, dass die Erkenntnisse aus einer Evaluationsberatung, die auch zu Folgen führen müssen, im Falle der universitären Evaluation in Zielvereinbarungen mit der Hochschul-

leitung einmünden. Das war auch für uns zunächst selbstverständlich bis wir erkennen mussten, dass dies viel zu eindimensional gedacht ist. Es unterstellt, dass die Lösungen aller lokalisierten Probleme ausschließlich in der Hand der Fachbereiche bzw. Fächer und der Hochschulleitungen lägen. Das hat auch damit zu tun, dass von hier aus das Evaluationsbegehren ausgeht. Man stößt jedoch auf Probleme, deren Beseitigung in der Hand Dritter liegt. Das ist z.B. dann der Fall, wenn das zuständige Ministerium eine Umstrukturierung mittragen muss, weil sie Verordnungs- und Gesetzesimplikationen mit sich bringt.

In sehr vielen Fällen liegen die beschriebenen Blockaden gerade in einem *Geflecht von Zuständigkeiten*, das in der Vergangenheit die Praxis ermöglicht hat, die Zuschreibung von Zuständigkeiten und Verantwortungen sozusagen hin- und herzuschieben. Dies ist geradezu ein typischer Fall von komplexen Transaktionsblockaden. Zunächst denkt man immer nur an interne Verkrustungen. Die Lösungen werden dort schwieriger, wo die *Organisationsgrenzen überschritten* werden. Die Logik unseres Verfahrens legt nahe, diese Situation in Verhandlungen und Zielvereinbarungen widerzuspiegeln.

Das *Kontraktmanagement und Zielvereinbarungen* sind deshalb auch mit einem größer werdenden Kreis an Einrichtungen aufzunehmen. Unser Ministerium hat das in einem wichtigen Paradedfall akzeptiert und anerkannt, dass mit der Aufforderung zur Evaluation und mit der In-die-Pflicht-Nahme der Fachbereiche es selbst Änderungen mittragen muss, wenn es selbst mit im Boot sitzt. Ich glaube, das ist ein sehr wichtiger Schritt gewesen, der für die Evaluation zukunftsweisend ist. Aber auch das wäre nicht denkbar gewesen ohne eine Einrichtung, bei der die Möglichkeit der Ausarbeitung und Umsetzung solcher Erkenntnisse liegen kann. Beratung erstreckt sich hier wie in anderen Fällen auf eine Mehrzahl von Ansprechpartnern. Im Übrigen ist dem beschriebenen Fall zu entnehmen, dass das Prinzip der Unterscheidung zwischen veranlassenden Instanzen der Evaluation – beispielsweise das Ministerium – und veranlassenden Einrichtungen – beispielsweise ein Fachbereich – durchbrochen ist. Das ist eine gute Entwicklung für umfassende Transaktionsprozesse. Reformen sind keine Einbahnstraße des Bewirkens. In der Zukunft könnten vielseitige Evaluationsabmachungen solche Überraschungen der wechselseitigen Verpflichtungen mit sich bringen.

4.3 Sprache und Loyalität

Wir haben gesehen, dass Beratung auch innerhalb der Evaluation eine prinzipielle Gleichwertigkeit der Gesprächspartner voraussetzt. Darin besteht eine personelle und finanzielle Herausforderung an die Evaluation der Zukunft, auf die es noch keine ausreichenden Antworten gibt. Jedenfalls ist eine sehr *weitgehende Spezialisierung* erforderlich, die nur im Rahmen von Agenturen, starken Funktionsstellen oder vergleichbaren Einrichtungen gewährleistet werden kann, die über einen entsprechenden Personalbestand verfügen. Was aber hat es mit der Methodik der Evaluation und Beratung auf sich, wie ist ein Vertrauensverhältnis herstellbar und wie kann die intime Neutralität verwirklicht werden?

Wir müssen feststellen, dass es eine theoretische Fundierung für unsere praktischen Erfahrungen in einer zufriedenstellenden Weise noch nicht gibt. Zunächst zu unserer Praxis: in unseren leitfadengestützten Gesprächen mit Studierenden, akademischen Mitarbeitern und anderen Gruppen haben wir mit steigender Intensität die Erfahrung gemacht, dass mit dem thematisch bestimmten Gespräch in einer einerseits festgelegten und andererseits offenen Situation der Thematisierung von Zusammenhängen aller Art Platz für überraschende Begegnungen entsteht. Voraussetzung dafür ist bei den Interview-Partnern auf Seiten der Evaluatoren eine Feldkenntnis und entsprechende Sensibilisierung für zentrale Probleme des Feldes, welches es auch sei. Die Kenntnis von exemplarischen Hochschulabläufen mit ihren offenen Fragen ist z.B. unerlässlich. Wer selbst die Leistungsanforderungen kennt, die an diejenigen gestellt sind, die Vorlesungen halten, kann unbefangener und glaubhafter mit Kollegen und Studierenden über dieses heikle Thema sprechen.

Wir haben diese Art der Befragung in unserer Arbeit eine ‚*Ausstrahlung der Evaluation*‘ genannt. Darunter ist zu verstehen, dass Rückfragen und intensive Nachfragen nicht nur dazugehören, sondern auch Methode sind. Wir gehen davon aus, dass unsere Gesprächspartner über ein großes Wissen darüber verfügen, was sie tun und was mit ihnen geschieht. Wir begreifen sie als Akteure, welche nicht nur den Folgen der Intentionen und Handlungen anderer, sondern auch denen ihrer eigenen Handlungen begegnen. Auch Studierenden im ersten oder zweiten Semester unterstellen wir Wissen, Sprach- und Artikulationsfähigkeit sowie einen hohen Grad an Reflexionsvermögen. Entsprechend fragen wir und erlauben uns, gegebenenfalls ungläubig nachzufragen, immer mit dem Hintergrund, als außenstehende und fachfremde Insider der Bildungs- und Organisations-

systeme in der eignen Sprache nachvollziehen zu wollen, was als wertvoll, was als Problem und was als Wunsch definiert wird.

Der *Maßstab der Richtigkeit der Urteile* interessiert zunächst nicht, sondern nur derjenige der Konsistenz der Urteile. Wenn sich Studierende in auffallender Konzentration nach Anzahl und Dichte der Beschreibung für uninformiert hinsichtlich relevanter Abläufe und Zuständigkeiten an ihrem Fachbereich erklären, wird das bis zum Beweis des Gegenteils als *Argument* ernstgenommen. Es kommt darauf an, Argumente zu sammeln und Behauptungen erst einmal zu formen. Es wird häufig davon ausgegangen, die strittigen Argumente in der Bewertung des Alltags von Organisationen stünden fest, es komme lediglich keine Einigung zustande. Weit gefehlt: Auch Organisationen leben von Geschichten, Mythen, ritualisiertem Gerede und Klatsch, von Formeln der Verteidigung und Anklage, von Erzählungen des heroisierenden Erfolges, über Konspirationen und schmähhches Versagen.

Organisationsentwicklung und Organisationsberatung müssen neben der selbstverständlichen Erhebung quantitativer Daten auch auf *Sprache* setzen. Unsere Welten in und außerhalb von Organisationen sind sprachlich verfasst, und es gibt keinen Grund die damit verbundene Imperfektion für ein Ausschlusskriterium zu halten. Der Denkfehler liegt darin, die Unvollkommenheit der natürlichen Sprache für endgültig heilbar zu halten. Das könnte nur zu einer Verweigerung der Sprache, zur Sprachlosigkeit führen. Stanley Cavell hat gesagt: „Der Traum von einer technischen, einer perfekten Sprache war für mich auch der Traum von einer stimmlosen Sprache.“³⁶

Ich glaube, so muss das auch im Geschäft der Evaluation gehalten werden. Die Menschen in Organisationen *kontrolliert zur Sprache zu bringen*, muss ein Ziel bleiben. Das heißt nur, sie auf das ernsthaft zu verpflichten, was sie ohnehin ständig tun. Und es muss heißen, das Erfahrene explizit zu machen, um ihm nicht länger ausweichen zu können. Hierin liegt eine Art erster Heilung der eingeborenen Ungenauigkeit von Sprache. Zur Pathologie von Organisationen gehört z.B., dass die Verantwortlichen den sprachlichen Austausch vermeiden.³⁷ Das heißt nicht, dass sie überhaupt nicht sprächen, sondern dass sie problemaussparende Ersatz- und Ritualkommunikationen unterhalten. Das sollte nicht einfach als eine oberflächliche Fahrlässigkeit verstanden werden, sondern es handelt sich eher

³⁶ Stanley CAVELL (2001): *Nach der Philosophie*, 244.

³⁷ Gerrit BROEKSTRA (1998): *An Organisation is a Conversation*, 162.

um ein Symptom tieferliegender, ausgeblendeter Konflikte und eines Verlustes an Loyalität gegenüber den Organisationszielen.

Unser Verfahren der *Explizität* beruht zunächst auf der beschriebenen Eindringlichkeit der Befragung, die *von der Sprache zum Argument* führen soll. In einem ersten Schritt ist die Unterscheidung von persönlicher Äußerung und einem allgemeinen Urteil aufschlussreich. Prozessbeteiligte kennen sich auf Nachfrage hier sehr gut aus und können unterscheiden. Bereits das Aussprechen eines Urteils und das Ausformen eines Argumentes stellt einen Prozess der Reorganisation dar. Das entspricht nicht lediglich einer Befindlichkeit der Evaluatoren, sondern auch einer oft geäußerten Einschätzung der Befragten. Zwei Beispiele: Studierende sagen uns oft, dass sie noch nicht über die befragten Dinge nachgedacht hätten, aber die im Gespräch eingeführte Perspektive zu schätzen wüssten. Ein Professor hat einmal in einer auswärtigen Evaluation zu Beginn eines Gespräches sozusagen programmatisch gesagt, er sei jetzt zwölf Jahre im Amt, aber bisher habe noch niemand von ihm wissen wollen, was er von seiner Arbeit an der Hochschule erwarte.

Zwei beispielhafte Äußerungen, die bestätigen sollen, dass Sprache trotz all ihrer Grenzen das erste und zentrale und vielleicht entscheidende *Medium der Rechenschaftslegung* der Beteiligten von Organisations- und Kooperationsvorgängen aller Art darstellt. Es handelt sich um Rechenschaftsablegung sich und anderen gegenüber. Wahrscheinlich arbeitet der Mechanismus in umgekehrter Reihenfolge: die sich selbst geschuldete Rechenschaft wird durch das Eingeständnis anderen gegenüber gewonnen. Damit komme ich zu einem der Sprache eingebauten Korrektiv ihrer primären Begrenzung, das im Alltag so selbstverständlich wie im sozialwissenschaftlichen Gebrauch noch unausgestanden und gleichwohl von sehr weitragender Bedeutung ist. Ich meine die Kommunikation und besonders den *Dialog*.

4.4 Logik des Dialogs

Wen wir von der besonderen Funktion und der Philosophie des Dialoges sprechen, sind zwei Namen zu diesem Problemkreis besonders hervorzuheben. Damit ist die Spannweite der Diskussion ebenso gekennzeichnet wie deren Unausgestandenheit. Es handelt sich um Jürgen Habermas und David Bohm. Habermas ist sozialwissenschaftlicher Philosoph, Bohm ist

Physiker und Philosoph. Zunächst zu Habermas: Er erkennt, dass zwischen lebensweltlichen Bezügen einerseits und funktionalen Mechanismen der Integration andererseits Welten liegen. Verständigungsorientiertes Handeln hier, anonyme Systemsteuerung durch Medien wie Geld und Macht dort. Ohne sprachliche Nahbegegnung reichen die Konsensvorräte nicht aus und *Konsens* ist nicht mehr in Sicht.³⁸ Habermas' Diagnose von der ‚Kolonialisierung der Lebenswelt‘ stellt m.E. eine vehemente Aussage und ohne zusätzliche Differenzierung eine schwache Analyse dar. Seine Referate und ihre Auswertung der soziologischen Klassiker sind allerdings meisterlich. Einige Ableitungen aus seinen Hauptthesen erscheinen für unsere Zwecke anregend und der Mühe der Aufnahme wert zu sein.

Selbstverständlich sind *Organisationen medial gesteuert*. Sie sind, verglichen mit einer wie auch immer gearteten natürlichen Sprechsituation, höchst künstliche Gebilde der Programmierung, Koordination und Synchronisation von Handlungen auf ein *kollektives Ziel* hin, welches die zahlreichen Akteure von sich primär aus niemals anstreben wollten. Eben deshalb wird Motivation zu einem überaus knappen Betriebsstoff. Andererseits umspannen Organisationen solchermaßen intensiv das Alltagsleben der Mitglieder, dass es ohne die Vereinnahmung von natürlichem Austausch zwischen ihnen gar nicht zugehen kann. Hier gibt es zwar keinen Platz für idyllische Partizipationsideale, aber die Steuerungsfragen und Gestaltungsaufgaben erlauben auch nicht, *kommunikative Belange* als Fremdkörper und Störfaktoren zu betrachten.

Im Gegenteil: Ein zentrales Anliegen der neueren Organisationsberatung lautet in unterschiedlichen Fassungen, ‚gemischte Botschaften‘ müssten aus Verständigungs- und Kooperationsprozessen in Organisationen mit Hilfe komplexerer Lernformen aufgespürt und ausgeschlossen werden. *Gespräche* sollen eine *Form des anspruchsvollen Lernens* sein.³⁹ Neben der Forderung nach Rücknahme der Kolonialisierung der Lebenswelt steht Habermas' Appell an die Bereitschaft zur ‚gewaltlosen Intersubjektivität‘.⁴⁰ Wer wie ich diesen Appell an eine politische Öffentlichkeit als unrealistisch und philosophisches Missverständnis bewertet, muss doch dessen Nähe zu *Empfehlungen an die Organisationswirklichkeit* zur Kenntnis nehmen. Habermas fordert die Rückführung der verselbstän-

³⁸ Jürgen HABERMAS (1981): *Theorie des kommunikativen Handelns*. Bd. II, 341, 381, 393.

³⁹ Chris ARGYRIS, Donald A. SCHÖN (1999): *Die lernende Organisation*, 146.

⁴⁰ Jürgen HABERMAS (1981): *Theorie des kommunikativen Handelns*. Bd. I, 523.

digten Imperative institutioneller Regeln zu ‚motivationaler Verankerung‘.⁴¹ Das ist auch ein Anspruch moderner Organisationslehren. So hören wir von Peter M. Senge, dass es um die Erzeugung von ‚Aufregung‘ gehe, „die eine Organisation aus dem Profanen heraushebt.“⁴² Es sei der *Dialog*, der die Korrektur des demotivierten Denkens in Organisationen zustande bringen könne.⁴³ Im Dialog soll nach dieser Vorstellung offenbar für die Organisation mobilisiert werden, dies ist zunächst eine Umkehrung der Argumentation von Habermas, der die Verhältnisse in ein menschliches Maß bringen möchte. Ich glaube aber, dass auch der zweite Aspekt für Organisationen genutzt werden sollte.

Eine hinreichend konsistente theoretische Fundierung des Dialogs liegt noch nicht vor, aber es gibt Ansätze, welche eine sozialwissenschaftliche Verankerung erkennen lassen. Bemerkenswerterweise hat David Bohm wichtige Überlegungen zu diesem Thema beigesteuert.⁴⁴ Um seine Gedankengänge zu verstehen, muss man sich mit einem zentralen Baustein sozialwissenschaftlichen und kognitionswissenschaftlichen Denkens vertraut machen. Es handelt sich um die Tatsache, dass *Persönlichkeitssysteme* und *soziale Systeme* in enger Wechselwirkung entstehen und einer *Koevolution* unterliegen.⁴⁵ Es gibt keine Individuen ohne soziale Vorlagen oder Programme und ebenso wenig existieren soziale Systeme unabhängig von Kognitionen. Ich verstehe Bohm so, dass er den Dialog als Ort der Entstehung individueller und überindividueller Strukturen auffasst. Mehr noch, der Dialog stellt gleichsam die drahtlose Vernetzung mehrerer Gehirne her.

Der Dialog ist für ihn „ein Prozess direkter Begegnung von Angesicht zu Angesicht“⁴⁶ und muss von der Diskussion unterschieden werden, bei der es immer um Gewinnen wie in einem Kampfspiel geht. Beim Dialog, sagt Bohm, gewinnen alle.⁴⁷ Der Dialog setzt Energien der Kohärenz frei und ermöglicht eine Veränderung des Individuums und der Beziehungen, er ist eine Teilhabe an einem Zusammenhang. „Eine Transformation des Wesens des Bewusstseins ist möglich, auf der individuellen und auf der

⁴¹ Jürgen HABERMAS (1981) in der Kritik an Max Weber vgl. *Theorie des kommunikativen Handelns*. Bd. I, 278.

⁴² Peter M. SENGE (2001): *Die fünfte Disziplin*, 254.

⁴³ Ebenda, 294.

⁴⁴ David BOHM (2000): *Der Dialog*.

⁴⁵ Armand und Michèle MATTELART (1998): *Theories of Communication*, 135.

⁴⁶ Lee NICHOL (2000): Vorwort in: David BOHM, 14.

⁴⁷ David BOHM (2000): *Der Dialog*, 33f.

kollektiven Ebene.“ Ich unterschlage hier einige pathetische Zwischentöne Bohms um die Kerngedanken nicht zu gefährden.⁴⁸

Bohm spricht indirekt auch das Problem der *Motivation* an. Es gibt im Dialog so etwas wie eine kollektive Erzeugung von erkannter Notwendigkeit. „Notwendigkeiten schaffen mächtige Triebkräfte. ... Wenn die Leute überzeugt sind, daß etwas notwendig ist, verstoßen sie sogar gegen den Selbsterhaltungstrieb.“⁴⁹ Diese Aussage erinnert daran, dass Selbstinteresse der Ausgangspunkt von Motivation ist, dass es aber durch Bindung an Programme gerichtet und bis zur Auflösung des unmittelbaren Selbstinteresses verändert werden kann.

Ist das alles wirklich auf Organisationsberatung übertragbar? Hier muss zunächst unterschieden werden zwischen solchen Theorien der Kommunikation, deren Ziele primär in einer Moderationsschulung liegen und solchen, die das Vorgehen mit der Analyse von Situationen verbinden. Hierzu gehört z.B. das Konzept der situationsgerechten Kommunikation von Schulz von Thun.⁵⁰ Nach meiner persönlichen Erfahrung sind in der Organisationsberatung Virtuosen des kommunikativen Stimmungsmanagements besonders dann in Gefahr, nicht überzeugen zu können, wenn das Kapital des Wissens eine große Rolle spielt. Aber dann gewinnt die Fähigkeit, offene Fragen und Probleme sprachlich möglichst exakt zu fassen, zunehmend an Bedeutung. Und das auf beiden Seiten: Organisationsangehörige legen selbst großen Wert auf möglichst präzise Darstellungen, erwarten aber auch, dass Berater oder Evaluatoren, die ihnen gegenüber sitzen, grundsätzlich mit *Klarheit des Denkens und Argumentierens* aufwarten können.

Interessant erscheint mir nun, dass diese wechselseitige Bereitschaft, die Dinge möglichst genau beim Namen zu nennen, tatsächlich einen Effekt haben kann, welcher der von Bohm beschriebenen Teilhabe ähnlich ist. Ab einem bestimmten Punkt der Sammlung und Zusammensetzung von Wissen und Daten über das Geschehen in Organisationen nimmt die Neigung der Zurschaustellung eines bestimmten Habitus ab, der als selbstgefällige Abwertung fremder Perspektiven bei gleichzeitiger Selbstbeanspruchung von umfassender Kompetenz bezeichnet werden könnte. Besonders wenn Organisationsmitglieder in größerer Zahl zusammen sind, bemühen sich viele Fachleute und Führungspersonen um den Gestus eines

⁴⁸David BOHM (2000): *Der Dialog*, 99f.

⁴⁹Ebenda, 59.

⁵⁰Friedemann SCHULZ VON THUN (1998): *Miteinander reden*, 284.

unanfechtbaren Beurteilungsmonopols. Das ändert sich signifikant, wenn Auffassungen und Sichtweisen in *kleinerem Kreis* konfrontiert und zusammengesetzt werden können. So wird neben dem Wissensfortschritt und der Datenkenntnis auch die natürliche Gesprächssituation in kleineren Gruppen ermöglicht und befördert. Unter diesen Bedingungen kann es eher zu einem Austausch von Perspektiven kommen, zu dem, was vorher Transaktion von einer Person zu einer anderen genannt worden ist.

Wer diese Erörterungen für managementferne Spekulation hält, täuscht sich. Seit geraumer Zeit bemühen sich Management-Kalküle wie das *Harvard-Konzept* um die kommunikative Zusammenführung von Argumenten. Diese Darstellungen unterscheiden sich nicht von dem hier Vorgetragenen.⁵¹ Aufnahmefähigkeit und Sprachvermögen spielen sicherlich hier eine große Rolle. Das gilt natürlich auch für die Evaluatoren. Wer als Berater oder Evaluator zurückfragen kann, ob eine eigene Version der Darstellung den Sachverhalt treffe und die Antwort erhält, das sei ungewohnt, aber treffe den Kern der Sache, hat einen großen Schritt nach vorn getan. Das ist auch als ‚*Intervention durch Deutung*‘ bezeichnet worden.⁵²

Aber es geht im Kern noch um etwas Anderes. Um komplexes Wissen zu vermitteln bedarf es eines besonderen Schutzes vor Fehldeutungen und Missverständnissen, die entweder auf unterschiedlichen Perspektiven oder aber auch auf unterschiedlichen Interessen beruhen können. Dabei ist von einer Skala des Fehlverständnisses auszugehen, die von wohlwollenden über neutrale bis zu missgünstigen Grundhaltungen reicht. In allen Fällen ist bei fehlendem Primärverständnis für die Positionen anderer, ja sogar vor jedem Versuch der Herbeiführung von Verständigung von einer Art grundlegendem *mental*en Vorbehalt auszugehen. Wahrscheinlich liegt das an einem in den Sozialwissenschaften hinlänglich bekanntem Mechanismus der *Privilegierung des Selbstschemas* in all seinen Formen.

Interessant ist für das hier vertretene Anliegen jedoch, dass wir auf ein *Gefälle des mentalen Vorbehalts* stoßen, das von der medialen Vermittlung abhängt. Obwohl dieses Gefälle zum festen Bestand des alltäglichen Organisationswissens gehört, scheint es dennoch außerordentlich schwer zu sein, dieses intuitive Wissen systematisch umzusetzen. Im Alltag von Behörden, Organisationen oder Unternehmen werden die verbreitenden Nachrichten in sehr vielen Fällen deshalb angefeindet und müssen deshalb systematisch mentale Vorbehalte erzeugen, weil die *Informationsge-*

⁵¹ Z.B. Cornelia VEIL (1993): *Relationale Kommunikation*, 114, 128.

⁵² Heiner MINNSEN (1998): *Soziologie und Organisationsberatung*, 67.

schichte als grundsätzlich begehrtes Hintergrundwissen aus unterschiedlichen Gründen entweder schwer mitteilbar ist oder fälschlicherweise vorausgesetzt wird oder aber absichtlich zurückgehalten werden soll. Oft erschiene es absurd, vielleicht auch skandalös, diese Informationsgeschichte in eine formelle Verwaltungsmitteilung aufzunehmen. In Dienstbesprechungen könnten dagegen ergänzende Andeutungen und Rechtfertigungen einfließen und um Verständnis heischen. Die Größe des beteiligten Personenkreises ist von ausschlaggebender Bedeutung. In einem Abteilungsleitergespräch ist es eher möglich, mit dem Hinweis auf strenge Vertraulichkeit die wahren Hintergründe zu benennen. Ein Gespräch unter vier Augen erlaubt es, sich noch weiter vorzuwagen und Vertrauen zu schöpfen. Die Bedingungen für einen *Gedankenaustausch ohne jeden Vorbehalt* kommen dem sehr nahe, was hier als Dialogsituation beschrieben worden ist.

Damit ist das Gefälle der Kommunikation als eine *gestufte Verständigungschance* beispielhaft beschrieben. Offensichtlich handelt es sich um eine Skala mit der *Dimension der Anonymität/Nähe*, bzw. *Vorbehalt/Vertrauen*. In dem hier interessierenden Zusammenhang kommt es aus zwei Gründen darauf an, in der Evaluation im Sinne einer Organisationsberatung auf diesen zentralen Mechanismus das Augenmerk zu richten:

1. Die Qualität von Organisationsabläufen hängt in hohem Maße von der Berücksichtigung dieser *Kommunikations-Architektur* ab.
2. Für die Gewinnung von Einblicken in das Organisationsgeschehen ist es für die Evaluatoren unverzichtbar, die *Dialogseite* der Skala gestufter Kommunikation erreichen zu können.

Bisher haben sich die *Stufen des Kommunikationsgefälles* als einander ausschließend verstehen lassen. Das ist aber nur der erste Schritt. Es ist durchaus möglich, die *Informationswege miteinander zu verknüpfen*. Darin scheint mir sogar der entscheidende Erkenntnisgewinn zu liegen. Wo es um die regelmäßige Vermittlung und Umsetzung sehr schwieriger Sachverhalte geht, reichen formalisierte und technisierte Informations- und Kommunikationssysteme nicht aus, mehr noch, sie sind Konfliktherde. Zum Abbau mentaler Vorbehalte sind *institutionalisierte Dialoginstanzen* notwendig, die mit formellen Schalt- und Übertragungsstellen für Information verknüpft sind. In Organisationen können das z.B. Stabsstellen wie persönliche Referenten, Assistenten für besondere Aufgaben oder Leiter der Public-Relations-Abteilung usw. sein. Evaluationsverfahren

haben dort ihren Platz, wo selbst solche Stellen, die quer zur Hierarchie stehen sollen, in ein unausgesprochenes Interessengeflecht der Blindheit oder Interventionsblockade gegenüber strukturellen Missständen einer Organisation einbezogen sind. Hier helfen eventuell nur unbelastete Dritte mit dem Blick des Außenstehenden um die Verbote, die Dinge beim Namen zu nennen, zu durchbrechen. Es sei an das erinnert, was zu *intermediären Instanzen* gesagt worden ist.

Dialogisch gewonnene und fixierte Diagnosen dieser Art beruhen zweifellos auf Informationen und Daten, aber sie benötigen ebenso die interne Anerkennung, das interne Eingeständnis, das auch sprachlich aus der Mauer des Schweigens, Leugnens und der Abwehr gelöst werden muss. Die Logik des Dialogs ist und bleibt ein Medium für diese Mobilisierung. Blindheit, Selbstbetrug oder gar Zynismus gegenüber struktureller Fehlentwicklung und organisationalem Verfall benötigen die Blockade der Benennung und damit den Schutz interner Anonymisierung. Dieses Mittel der Mobilisierung durch externe Perspektiven und Beratung trifft exakt den Sachverhalt, der als *Organisationsentwicklung* beschrieben wird.

Es muss nicht weiter hervorgehoben werden, dass der Erwerb des *Deutungswissens* sehr voraussetzungsvoll ist und kaum und nur sehr begrenzt routinisiert werden kann. Aber es gibt hinführende Wege. Ein wichtiger Weg ist darin zu sehen, viele Dialoge mit unterschiedlichen Gruppen oder Einzelpersonen der Organisation zu führen. So kommt man zu Kontrastwissen, das sich in der Form der Frage und des Verweisens anwenden und in weitere Gespräche einführen lässt. Auch in diesem Fall kommt der Qualitätsbegriff über einen Maßstab zum Tragen. Der Vergleich entdeckt und wird zu einem wichtigen Element des Dialogs.⁵³

Ein weiteres Mittel besteht darin, das sogenannte *initial commitment* nachdrücklich zu erzeugen.⁵⁴ Der Dialog sollte nicht das Thema aussparen, ob der Dialog überhaupt geführt werden sollte. Davon habe ich oben bereits gesprochen. Die Erzeugung von ernsthaftem Beginn und die Bindung der Beteiligten an das Verfahren kann je nach Feld unterschiedlich sein und von einer unspezifischen Verabredung bis zu einem Vertrag reichen. Wir haben uns in unseren Verfahren, wie berichtet, mehr und mehr auf vertragliche Abmachungen gestützt und haben erfahren, wie

⁵³ Vgl. die Grundlagen der Comparison-Theorie Festingers u.a.: Steve DUCK (1998): *Human Relationships*, 135.

⁵⁴ Steve DUCK (1998): *Human Relationships*, 144.

hilfreich das ist. Es sollte in Zukunft keine Evaluationen mehr geben, die als Basis nichts anderes haben als ein loses Übereinkommen über ein wohlwollendes Hinsehen und Besprechen im Rahmen einer Evaluation.

Das Verfahren des Dialogs könnte auch methodisch ausgearbeitet werden. So haben wir beispielsweise in einer hoch kontroversen Gruppe von Führungspersonen sehr intensive Einzelgespräche, z.T. auch Kleingruppengespräche geführt, je nach Konstellation anstehender Entscheidungen. Die Protokollierung wurde autorisiert und anschließend eine aufwendige *Synopse* erstellt. Die Ergebnisse sind für die Betroffenen äußerst verblüffend gewesen. Man stelle sich vor: fast alle einzeln Interviewten erklären, es gebe einen beklagenswerten Missstand, man selbst sei bereit ihn zu beheben, aber die Anderen seien für die Blockade einer entsprechenden Initiative verantwortlich. In einem solchen Fall, der zugegebenermaßen der krasseste war, kann sich niemand mehr der Verantwortung entziehen. Die besondere Form des Dialogs hat auch erst die Intimität der Zuschreibung von Verantwortung und Nicht-Verantwortung ermöglicht. Der *herrschaftsfreie Diskurs* Habermas'scher Färbung ist hier sozusagen *synthetisch* mit Hilfe von offener Gesprächssituation und methodischer Auswertung hergestellt worden. Was Bohm die ‚Idee des Dialogs‘ nennt, ist ebenfalls gegeben, denn es entsteht eine überindividuelle Situation der Klärung.

Die weiter oben entwickelten diversen Formen des Beratens i.S. einer intermediären Instanz kann in eine Theorie des Dialogs integriert werden. Die Betreiber von Evaluationen sind Außenstehende, sie sind Dritte, die geduldet und sogar gerufen sind, um sich ein Bild von Transaktionsprozessen zu machen. Sie stehen sozusagen zwischen denen, die administrative oder politische Entscheidungen zu treffen haben und über Anordnungs- und Entscheidungsbefugnisse verfügen. Es ist die Frage, inwieweit diese Dritten als unbeteiligt Beteiligte sich langwierig selbst Bilder machen sollten vom Organisationsgeschehen oder ob es abkürzend einfacher wäre, Verantwortliche der Organisation zu befragen. Ich glaube beim jetzigen Stand unseres Wissens, dass sich eine Mischung empfiehlt. In dem Maße, in dem Evaluationswissen auf einem soliden Selbstverständnis als Diagnose und Beratung aufgebaut ist, kann es Sachdialoge geben. Man wird feststellen, dass diese mit fortschreitender Zusammenführung unterschiedlicher Situationsdefinitionen und mit immer unabweislicher werdender Offenlegung divergierender Interessen zunehmend ernster und verantwortlicher werden. Aber diese Schwelle zum echten Dialog muss zunächst einmal mit Hilfe der Evaluatoren die eigenen Vermeidungsan-

strenge, Verweigerungen und mentalen Vorbehalte gegenüber der Außensicht überwunden werden. Das erfordert sachliche und intellektuelle Arbeit beträchtlichen Ausmaßes.

Es gelingt mir hoffentlich die Klärung darüber, dass es keinen Gegensatz zwischen Datenerhebung, Situationsverständnis und Gesprächsdialog grundsätzlicher Art gibt. Kompetenz fördert Entspannung, wenn der Austausch im Stile einer wirklichen Transaktion – der Begriff Dialog kommt dem nahe – gelingt. Dialog befördert eine neue Bewertung, im Grenzfall die Freude, die *Wahrheit hinter den Wahrheiten* zu sagen. Auch der Organisationstheorie sind solche Überlegungen nicht mehr fremd. Der bekannte Karl E. Weick besteht darauf, dass es wichtig sei, ‚intrasubjektive Bedeutungen‘ von ‚intersubjektiven‘ zu unterscheiden. Gespräche können nach seiner Ansicht in Organisationen eine *Transformation vom Ich zum Wir* herbeiführen. Daneben gibt es für ihn noch ein systemgebundenes ‚Gattungsselbst‘, in dem spezifische Kulturmuster aufgehoben sind.⁵⁵ Diese Dreiteilung repräsentiert eine Mehrebenenanalyse, in der sich Programme mit Personen und Personen mit Personen über Programme verknüpfen.

Hier muss die Fähigkeit der Evaluatoren, Interessen zu verstehen, Standpunkte nachzeichnen zu können, Perspektiven zusammenzuführen und Motivationen modellieren zu können, nicht erneut erörtert werden. Verhandlungs- und Beratungsfähigkeit sowie Transaktionsleistungen erfordern *Kommunikation* und besonders die *Sprache*. An Methoden formalisierte Formen der Darstellung und des Berichts aus natürlicher Kommunikation entstehen zu lassen muss m.E. noch gearbeitet werden. Die Welten, in denen wir leben, sind aus Sprache gemacht und wir sollten nicht zögern, auch *Sprache* zu einem *Medium der Qualitätsermittlung* zu erklären. Sprache ist aber nicht nur ein Medium, sondern immer auch die Sache selbst. Die Tatsache, dass die Sondersprache z.B. der Disziplinen in Alltagssprache oder auch eine andere Sondersprache, z.B. der oben beschriebenen Disziplinen, welche das Evaluationshandwerk bedienen können, übersetzt werden müssen, kann detektivisch sein und Betriebsblindheiten aufdecken. Hinzu kommt, dass in anspruchsvollen Feldern der Organisationssteuerung die Sprachfassung ohnehin nicht nur ein darstellendes, sondern auch ein *motivierendes Element* ist.

Die *Sprachfassung des Verhaltens* sollte also nicht grundsätzlich methodisch verdächtig, sondern als Motor zur Offenlegung von zunächst als

⁵⁵ Karl E. WEICK (1995): *Sensemaking in Organizations*, 71.

privat deklarierten Angelegenheiten angesehen werden. Offenlegung ist ein paradigmatisches Mittel der Klärung und Kooperation. Offenlegung ist der Schritt in die Erweiterung des Bewusstseins, sie schafft die Verbindungen zur Zusammenführung der zunächst disparaten Perspektiven. *Sprache ist kollektiv*⁵⁶ und privat, sie ist eine Brücke für Engagement. Sie eröffnet die Einrede durch Dritte und zeigt auf, dass die Menschen, was immer sie tun und welche Position sie einnehmen, mit ihren partikularen Interessen nicht nur einsam, sondern auch unvollkommen sind. Sprache ist ungenau, voller Fallen und doch auch reich an Kombinationsmöglichkeiten und Erfindungen. Sprache ist Festlegung und Kreativität, damit Ausdruck einer zentralen menschlichen Befindlichkeit nicht nur in Organisationen.

Sind Sprechen und Kommunikation für diese Zwecke lehrbar? Ich glaube nur indirekt, nämlich in der Form des situativen Lernens, das auch als Habituslernen zu bezeichnen ist. An anderer Stelle haben wir das als *Wissen 2* bezeichnet. *Wissen 1* hat die Standardform des Erwerbs von kodifiziertem Wissen wie es in Büchern zu stehen pflegt. Die Dominanz dieser Wissensform lässt vergessen, welche Bedeutung *Wissen 2* zukommt.⁵⁷ Es gibt gute Versuche die Vielfalt der Ansätze der *Kommunikation in der Organisation* darzustellen,⁵⁸ aber hiervon verspreche ich mir nicht allein einen Lernerfolg im Sinne von *Wissen 1*. Was ich von einer Vermittlung der Kenntnis über Kommunikationsgeschehen in Buchform erwarten zu können glaube, ist allenfalls eine Art der Sensibilisierung für die beschriebenen Vorgänge in ihrer Komplexität und ungeheueren Bedeutung. Um den Anforderungen praktischer Interaktion gerecht werden zu können, muss sich dieses Wissen mit demjenigen Augenmaß und Fingerspitzengefühl verbinden, das nur in Szenen der Anschauung und lebensweltlichen Rückkopplung gewonnen werden kann. Evaluationsverfahren sollten darauf vertrauen können, dass diese Fertigkeiten ihren Verantwortlichen nicht fremd sind.

⁵⁶ David BOHM (2000): *Der Dialog*, 40, 45.

⁵⁷ Manfred HENNEN, Elisabeth SPRINGER (1994): *Modellprojekt zur Förderung von Studium und Lehre*, 63f.

⁵⁸ Anton HAHNE (1998): *Kommunikation in der Organisation*.

4.5 *Der Mechanismus der Defektion*

Bisher bin ich stillschweigend davon ausgegangen, dass mangelnde Verständigung und Blockade einer Überbrückung unterschiedlicher Perspektiven für Missstände in der Organisation verantwortlich sind. Sie zu beheben macht eine Investition an außeralltäglicher und von außen unterstützter Mobilisierung notwendig. Eine so verstandene Evaluation unterstellt sozusagen *good will* unter der Schicht des alltäglichen Flickwerks und der organisationalen Unzulänglichkeit. Dagegen ist einzuwenden, diese Sichtweise sei zu optimistisch und in der Wirklichkeit gebe es nicht selten eine Dauerverweigerung gegenüber kooperativem Problemlösen. Das nennt die Sozialtheorie und Organisationswissenschaft ‚Defektion‘ – ein quasi-technischer Begriff für Verrat und Verweigerung.⁵⁹

Solche Fälle gibt es in der Tat. Sie lassen sich u.a. daran erkennen, dass von den Organisationsmitgliedern quer durch die Hierarchien erhebliche Anstrengungen unternommen werden um die Tatsache der Defektion zu verschleiern und ein Bild engagierter Kooperation vorzugaukeln. Sich für die tatsächlichen Problemlösungen mit den notwendigen Differenzierungen einzusetzen wird in persönlichen Gesprächen als Blauäugigkeit und Sentimentalität gekennzeichnet. Sich einer optimistischen Sicht kollektiver Einrichtungen und Organisationen hinzugeben gilt als ein Zeichen von Weltfremdheit oder gar Dummheit. Mir sind diese Argumente in ihrer Abstufung nicht unvertraut, aber es erscheint mir höchst aufschlussreich zu sein, dass sich in diesem Szenario genau das abspielt, was oben beschrieben worden ist. Die Defektion wird legitimiert mit dem Verhalten der anderen; die Welt ist eben so und will es so. Die Hölle, das sind die anderen, wie wir wissen. Mit anderen Worten: Für die eigene Person wird Defektion verworfen und Kooperationswille unter anderen sozialen Randbedingungen zumindest eingeräumt. Voraussetzung für das Eingeständnis einer bitteren Wahrheit dieser Art ist eine Dialogsituation; öffentlich muss verschleiert und geleugnet werden.

Robert Axelrod hat mit seinen aufwendigen spieltheoretischen Simulationen zu Kooperation und Defektion die Mechanismen der Überwindung von Defektion nachgewiesen, indem er unterschiedliche Strategien der Gewinnung und Erhaltung von Kooperation testete. Die erfolgreichen Strategien hat er als ‚evolutionär stabil‘ bezeichnet. Die Ergebnisse der Spieldurchgänge „zeigen, daß unter geeigneten Bedingungen tatsächlich Kooperation in einer Welt von Egoisten ohne zentralen Herrschaftsstab

⁵⁹ ROBERT AXELROD (1988): *Die Evolution der Kooperation*.

entstehen kann.⁶⁰ Das Geheimnis besteht darin, dass die Menschen unterschiedliche Neigungen haben, den zuvor beschriebenen – wie ich es nennen möchte – *organisationalen Zynismus* zu überwinden. Die kooperativeren Personen, die unter dem allgemeinen Zustand der Defektion mehr leiden als andere und stärkere moralische Anlagen zur Reorganisation haben, sollten zunächst gewonnen werden können. Die Spielregel, die in allen Versuchen den Sieg davongetragen hat und die es für die Entstehung von Kooperation herzustellen gilt, heißt ‚TIT FOR TAT‘. Das heißt soviel wie ‚wie du mir so ich dir‘. Diese Regel oder Haltung besteht in meinen Worten in einer *Grundbereitschaft zur moralischen Investition*, die solange aufrecht erhalten wird wie die anderen keinen Verrat begehen. *Moralisch* heißt hier nichts anderes als die Grundhaltung aufzubringen, zum ersten Schritt der Wiederherstellung von Kooperation, der Überwindung des Dilemmas der wechselseitigen Verdächtigung und der mentalen Vorbehalte bereit zu sein. Es lohnt sich die vier Entscheidungsregeln der genannten Strategie unter dem Aspekt der Zielsetzung von Evaluation zu lesen:

„1. Vermeidung unnötigen Konflikts durch eigene Kooperation solange der andere Spieler kooperiert, 2. Provozierbarkeit angesichts einer unnötigen Defektion durch den anderen, 3. Nachsicht nach der Antwort auf eine Provokation, 4. Verständlichkeit des Verhaltens, damit sich der andere Spieler an das Verhaltensmuster des Spielers anpassen kann.“

Es besteht eine gewisse Parallele zwischen dem Verfahren der Evaluation und den Mechanismen sich bildender Kooperation nach dem Schema von Axelrod. Beiden ist gemeinsam, dass die Reorganisation von Kooperation und die Überwindung von Defektion mit Hilfe von Selbstorganisationskräften eingeleitet werden soll. Evaluation ist nach dem hier vertretenen Verständnis nicht als kontrollierende Sanktion zu verstehen, sie soll vielmehr durch Konfrontation mit internen Organisationskonflikten in Gang kommen. Dazu muss für moralische Investitionen im oben beschriebenen Sinn geworben werden. Sanktionen sind aber keineswegs völlig abwesend; sie wirken dadurch, dass jede ‚unnötige Defektion‘ nach der Strategie 2 mit unmittelbarer Verweigerung der Kooperation bestraft wird. Auch hier gilt ‚wie du mir, so ich dir‘. Auch in Evaluationen soll jeder Defektionsmöglichkeit mit dem Aufweis gegnerischer Defektionsmöglichkeiten begegnet werden. Die Überbrückung der oben beschriebenen

⁶⁰ ROBERT AXELROD (1988): *Die Evolution der Kooperation*, 18.

nen Differenz zwischen öffentlich und persönlich soll ja dort gelingen, wo die persönliche Einschätzung öffentlich gemacht werden kann und die öffentliche Einschätzung durch persönliche Revision korrigiert wird. Defektion lebt von der bewussten Aufrechterhaltung des öffentlichen organisationalen Zynismus. Die Spannung zwischen öffentlicher Darstellung und persönlicher Motivation tritt besonders dann zu Tage, wenn es um Handlungsvorbilder für Organisationen geht. Wer sich mit den Grundlagen der Evaluation und ihrer Funktion befasst, muss sich damit auseinandersetzen. Das soll im folgenden Kapitel geschehen.

5. Leitbilder

5.1 Externer und interner Werbefeldzug

Der politische, administrative und industrielle Westen befindet sich in einem Zustand der meist ökonomisch motivierten Reformbesessenheit. Damit wird Sinnvolles avisiert, jedoch gleichzeitig beträchtliche Instabilität erzeugt. Diese Instabilität kommt u.a. von zu vielen Baustellen, die für gegenseitige Behinderung sorgen. Wenn man viele Variablen verändert, wird die Kontrolle der Wirkungen im Einzelnen unmöglich. Die Diagnose des Reformstaus, der *Reformblockade*, ist in weiten Bereichen glaubhaft. Dem steht entgegen, dass die großen Reformen – und dazu gehört sicherlich in Deutschland die *Bildungsreform* – nicht gelingen wollen. Zwischen diesen beiden Befunden der Blockade und der unkontrollierten Wechselwirkungen taumelt der *Reformaktivismus*. Er kommt mit vielen neuen Namen einher, hinter denen sich im günstigen Fall alte Weisheiten verbergen. Dazu gehört der Begriff der Evaluation, dessen reinigende Klärung mein Anliegen hier ist; dazu gehört auch der *Begriff des Leitbildes*. Ich habe die Überschrift ‚Leitbilder‘ gewählt, um deren *Inflation* zu kennzeichnen. Inflation bedeutet nicht nur in der Ökonomie Vermehrung bei gleichzeitiger Entwertung, wobei sich diese Prozesse wechselseitig bedingen.

Ob Energieversorgungsunternehmen, Krankenhäuser, Kommunen oder Universitäten, Leitbilder werden als Darstellungs- und Mobilisierungsmedium erfunden und eingesetzt. Das ist für sich genommen keineswegs anrüchig, suspekt ist nur das *Maß an Uniformität*. Die Häufung der werthaltigen Begriffe und Superlative zur Bestimmung der Leitbilder ist atemberaubend. Sie folgen überwiegend einer wissenschaftlichen und

gesellschaftlichen Leitwährung der Ökonomie. Als Volkswirt möchte ich betonen, dass es sich um eine Spezifikation der Ökonomie handelt, nämlich der Betriebswirtschaft und noch enger der Geschäftsführungswissenschaft.⁶¹ ‚Effektivität‘, ‚Effizienz‘, ‚Transparenz‘ und vor allen Dingen ‚Kundenorientierung‘ dominieren.

Es ist unbezweifelbar, dass es sich hier um die Beschreibung von Mechanismen handelt, deren zentrale Funktion in Koordinierungsleistungen besteht, es handelt sich durchgehend um *Markt- und Allokationsmechanismen*. Diese Mechanismen sind jedoch keine Repräsentanten einer institutionellen Verfassung. Sie sind gänzlich unspezifisch. Die ihnen in den Leitbildern regelmäßig zugeschriebene Kraft zu Mobilisierung und Motivation ist nicht gänzlich in Abrede zu stellen, bedenklich stimmt nur die Ferne zum jeweiligen Typus von Organisation oder Institution. Motivieren kann man nicht in Breitstreuung, allenfalls kurzfristig in einer Art Strohfeuer. Motivation bedeutet das Vorhandensein oder die Erzeugung von Selbstbindungen sachlicher und persönlicher Art. Wie soll das erfolgen, wenn der eigene Lebens- und Verantwortungsbereich keinerlei Konturen in der Aufforderung zur Schaffung von Transparenz Gestalt annimmt? Die Menschen in Organisationen verpflichten sich nicht für einen abstrakten Begriff; *Motivation* ist immer gegenüber seinen jeweiligen Objekten der Verpflichtung ein sozusagen *hautnaher Zustand oder Prozess*.

Am häufig verwendeten Begriff der *Kundenorientierung* lässt sich das besonders klar machen. Von ihm werden Auswirkungen auf die Motivationsleistungen der Organisationsmitglieder ausdrücklich erwartet. Meine Frage lautet im Anschluss an die obigen Ausführungen über Defektion und organisationalen Zynismus: Was soll einen Mitarbeiter dauerhaft in seinen Verhaltensdispositionen beeinflussen und verändern, deren Fernziele mit der Ausrichtung auf Kundenorientierung beschrieben werden, wenn er gleichzeitig ein Modell desinteressierter Vorgesetzter und Kollegen zu erleben glaubt? Oder: Was soll ein Referatsleiter an Motivations-schub empfinden, dem nicht reiner Wein eingeschenkt worden ist, während man ihm Versprechungen hinsichtlich seiner personalen Ausstattung macht und während öffentlich von Transparenz und Kundenorientierung geschwärmt wird?

Kundenorientierung bedeutet so etwas wie freundlich, aufgeschlossen zu sein und seine Kooperationspartner ernst zu nehmen. In Organisationen

⁶¹ Im Französischen heißt das treffend: sciences économiques relatives à la gestion.

sollen die Organisationsmitglieder die Kunden sein. Das Kooperations-Dilemma ist deutlich wieder zu erkennen. Das Verständnis als Kunde soll die Krankheit der Defektion heilen. Wer möchte das nicht unterschreiben wenn dem so wäre? Der Kundenbegriff ist aber in der einschlägigen Literatur der OrganisationsökonomInnen firmenintern nicht zu verwenden.⁶² Das Handeln der Organisationsmitglieder ist gerade nicht als sehr enges und spezifisches Nachfrageverhalten zu begreifen, das von klaren Präferenzen für ein Produkt ausgeht. Parastaatliche Organisationen der Konsumentenberatung beschreiben z.B. weit treffender das Problem von Studierenden, die sich auf eine Universitätsausbildung einlassen wollen als deren Definition als Kunden. Der Markt setzt vollzogene Entscheidungen voraus, die Organisation oder Firma verlangt dagegen die Bereitschaft des sich einordnenden Beitrags und der Kooperation. Die Leistungen des Marktes sind geradezu unbegreifbar wenn dieser Unterschied nicht verstanden ist.

Mit anderen Worten: Wir reden in Organisationen von der Herstellung von adäquaten Entscheidungen, im Marktmodell gelten sie als vollzogen und vorauszusetzen. Motivation ist im letzten Fall fraglos, im ersten das zentrale Problem. Die Kundenorientierung lässt uns mit der Aufklärung über die Herstellung von Motivation allein. Der Leitbild-Begriff ist als eine Art unspezifischer *Werbung um Motivation nach außen und innen* zu verstehen. Nach außen insofern, als die entsprechende Einrichtung sich engagiert zeigen soll. Das spielt eine große Rolle im heutigen *Marketing*-zentrierten Leben. Schmucke Broschüren tun hierbei ihre Wirkung. Es gibt aber auch eine Marketing-Aufgabe nach innen. Die Mitarbeiter der Organisationen sollen werbend gewonnen werden für einen stärkeren Einsatz auf der Basis eines stärkeren Engagements für die Belange der Organisation. Das Geschäft der Werbung und dessen Wirkungen ist uns allen vertraut, aber auch dessen Grenzen. Die Mundpropaganda stellt ein weitaus stärkeres Gegengewicht dar. Unspezifische propagandistische Entwürfe ohne Realitätsanbindung sind immer kurzatmig.

Diese kritischen Überlegungen zum Leitbildbegriff sollen aber nicht außer Acht lassen, dass hinter dem Leitbildbegriff *fundierte und traditionsreiche verwandte Begriffe* stehen, denen es sich zuzuwenden lohnt, gerade wegen der großen Nachfrage nach Leitfunktionen.

⁶² Reinhard PIRKER (1997): *Die Unternehmung als soziale Institution*.

5.2 Idee und Leitidee

Zunächst ist dem Begriff der ‚Idee der Universität‘ nachzugehen. Er mag auf den ersten Blick antiquiert erscheinen, er ist es aber keineswegs. Diese Bezeichnung ist mindestens 200 Jahre alt und bis auf den heutigen Tag immer wieder zur Kennzeichnung der Aufgabenstellung der Universität und ihres Selbstverständnisses für Wissenschaft, Bildung, Ausbildung und Gesellschaft herangezogen worden. Wichtig scheint mir, dass er besonders in deren Krisenzeiten bemüht und dabei z.T. aktualisiert worden ist. Die Idee der Universität beschreibt zwischen 1800 und 1900, einige Autoren sagen bis zum Ausbruch des zweiten Weltkrieges, die in dieser Zeit international vorbildliche deutsche und z.T. auch angelsächsische Universität. Seit etwa 1850 wird unter dieser Bezeichnung auch ein Modell der erfolgreichen Forschungsuniversität, nach 1900, später nach nordamerikanischem Vorbild eine erfolgreiche Reformuniversität verstanden.

Ich kann hier nur einige wichtige Autoren und Zeitpunkte der Thematisierung des Begriffs aufführen. Um 1800 haben Schelling, Fichte, Schleiermacher und W. v. Humboldt diesen Begriff zum Motor ihrer Reformmaßnahmen gemacht.⁶³ Um 1850 entwickelte der Kardinal Newman eine ausführliche Konzeption der Idee der Universität, in deren Mittelpunkt die Vorstellung stand, dass Wissenschaft und Bildungsvermittlung ein Selbstwert seien.⁶⁴ *Selbstwert* ist ein starker Ausdruck für die höchste Form von Motivation. Die Basis des primären Selbstinteresses ist durch eine geglückte und faszinierende Form der Bindung an ein Ziel sozusagen ausgeblendet und vergessen.

Um 1920 erneuerte der preußische Kultusminister C. H. Becker die Vorstellung, dass die Spannung zwischen einer starken Regierung und einer starken Universität deren beider Erfolg garantiere.⁶⁵ Karl Jaspers erneuerte 1945 in einer schwierige Zeit mit Nachdruck die Verwendung des Begriffs der Idee der Universität.⁶⁶

Auch danach hat der stehende Begriff Idee der Universität seine Aktualität behalten. Helmut Schelsky, der Gründer der Universität Bielefeld, hat diesen Begriff in den Mittelpunkt seiner theoretischen Überlegungen

⁶³ J.J. ENGEL u.a. (1990): *Gelegentliche Gedanken über Universitäten* [zuerst 1802-1810].

⁶⁴ John Henry NEWMAN (1996): *The Idea of a University*. Jaroslav PELIKAN (1992): *The Idea of the University*.

⁶⁵ Helmut SCHELSKY (1971): *Einsamkeit und Freiheit*.

⁶⁶ JASPERS, Karl (1946): *Die Idee der Universität*.

gestellt.⁶⁷ Talcott Parsons und Gerald M. Platt haben für das amerikanische Universitätswesen vergleichbare theoretische Anstrengungen gemacht und der Sache nach diesen Begriff ausdifferenziert.⁶⁸ Die Verwendung dieser Formel ist auch in den Naturwissenschaften eine Selbstverständlichkeit gewesen. Der Nobelpreisträger Manfred Eigen verwendet sie gemeinsam mit anderen Autoren oberbegrifflich.⁶⁹ Bemerkenswert ist, dass auch aktuell diese Bezeichnung als eine Art *Erkennungszeichen* aufrechterhalten wird. Es gibt sogar so etwas wie eine aktuelle Renaissance dieses Begriffes offenbar in Reaktion auf fehlgeschlagenen Reformübereifer.⁷⁰ Wie erklärt sich das und wie gehen wir damit um?

Das Schwerste scheint zu sein, heute überhaupt verständlich zu machen, was als *Idee* in diesem Zusammenhang aufzufassen sei. Einer der bedeutendsten Hochschulforscher Eduard Spranger hat das 1930 so treffend formuliert, dass es bis heute Gültigkeit hat. Spranger hat auch deshalb als einer der Großen zu gelten, weil er dem Nationalsozialismus widerstanden hat. Er soll deshalb hier ausführlich zu Wort kommen: Nach seinem Verständnis liege die Stärke der Universität in:

„der Anschauung, daß Wissenschaft nicht nur für den unmittelbaren Bedarf und Verbrauch produziert werden dürfe, sondern ihrem eigenen, inneren, tendenzfreien und beschränkungsreichen Gesetz überlassen werden müsse. In dieser freien Selbstentfaltung werde sie dem Leben am besten dienen. Denn für unvorhergesehene Bedarfsfälle halte sie dann Lösungen bereit, deren Verwendbarkeit die Praxis allein niemals geahnt hätte. Man sagt nicht zu viel, wenn man in dieser Idee den eigentlich entscheidenden wissenschaftlichen Grundgedanken der deutschen Universität findet. Sie ruht – in einem Bilde aus der Wirtschaftswissenschaft – bewußt auf dem Prinzip der freien, unregelten Wissensproduktion.“⁷¹

Hier wird die sinnvolle Grenze der Differenzierung in Universitäten und Fachhochschulen sichtbar. Wäre diese Grenze mit ihren Möglichkeiten des Übergangs doch verwirklicht! Zurück zum Begriff der Idee. Eine Idee

⁶⁷ Helmut SCHELSKY (1971): *Einsamkeit und Freiheit*.

⁶⁸ Talcott PARSONS und Gerald M. PLATT (1990): *Die amerikanische Universität*.

⁶⁹ Manfred EIGEN u.a. (1988): *Die Idee der Universität*.

⁷⁰ Helmut BACHMAIER und Ernst Peter FISCHER (Hrsg.) (1997): *Streit der Fakultäten. Oder die Idee der Universität*; Peter BAGGEN (1998): *Knowledge Society and the Idea of the University*; Michael DAXNER (1999): *Die blockierte Universität*; Wolf LEPENIES (1989): *Gefährliche Wahlverwandtschaften*.

⁷¹ Eduard SPRANGER (1973): *Das Wesen der deutschen Universität*, 107.

ist kein Ideal, das lediglich die Formulierung eines Höchstwertes darstellt, sie ist ein Entwurf einer spezifischen Lebensorientierung mit utopischen Zügen und der gleichzeitigen Kraft der handlungsleitenden und sinnstiftenden Regulation. Seit der frühen Neuzeit ist die Gelehrsamkeit als Wissenssystem und eigenständiges Sozialsystem eine ausschlaggebende ‚Leitformel‘ für die Universitätsidee.⁷² Damit ist das, was über einen langen Zeitraum ‚Idee‘ genannt worden ist, die *gehaltvollere Variante* dessen, was heute Leitbild heißt.

Universitäten sind primär weder Marktsysteme noch bürokratische Systeme noch Systeme demokratischer Vereinigungen.

„Es ist ein Mißverständnis unserer Zeit, bei der Bewertung von Institutionen aller Art von feststehenden Präferenzen der Klienten und von Preisvergleichen ausgehen zu wollen. Universitäten, deren Leistungen gerade darin bestehen, Studien-, Lehr- und Forschungskulturen zu etablieren, fordern von allen Beteiligten Leistungsbereitschaft, insofern verlangen sie Kredit und Motivation. Das ist aber das Gegenteil von Konsum, es ist Konsumverzicht.“⁷³

Der Kanadier Bill Readings hat exemplarisch am Begriff der ‚excellence‘ als einem der inhaltsleeren Leitbildsuperlative nachgewiesen, dass er symptomatisch für den Verlust eines motivierenden Bildes von der Universität sei. Die Idee der Universität bedürfe eines *spezifischen Wertes*, dem zu folgen es den Mitgliedern als lohnend erscheine. Eine Idee müsse bei aller Unschärfe Ziele und Zwecke des Handelns lebendig halten können. Die Idee der Kultur und der Bildung hätten das während der Blüte der deutschen Universität erfüllt, es gebe keine Entsprechung außer der Leerformeln wie excellence.⁷⁴

Ich hoffe verständlich machen zu können, wie verhängnisvoll Begriffe wie excellence und Kundenorientierung für die Universität sind. Die Orientierung an Wissenschafts- und Bildungswerten sind Ideen eines *höchst spezifischen Sozialsystems*. Jedes soziale System lebt von seiner Spezifikation und damit von seinen *Grenzen*. Es bleibt deshalb ein offenes System. Aber wenn die Schleusen geöffnet werden, diffundieren die Aufgaben und Motivationen verfallen. Ich fürchte, die Universität ist heimlich, still und leise eine ‚*Gesamtuniversität*‘ geworden in dem Sinn,

⁷² Rudolf STICHWEH (1991): *Der frühmoderne Staat und die europäische Universität*, 114.

⁷³ HENNEN, Manfred (1998): *Leitbilder zur Evaluation der Lehre*, 70.

⁷⁴ BILL READINGS (1997): *The University in Ruins*, 17, 54, 160.

dass sich in ihr die unterschiedlichsten und ungeschiedenen Motivationen, Ziele und Lebensentwürfe der verschiedenen Mitglieder finden. Der politische Entwurf ist hiermit weder angesprochen noch kritisiert.

Beginnen wir bei den Motivationen der Studierenden: Motive von Berufserfolg, Karriere, Unentschiedenheit des Lebensentwurfs, fehlender Alternativen am Arbeits- und Ausbildungsmarkt, Traditionalismus – die Eltern sind Akademiker –, Wissensdurst und Bildungsinteresse existieren ungeschieden nebeneinander. Das löst eine Institution von innen auf. Alle Floskeln der ‚corporate identity‘ sind nur Symptom, nicht Lösung dieses Dilemmas. Es ist nicht einmal sichergestellt, dass Wissens- und Bildungsarbeit in der Universität noch durchgehend in allen Fächern als wirklicher Lebensmittelpunkt anzusehen sind. Die Leistungsanforderungen sind auseinandergefallen und teilweise wird als Programm formuliert, man solle außerhalb der Universität die Voraussetzungen für den Berufserfolg suchen. Das sind, ob es zutrifft oder nicht, Auflösungserscheinungen für ein System. Es ist symptomatisch, dass sich gleichzeitig private Hochschulen, die hohen Einsatz und viel Studiengeld verlangen, hoher Nachfrage erfreuen.

Die Motivationen der Professoren sind in ihrem Selbstverständnis ebenfalls weniger einheitlich geworden. Neben die handlungsleitenden Ideen der Wissenschaft, Forschung und Wissensvermittlung sind m.E. mehr und mehr *universitätsunspezifische Karriere-Muster* getreten. Ich sage das mit Zurückhaltung, glaube aber nicht übersehen zu können, dass ein maximaler akademischer Titel ohne korrespondierende akademische Verpflichtungen dem Muster der privat erfolgreichen und öffentlich verheerenden *Defektion* entspricht. Ich weiß von Professoren, die in ihren Vorlesungen ihren Hörern erklären, es sei ihnen unmöglich, von dem erbärmlichen Professorengehalt existieren zu können. Andere geben zu verstehen, dass sie die Studierenden für ihre Dummheit verachten ohne sich ernsthaft für eine Differenzierung oder gar Änderung dieses Zustandes zu engagieren. Nicht wenigen ist die Karriere nach der Universität das eigentliche Ziel. Leitbilddiskussionen hätten hier anzusetzen und nachzuforschen. Aber das ist nicht leicht.

Schließlich sei zu dieser Frage noch auf eine möglicherweise unerwartete Parallele in der anspruchsvollen Unternehmungsberatungsliteratur hingewiesen. Peter F. Drucker spricht in einer sehr ernst zu nehmenden Weise von *‘Unternehmensphilosophie’* und kommt damit dem, was als Idee oder Leitidee beschrieben worden ist, sehr nahe. Er zählt eine Reihe von Philosophien auf, mit denen Unternehmen erfolgreich gewesen sind. Ihnen

allen ist gemeinsam, dass sie zwar visionäre Inhalte aufwiesen, aber gleichzeitig ein hohes Maß an bereichsspezifischer Konkretisierung. Interessant ist nun besonders, dass er diese Aufzählung auch auf die Universität ausdehnt. Er schreibt:

„Jede Organisation, sei sie nun wettbewerbsorientiert oder nicht, hat eine eigene Unternehmensphilosophie. Und tatsächlich ist eine solche Theorie, sofern sie deutlich, konzentriert und in sich stimmig ist, ein sehr wirkungsvolles Instrument. Im Jahr 1809 gründete der deutsche Staatsmann und Wissenschaftler Wilhelm v. Humboldt die Universität Berlin auf der Grundlage einer radikal neuen Theorie des Bildungssystems. Und mehr als ein Jahrhundert lang, ... bestimmte diese Theorie das deutsche Bildungswesen, insbesondere die Forschung und Lehrinhalte.“⁷⁵

Auf die verhängnisvolle Wirkung der Inflation von Leitbildern im Zusammenhang mit der ungeklärten Kompatibilität und Gewichtung hat sehr überzeugend Uwe Schimank hingewiesen. Er spricht von 'Zielwirrwarr', der u.a. den Evaluationen die Grundlage entziehe.⁷⁶

5.3 Psychische und soziale Mechanismen

Wenn ich die bisherigen Überlegungen über die Leitbildentwicklungen zusammenfasse, komme ich zu dem Ergebnis, dass Leitbilder eine Idee haben müssen und dass diese Idee die *Kraft der inneren Überzeugung* für viele der in den Einrichtungen arbeitenden Menschen haben muss. Sie muss individuell überzeugen und sozial wirksam sein. Man könnte sagen, sie müsse beflügeln und eine Funktion erfüllen. Leitbilder, wenn sie realistisch und anspruchsvoll konzipiert sind, verweisen auf ein Mikro-Makro-Verhältnis. Sie artikulieren ein Grundverhältnis des sozialen Gemeinwesens. Menschen müsse mit ihren persönlichen Interessen und Bedürfnissen in sozial aufwendige und zumutungsintensive Programme eingebunden werden.

In diesem Zusammenhang ist an die richtungweisende Leistung des französischen Juristen Maurice Hauriou zu erinnern, der eine *Theorie der Institution* Anfang des 20. Jahrhunderts entwickelt hat, die bis heute unter

⁷⁵ Peter F. DRUCKER (2000): *Die Kunst des Managements*, 24.

⁷⁶ Uwe SCHIMANK (2001): *Festgefahrene Gemischtwarenläden – Die deutschen Hochschulen als erfolgreich scheinende Organisationen*, 224ff, bes. 229.

dem Gesichtspunkt der Leitbilddiskussion Aufmerksamkeit verdient.⁷⁷ Hauriou spricht von ‚*Leitideen*‘ als einer zentralen Charakterisierung von Institutionen. Bei diesem Begriff kommen wir dem, was hier unter dem Namen Leitbild zur Debatte steht, wahrscheinlich am nächsten. Leitideen sind primäre *Entwürfe zur Gründung* von Einrichtungen, die namhaft gemachte Ziele verfolgen. Sie entfalten ihre Kraft indem sie zT. unbewusst Entscheidungen und Handlungen beeinflussen. Leitideen sind in uns wohnende Objekte einer eigenständigen Ontologie. Sie sind nicht vollständig präzisierbar, haben aber die Kraft, die Vorstellungen Vieler zu besetzen und zu leiten.

Es geht Hauriou um die Beschreibung eines bis dahin nicht bekannten sozialpsychologischen Mechanismus der *freiwilligen Gleichrichtung der Intentionen* Vieler zum Zweck der Erreichung eines großen, aber im Detail unbestimmten Ziels. Das ist die *Beschreibung von Motivation* wie sie besser kaum sein könnte. Hier ist erkannt, dass Motivation grundsätzlich zweiseitig organisiert sein muss. Motivation ist in ihrer stärksten Form leidenschaftliches Engagement, aber auch immer Dienst an einer individuellen Sache. Das ist für die Beschäftigung mit Institutionen wie der Universität eine zentrale Nachricht. Sie lautet: Eine Leitidee kann Menschen im Rahmen einer Institution mit ihren Interessen und ihrem Willen zu spezifischem Handeln binden.

Neben diesem Aspekt der Formung individueller Zielsetzung steht ein zweiter, nämlich derjenige einer Belegung von Handlungsprogrammen durch und in Handlungssubjekten in Form ihres ‚*Handlungswillens*‘. Hauriou bezeichnet das als ‚*Personifizierung*‘.⁷⁸ In dieser Sichtweise ist die erst in jüngerer Zeit in den Sozialwissenschaften entworfene Errungenschaft einer Mehrebenenanalyse bereits ausgeprägt. Programme des Handelns und die von ihnen gebildeten Einrichtungen aller Art leben von individuellen Entscheidungen und ihrer Aggregation. Diese Entscheidungen sind in gewisser Weise freiwillig und stellen gleichzeitig *Selbstverpflichtungen* dar. Sie sind subjektiv in jedem Einzelfall und objektiv in der Summe. Eben das macht Motivation aus, die stets aus Akten der Bindung und Selbstverpflichtung hervorgeht und gleichzeitig Bindung i.S. einer ungeheuren *Bündelung von Entscheidungen* innerhalb einer Population darstellt.⁷⁹

⁷⁷ Maurice HAURIOU (1965): *Die Theorie der Institution*.

⁷⁸ Maurice HAURIOU (1965): *Die Theorie der Institution*, 36, 54.

⁷⁹ Ebenda, 40, 42, 59.

Man versteht, dass in diesem Entwurf eine Absage an jeden sozialen Determinismus liegt. Das hat Hauriou selbst deutlich gesagt. Ein Kollektivbewusstsein wollte er nicht gelten lassen.⁸⁰ Seine Leitideen stellen sozialpsychologisch gesprochen ‚Valenzen‘, Attraktoren oder Visionen mit Aufforderungscharakter dar.⁸¹

5.4 Organisation und Institution

Die beschriebene Konzeption zum Verständnis von Leitidee und Leitbild als einer dualen Vorstellung von Bindung schließt ausdrücklich ein, dass *Ideen verschleißen* und ihre Bindungs- und Motivationskraftkraft verlieren können.⁸² In unserer aktuellen Universitätsdiskussion wird uns das vorgeführt. Was Idee und Leitideen ausmacht, scheint verloren gegangen zu sein. Leitbilder sind grundsätzlich organisations- nicht institutionsbezogen. Das bedarf der Erklärung. Wenn wir von Organisationen sprechen, problematisieren wir das Verhältnis von Individuum und dem entsprechenden sozialen System als ein quasi-technisches.

Organisationen stellen ein Regelwerk von Vorschriften, Rechts- und Verwaltungsnormen zur Erzielung eines Systems von Teilzielen und Zwecken auf dem Weg der Abstimmung von definierten Leistungen dar. Organisationen sind Programmsysteme für die Herstellung von Entscheidungen, Leistungen und Koordinationsgewinnen.

Das entscheidende Merkmal von Organisationen ist also ihr *Funktionszusammenhang*. Institutionen kommen natürlich ohne das Moment des Organisiertseins nicht aus, aber man muss den Mehrwert, den sie leisten, erkennen. Darauf stellt die folgende Definition ab:

Institutionen untersuchen die Spannung von Individuum und Gesellschaft unter dem Aspekt der Selbstverpflichtung der Subjekte wie unter demjenigen einer treuhänderischen Obhut für kollektive Interessen oder kollektive Güter durch die organisierten Einrichtungen.

⁸⁰ Ebenda, 44.

⁸¹ Manfred HENNEN (1990): *Soziale Motivation und paradoxe Handlungsfolgen*, 86.

⁸² Maurice HAURIUO (1965): *Die Theorie der Institution*, 63, 65.

Die *treuhänderische Funktion* ist für die Universität von Parsons/Platt entwickelt worden.⁸³ Sie ist m.E. aber ein Merkmal kollektiver Einrichtungen, die wir Institutionen nennen, schlechthin. Oberste Gerichtshöfe, Zentralbanken und Bildungseinrichtungen sollten selbstverständlich gut funktionieren, aber sie sollten mehr sein, wir erwarten von ihnen, dass sie mit unserem Vertrauen verantwortungsvoll umgehen und wichtige kollektive Interessen in ihnen gut aufgehoben sind. Rechtsnormen und Verordnungen haben eine wichtige beschränkende Funktion, Institutionen sollten von Ideen leben, von denen eine *mobilisierende Kraft* ausgehen kann. Nichts anderes meint der Begriff der Leitidee.⁸⁴

Wir sagen mit Recht, dass die Bedeutung der Institutionen zurückgegangen ist, während diejenige von Organisationen zur lebensbeherrschenden Macht in unseren Gesellschaften herangewachsen ist. Die großen alten Institutionen Religion, Nation, Familie haben ihre Autorität eingebüßt und unterliegen Auflösungsprozessen. Das ist hier nicht zu bewerten. Wir reden in unserem Zusammenhang über die Universität und andere Bildungseinrichtungen. Ist auch ihr Niedergang beschlossene Sache? Viele Leitbilder sind sicherlich zusammengebrochen. Dazu gehört die Bindung an die Nation, die Verpflichtung auf Leitwissenschaften wie die Philosophie oder die Selbstverständlichkeit eines sozialen Korpsgeistes.

Darüber darf aber nicht vergessen werden, dass andere zentrale Leitideen nach wie vor aktuell sein könnten. Auch wenn wir uns in der z.T. geistlosen Wiederholung von Begriffen der radikalen Veränderung wie z.B. ‚Globalisierung‘ gefallen, sollten wir nicht *anthropologische Konstanten* übersehen, von deren Wirksamkeit wir abhängen. Die Universität kann nämlich dann als eine äußerst *moderne Institution* angesehen werden, weil sie anders als die anderen Genannten die Weitergabe, Erzeugung und ständige Überarbeitung und Überwindung von Wissen als ihren Kernprozess anzusehen hat. Es sei an die Wissensgesellschaft als Kennzeichen der Moderne erinnert. Andere Leitgedanken schließen sich dem nahtlos an. Ich lasse noch einmal Eduard Spranger mit verschiedenen Ausschnitten zu Wort kommen:

„ ... der Universität ist wirklich die Berufsanpassung nicht das *höchste* Ziel, sondern ein geistiger Menschentypus, der sich in seinem Beruf mit Freiheit bewegt. Und nicht nur das deutsche System ruht gerade auf dem wohlwogenen Gedanken, daß .. keine hö-

⁸³ Talcott PARSONS und Gerald M. PLATT (1990): *Die amerikanische Universität*, 176ff.

⁸⁴ Maurice HAURIUO (1995): *Die Theorie der Institution*, 35f.

here Kulturleistung ohne ein noch so bescheidenes Maß von Teilhabe an dem inneren Leben und Fortschritt der Wissenschaft möglich ist. Jede Trennung von Berufsausbildung und wirklicher wissenschaftlicher Bildung bedeutet „einen Rückschritt“ für die Universität. ... Auf jedem anderen Wege liegt die Gefahr scholastischer Erstarrung. Die echte Wissenschaft ist nie lebensfremd gewesen. ... Wird der deutsche Wissenschaftsbetrieb wirklich zu einem Bündel von Berufskunden umgestaltet werden? ... Der Mangel, wenn man will die Schuld der deutschen Universität liegt darin, daß sie noch immer mit der Fiktion arbeitet, als ob dieses quantitative Wachstum qualitativ in ihr nichts geändert hätte. An diesem Widerspruch, an dieser Unwahrhaftigkeit krankt sie. ... Zunächst kommt es darauf an, diesen ... nicht genug beachteten Unterschied zwischen der Universität als Anstalt und den Fakultäten als ihren Lebenszentren zu begreifen. ... ihre wesentliche Aufgabe liegt in der Niveauekontrolle. [Dazu ist] ... ein persönlicher Austausch von Mensch zu Mensch ... unverzichtbar; nur das unter Mühen selbst erarbeitete ist ganz echt – das ist die Idee.“⁸⁵

Die Häufung der klugen Argumente fasziniert m.E. Wer evaluieren will, darf sich solchen Einschätzungen nicht entziehen; dieses Versäumnis bedeutete keineswegs nur ein Manko an Bewertung, es hieße vor allen Dingen auch, die Chance der Verständigung zu verspielen, denn die hier aufgeführten Gedanken sind in den Köpfen vieler – nicht aller – Wissenschaftler lebendig und es gibt leider augenblicklich wenig Möglichkeiten, dafür das Verständnis einer größeren Universitätsöffentlichkeit zu finden.

Zusammenfassend lassen sich die Elemente einer Leitidee oder Idee der Universität für ein heutiges Verständnis in den folgenden sieben Dimensionen erfassen. Sie alle sind sowohl historisch zu belegen, als auch gegenwärtig und systematisch zu rechtfertigen

1. Wissenschaft als Engagement und Aufbruchstimmung
2. Wissenschaft als Kulturideal und Träger einer Zivilgesellschaft
3. Wissenschaft und Erkenntnis als Selbstzweck bei fortgesetztem Wandel
4. Wissenschaft als anspruchsvolle Leistung

⁸⁵ Eduard SPRANGER (1973): *Das Wesen der deutschen Universität*, 124f., 144, 151f., 154.

5. Wissenschaft als Zusammenhang und Interdisziplinarität
6. Wissenschaft als Gemeinschaft und Dialog
7. Wissenschaft als Abgrenzung zu Politik und Berufspraxis

6. Zusammenfassung

Zum Schluss meiner Ausführungen möchte ich die wichtigsten Ergebnisse zusammenfassen. Das Verfahren der Evaluation definiert sich über die Analyse und Bewertung von Einrichtungen, Programmen, Innovationen unter dem Aspekt ihrer Wirksamkeit und Qualität. Ihr Ziel besteht in der Situations- und Problemdiagnose sowie der Mobilisierung von Aktivität zur Reorganisation. Evaluation erfolgt unter der Zugrundelegung von Maßstäben, die aus dem Evaluationsfeld oder von außen stammen können. Diese Maßstäbe können selbst organisationsintern höchst widersprüchlich sein und ein großer Teil der Evaluationsarbeit besteht oft in einem Vergleich der Maßstäbe.

Auch der Qualitätsbegriff orientiert sich an Maßstäben. Qualität besteht in der Übereinstimmung bzw. Abweichung zwischen definierten Zielen einerseits und einer messbaren oder beschreibbaren Wirklichkeit andererseits. Qualität ist ein Maß der Kompatibilität von Zielen und Teilzielen mit ihrer Verwirklichung. Die angestrebten Reorganisationsprozesse machen Qualität zum Thema, weil für alle Formen sozialer Systeme gilt, dass Leistungen abnutzen und die Aufrechterhaltung von Qualität immer wieder besonderer Energiezufuhr und Anstrengungen bedarf. In Organisationen sind solche Anstrengungen besonders notwendig wegen der hohen Komplexität und des hohen Grades an Künstlichkeit der Kooperationsprogramme. Es gelingt selten, die interne Reorganisation zu einem selbstverständlichen Geschäft zu machen.

Evaluation soll deshalb Prozesse der außerordentlichen, außerplanmäßigen, extern veranlassten Introspektion der Qualität von Leistungen einleiten. Evaluation soll und kann nicht trivial kontrollieren, aber sie kann Selbstkontrolle über nicht-triviale Aktivierung veranlassen. Diese bietet sich besonders in Feldern an, auf denen Standardinitiativen und Verantwortungsbereitschaften brach liegen. Das ist überall dort der Fall, wo ein hoher Grad an Autonomie von Wissen unterschiedlichster Art zu Steuerungsresistenzen führen muss, so auch im Hochschulwesen. Hier ist nicht von ungefähr die Evaluation als ein Instrument eines ergänzenden, subsidiären Steuerungsinstrumentes eingeführt worden, als eine Hilfestellung oder ein Behelf für andernfalls missglückende Normalsteuerung. Solche

begleitenden, modifizierenden Interventionen werden als formative Evaluation bezeichnet.

Die zu behebenden Defizite bestehen in Qualitätsmängeln, die auf unzureichender Koordination, fehlenden Absprachen und ganz allgemein auf ausbleibender Verständigung beruhen; diese Mängel bezeichne ich als ein soziales Transaktionsdefizit. Transaktionen sind Leistungen von Person zu Person, von Kollektiv zu Kollektiv und soziale Leistungen. Transaktionen stellen Aushandlungs- und Übertragungsleistungen dar, die dort, wo sie nicht als selbstverständlich eingeführt und gesichert werden können, knapp werden und intern als erhebliche Zumutung angesehen werden. Von der Abstimmung dieser unterschiedlichen Leistungen und deren Zusammenführung hängt aber die Effektivität von Organisationen ab. Gerade in der Evaluationsarbeit muss man sich darauf einstellen, wie schwierig es ist, selbst in einer mittleren Organisation getrennte Wissenswelten zusammenzuführen und kompatibel zu machen. Es wird für diese Arbeit ein Wissen zweiter Ordnung benötigt, das im Evaluationsverfahren als Systematisierung für das Synthetisieren von organisationsinternem Teilwissen erzeugt werden soll.

Für diese Aufgabe stoßen standardisierte Verfahren wie fragebogengestützte Selbsteinschätzungen sehr schnell an ihre Grenzen. Evaluation ist angewandte Theorie und zwar besonders Handlungs- und Motivationstheorie im praktischen Interventionsvollzug. Theoretische Grundlagen der Evaluation müssen in verschiedener Weise und an verschiedenen Stellen unterschiedlicher Wissenschaften gewonnen werden. Evaluation beruht eher auf theoriebasierten Konzepten als auf strenger kausal-experimenteller Evaluationsforschung. Evaluation ist nach meinem Verständnis als wissenschaftlich gestützte Evaluation aufzufassen. Freilich gibt es für die Evaluationsverfahren wichtige und zentrale Anknüpfungspunkte an bestimmte sozialwissenschaftliche Konzepte. Dazu gehören *Grundlagen der Handlungstheorie*, darunter besonders solche, die einer Rational-Choice-basierten *Motivationstheorie* verpflichtet sind. Evaluation hat es überwiegend mit Phänomenen zu tun, die Albert O. Hirschmann mit ‚exit‘, ‚voice‘ und ‚loyalty‘ bezeichnet hat.⁸⁶

Der enge Zusammenhang von Evaluation und Motivation ergibt sich aus der Tatsache, dass die beschriebenen Transaktionsleistungen darin bestehen, dass durch die Aktivierung von Wissen Verantwortung und Motivation erzeugt werden. Motivation besteht immer in einem Mehrwert an

⁸⁶ Albert O. HIRSCHMANN (1974): *Abwanderung und Widerspruch*.

Handlungsbereitschaft und Belohnungsaufschub. Von den Sozialwissenschaften sind Formen des sozialen Engagements weitaus weniger untersucht worden als solche der Demotivation. Hierzu gehört die Thematik der öffentlichen Güter, des Trittbrettfahrers, des Gefangenendilemmas, der Verantwortungsdiffusion und des Shirking. Diese Erkenntnisse können für die Evaluation umgesetzt werden; wichtiger ist jedoch ihre Ergänzung durch die Entwicklung der Bedingungen für ihre Gegenbilder. Evaluation muss darauf bauen Empathie für den Umgang mit Perspektivenvielfalt zu erzeugen.

Um motivieren zu können, ist Evaluation darauf angewiesen bei den Organisationsverantwortlichen Wissen, Sprach- und Artikulationsfähigkeit sowie einen hohen Grad an Reflexionsvermögen zu erzeugen. Unsere Welten in und außerhalb von Organisationen sind sprachlich verfasst; wenn sie reorganisiert werden sollen, müssen die Menschen in Organisationen kontrolliert zur Sprache gebracht werden. Dieses Verfahren der Explizität beruht zunächst auf der beschriebenen Eindringlichkeit der Befragung, die von der Sprache zum Argument führen soll. Die Sprache ist vielleicht das entscheidende Medium der Rechenschaftslegung der Beteiligten von Organisationsprozessen.

Dazu sind die Bedingungen für einen vorbehaltlosen Gedankenaustausch durch entsprechende Befragungs- und Darstellungstechniken herzustellen. Sie sollen Kommunikationsgefälle überwinden um dem nahe zu kommen, was als Dialogsituation beschrieben werden kann. Dialogisch gewonnene und fixierte Diagnosen dieser Art beruhen zweifellos auf Informationen und Daten, aber sie benötigen ebenso die interne Anerkennung, das interne Eingeständnis, das auch sprachlich aus der Mauer des Schweigens, Leugnens und der Abwehr gelöst werden muss.

Evaluation bedeutet für die einzelnen Organisationsbeteiligten ausreichende Motivation zur Überschreitung der eigenen engen Perspektiven und Interessen zugunsten des Organisationskollektivs. Das bedeutet die Rücknahme von Defektion und die allmähliche Rückkehr zur Kooperation. Die spieltheoretischen Untersuchungen zu diesem Thema weisen interessante Entsprechungen zur Programmatik von Evaluationen auf.

Für die Zukunft der Evaluation sind neben der stärkeren Systematisierung ihrer Grundkonzeption weitere Veränderungen zu erwarten. Dazu gehört ihre stärkere Ausrichtung am Konzept der Beratung. Wahrscheinlich wird der Interventionsbedarf sich gesteigert auf Brennpunkte von Problemfeldern in Organisationen und Einrichtungen konzentrieren. Evaluationen

werden sich weiterhin stärker auf die Vernetzungsfelder – ich habe in diesem Zusammenhang den Begriff ‚Ausstrahlung der Evaluation‘ gewählt - einlassen müssen um weitere Akteure mit einzubeziehen.

Schließlich muss der Leitbild-Diskussion künftig mehr kritische Aufmerksamkeit geschenkt werden. Hierzu gibt es intelligentere Konzepte als die, welchen die Öffentlichkeit durch die einseitige Konzentration auf organisationsunspezifische Ökonomievorstellungen allenthalben ausgesetzt ist.

Evaluation – Neues Steuerungsmodell – Wettbewerb

Ulrich Druwe

Der Beginn: Das Mainzer Modell

Mit der Novellierung des Universitätsgesetzes von Rheinland-Pfalz 1995 wurden die Fachbereiche verpflichtet, jährlich über ihre Lehre zu berichten. Die Entwicklung des Lehrberichtsverfahrens führte an der Johannes Gutenberg-Universität Mainz zu der Erkenntnis, dass ein offensiver Umgang mit den in Studium und Lehre vorhandenen Problemen notwendig ist, der nur durch ein spezifisches, auf diesen Bereich ausgerichtetes Evaluationsverfahren erfolgen kann.

In Anlehnung an das Konzept der responsiven Evaluation wurde von den Soziologen Prof. Dr. Manfred Hennen, Elisabeth Springer, M.A. und Uwe Schmidt, M.A. sowie dem Leiter der Fachabteilung „Studium und Lehre“ an der Universität Mainz, Dr. Bernhard Einig, ein zweistufiges Evaluationsverfahren entwickelt, das mittlerweile als „Mainzer Modell“ bundesweit diskutiert wird.

Das Verfahren läuft wie folgt ab:

- Im ersten Schritt wird durch das zu evaluierende Fach eine „Selbstdiagnose“ verfasst und in einem Bericht schriftlich niedergelegt. Im Mittelpunkt stehen dabei durch Externe moderierte und in dem Prozess der Selbstdiagnose kumulativ verwendete Gruppengespräche mit Studierenden des Grund- und Hauptstudiums, Angehörigen des akademischen Mittelbaus und allen Professor(inn)en. Die Moderation lenkt die Gespräche immer wieder auf entscheidende Probleme im Fach, so dass der vom Fach unter Beratung der Moderatoren erstellte interne Bericht diese dann in erfreulicher Deutlichkeit anspricht.
- Im zweiten Schritt wird das Fach extern evaluiert. Die Gutachter werden von Vertretern des Faches selbst vorgeschlagen, jedoch von der Hochschulleitung in Absprache mit den Moderatoren eingeladen. Die Gutachter führen ihrerseits, auf der Basis des internen Berichts,

durch die Moderatoren vorbereitete und in deren Anwesenheit, ausführliche Gespräche mit den Angehörigen der verschiedenen Gruppen; bei den Studierenden und den akademischen Mitarbeitern ist die Auswahl der Gesprächsteilnehmer zufällig. Außerdem werden Studien- und Prüfungsordnungen sowie die Infrastruktur (Räume, Labore, Bibliotheken etc.) des Faches in Augenschein genommen. Über ihre Erkenntnisse verfassen die Peers ein externes Gutachten, welches sie sowohl dem Fach als auch den Moderatoren zur Verfügung stellen.

- Der interne Selbstbericht sowie das externe Gutachten stellen die Basis für die von ZQ formulierten Empfehlungen dar, die wiederum die Grundlage der Zielvereinbarung bilden. Sowohl diese Empfehlungen als auch der von ZQ erstellte erste Entwurf der Zielvereinbarung werden ausführlich mit dem Fach diskutiert.

Mit dieser Vorgehensweise können wichtige Einwände, die häufig gegen die Evaluation erhoben werden, entkräftet werden.

1. Das Mainzer Modell basiert auf modernen theoretischen und methodischen Standards der Evaluationsforschung; entsprechende Neuerungen werden permanent reflektiert und in das Verfahren integriert.
2. Das Verfahren ist freiwillig, d.h. es setzt bei den Interessen der Fachvertreter selbst an.
3. Die neutralen Moderatoren konzentrieren die Debatte der Fachvertreter und der externen Gutachter auf wesentliche Probleme in Studium und Lehre, die dadurch oft erstmals offen zwischen den Beteiligten diskutiert werden.
4. Die oft formulierte Kritik, die Peers würden sich ihrer Aufgabe nicht ernsthaft genug stellen, wird durch das Mainzer Modell gegenstandslos; hier werden die Gutachter von den Moderatoren durch das Verfahren begleitet und explizit auf Problemlagen aufmerksam gemacht.

Das Mainzer Evaluationsverfahren wurde in den Fächern und Fachbereichen sehr schnell äußerst positiv aufgenommen. Dies führte dazu, dass der Vizepräsident für Studium und Lehre 1999 den Senat der Johannes Gutenberg-Universität Mainz dafür gewann, das Zentrum für Qualitätssicherung und -entwicklung (ZQ) als wissenschaftliche Betriebseinheit zu

gründen. Die Koordination der Arbeit erfolgt in Absprache mit einem Senatsausschuss.

Seit der Evaluation in den Fachbereichen 21 – Biologie und 22 – Geowissenschaften,¹ bei denen auf deren eigenen Wunsch das Mainzer Modell erstmals auch explizit zur Evaluation der Forschungsleistungen angewendet wurde, wird in das Evaluationsverfahren generell eine Forschungsbeurteilung durch die externen Gutachter integriert. Die Plausibilität ergibt sich aus der für die Universität notwendigen Einheit von Forschung und Lehre: Nur Forschung, die nach neuem Wissen strebt, ermöglicht überhaupt Lehre auf höchstem Niveau. Gute Forschung ist folglich eine notwendige, wenn auch keine hinreichende Bedingung für gute Lehre.

Die Ergebnisse des Mainzer Modells können sich sehen lassen. So gelang es in allen Fällen einen Konsens über

- wichtige Ziele und Inhalte des Studiums,
- bislang diffuse Leistungsstandards und -erwartungen,
- eine notwendige Verbesserung der Betreuung der Studierenden,
- Verbesserungen in der Organisation des Studiums - insbesondere die Verzahnung der verschiedenen Fachgebiete – und den Prüfungen sowie
- die Qualität der Forschungsleistung einzelner Bereiche herzu stellen.

Besonders positiv ist es zu werten, dass die bislang individualisierte Verantwortung für Studium, Lehre und Forschung durch das Verfahren nun als Gesamtverantwortung der Kolleg(inn)en des Faches bzw. des Fachbereichs wahrgenommen wird.

Zusammenfassend kann man daher für die Johannes Gutenberg-Universität Mainz feststellen, dass es für sie eine Selbstverständlichkeit ist, über die Qualität des Studiums, der Lehre und der Forschung regelmäßig zu reflektieren und darüber auch zu berichten. Dies erfordert ein vernünftiges Verfahren der Evaluation, wie sie das Mainzer Modell darstellt.

¹ Vgl. zu den durchgeführten Evaluationen an der Johannes Gutenberg-Universität ausführlicher den Aufsatz von Elisabeth SPRINGER in diesem Band.

Die Gegenwart: Evaluation und Neues Steuerungsmodell

Evaluation im Sinne der Selbstreflexion zu betreiben rechtfertigt allein nicht den hohen Aufwand (Personal- und Sachmittel, Zeit). Evaluationen müssen auch zu Konsequenzen führen.

Um dies zu erreichen, wird das Mainzer Modell der Evaluation in Lehre (und Forschung) gegenwärtig zu einem Konzept der Qualitätssicherung als integrativer Bestandteil im Rahmen einer konsequenten Neuausrichtung der Johannes Gutenberg-Universität Mainz weiterentwickelt. Dieser Prozess der Reorganisation hat 1999 begonnen und firmiert unter dem Namen „Neues Steuerungsmodell“.

Hintergrund dieser Entwicklung waren einerseits Veränderungen in der europäischen bzw. internationalen Hochschulpolitik und andererseits die neue Hochschulpolitik der rheinland-pfälzischen Landesregierung.

Die Denationalisierung der Hochschulpolitik führte die Mitgliedsstaaten der Europäischen Union dazu, im Maastrichter Vertrag u.a. Grundzüge einer europäischen Hochschulpolitik festzuschreiben, was im sog. „Bologna-Prozess“ von 1999 dann seine konkretere Beschreibung fand. Demnach soll bis 2010 ein einheitlicher europäischer Hochschulraum gebildet werden.

Als allgemeines Ziel haben über 300 europäische Universitäten und die wichtigsten Organisationen, die sie vertreten, im März 2001 in Salamanca Folgendes formuliert:

„Als autonome und verantwortliche Rechts-, Bildungs- und Sozialeinheiten bekennen ... sich (Universitäten, U.D.) zu den Grundsätzen der Magna Charta Universitatum von 1988 und vor allem zum Prinzip der akademischen Freiheit. Die Universitäten müssen die Möglichkeit haben, ihre Strategien festzulegen, ihre Prioritäten für Forschung und Lehre zu setzen, ihre Mittel zuzuteilen, ihre Studienpläne zu erstellen und ihre Kriterien für die Aufnahme von Dozenten und Studenten zu bestimmen. Die europäischen Hochschulen nehmen die Herausforderung einer Tätigkeit in einem wettbewerbsorientierten Umfeld im eigenen Land, in Europa und in der ganzen Welt an, sie müssen jedoch zu diesem Zweck über die notwendige Freiheit des Managements, ... unterstützende Regulierungen und eine faire Finanzierung verfügen.“²

² CONVENTION OF EUROPEAN HIGHER EDUCATION INSTITUTIONS (2001): <http://147.83.2.29/salamanca2001/>

Die konkrete Ausgestaltung des europäischen Bildungsraumes bezieht sich auf:

- die Anpassung der Studienstruktur (akademischer Grad, Studienpläne),
- eine Mobilitätsförderung für Dozenten und Studierende,
- die gegenseitige Anerkennung von akademischen Graden und Studienplänen,
- die Internationalisierung von Curricula bis hin zu den Hochschulen selbst,
- die Entwicklung von gegenseitig akzeptierbaren Mechanismen für die Bewertung, Gewährleistung und Zertifizierung der Qualität und
- den Abbau der gegenwärtigen Überregulierung im Hochschulbereich.³

Letzteres ist seit Mitte der 90er Jahre auch Ziel der rheinland-pfälzischen Landesregierung. Daher hat Wissenschaftsminister Zöllner das Muster der hochschulpolitischen Steuerung verändert: Dominierte bislang die kameralistische Inputsteuerung, so wird nun den Hochschulen des Landes ein Personal- und Sachmittelbudget zugeteilt, das über Modelle berechnet wird, welche sich an Belastungen (Studierendenzahl, Prüfungen etc.) und Leistungen (z.B. eingeworbene Drittmittel, Habilitationen, Sonderforschungsbereiche) orientieren und zu 75% (Personal) bzw. 100% auf das folgende Haushaltsjahr übertragbar sind. Ab 2003 werden neben dem Personal und den Sachmitteln auch die Flächen und Räume der Landeshochschulen über bestimmte Parameter (z.B. Studienplatzzielzahl, Flächenrichtwerte) budgetiert.

Das durch Globalbudgets und outputorientierte Steuerung gewonnene Mehr an Autonomie der Hochschulen verlangt nun von diesen, dass sie nach eigenen Zielen und Kriterien Personal, Sachmittel und in Zukunft auch Flächen unterverteilen, also eigenverantwortlich, effektiv und effizient steuern.

Die Johannes Gutenberg-Universität Mainz hat auf die skizzierten Herausforderungen reagiert. 1999 haben der Kanzler der Johannes Gutenberg-Universität Götz Scholz, Vizepräsident Prof. Ulrich Druwe sowie Claudia Knoth-Weiler und Wiebke Koerlin als Projektleiterinnen in Ab-

³ vgl. <http://www.unige.ch/cre/activities/Bologne> 1999.

sprache mit dem Universitätspräsidenten Prof. Josef Reiter ein Reformkonzept entwickelt, das fünf Ziele verfolgt:

- a) **Universitätsziele** – Die Selbststeuerung einer Universität ist ohne eine Verständigung auf grundlegende Ziele nicht möglich. Das Projekt „Leitbild der Johannes Gutenberg-Universität Mainz“ soll diesen Konsens der Mitglieder der Universität ermöglichen.
- b) **Strukturreform** – Eigenverantwortliche Steuerung ist nur möglich auf der Basis leistungsfähiger Organisationseinheiten. Dies verlangt beispielsweise eine Neugliederung der Fachbereiche oder die Reorganisation der Bibliotheksstruktur, aber auch die Entwicklung neuer Servicebereiche wie etwa die Einrichtung eines Studienbüros oder von Prüfungssätern.
- c) **Kommunikation** – Die neuen Aufgaben der Universität erfordern deutliche Verbesserungen der internen und externen Kommunikation. Um nachvollziehbare Entscheidungen treffen zu können, müssen entsprechende Daten schneller bereitgestellt werden. Außerdem gilt es, die Universitätsöffentlichkeit in notwendige Maßnahmen einzubinden, um Konflikte konstruktiv lösen zu können. Schließlich ist die Öffentlichkeitsarbeit und das Marketing der Universität auszubauen. Daher sind das edv-gestützte Berichtswesen, der Internetauftritt der Universität und Strategien zum Fundraising und Sponsoring zu konzeptionieren.
- d) **Mitarbeitermotivation** – Zentral für eine Universität sind ihre leistungsbereiten und motivierten Mitarbeiter; dies gilt für die Wissenschaftler ebenso wie für die Mitarbeiter in den wissenschaftsstützenden Bereichen. Die Erhöhung der Motivation sowie die kontinuierliche Weiterqualifikation sind daher wichtige Elemente des Neuen Steuerungsmodells. Entsprechend werden in verschiedenen Teilprojekten Maßnahmen der Personalentwicklung, Zulagen und Prämien, Modelle der Arbeitszeitflexibilisierung ebenso erörtert wie Mitarbeitergespräche und Verhalten am Arbeitsplatz.
- e) **Steuerungsinstrumente** – Die Umsetzung der Ziele der Universität verlangt adäquate Steuerungsinstrumente. In einem ersten Schritt hat die Johannes Gutenberg-Universität in Anlehnung an die Landesmodelle Personalbemessungskonzept PBK und Mittelbemessungskonzept MBK universitätsinterne Verteilungsmodelle für das Personal und die Sachmittel (der Titelgruppe 71) ausgearbeitet. Gegenwärtig wird ein Modell für die Verteilung der Flächen/Räume (Flächenmanagement) entwickelt. In einem zweiten Schritt geht es hier aber

insbesondere um die Entwicklung und Umsetzung einer universitäts-spezifischen Kosten -und Leistungsrechnung.

Um diese Ziele zu erreichen wurden 1999 sechzehn Teilprojekte definiert; vier davon konnten schon erfolgreich beendet werden, neue Teilprojekte wurden ergänzt, aktuell werden die genannten Ziele in 22 Teilprojekten verfolgt; ca. 80 Mitarbeiter(innen) sind in die Arbeit eingebunden.

Unter dem Aspekt der Evaluation – der hier im Mittelpunkt steht – sind die genannten Ziele wie folgt zu interpretieren: Die Johannes Gutenberg-Universität versucht gegenwärtig im Rahmen des Neuen Steuerungsmodells ein umfassendes Konzept der Qualitätssicherung und -entwicklung aufzubauen.

Die Bestimmung dessen, was die Mitglieder der Johannes Gutenberg-Universität als „Qualitäten“ in Forschung, Lehre, Studium, Weiterbildung und Verwaltung erachten, ergibt sich aus dem Leitbild der Universität. Dieses stellt folglich den verbindlichen Rahmen von anzustrebenden Zielen und Werten dar. Erst in diesem Kontext sind strategische Entscheidungen, wie beispielsweise über das Profil der Universität, Schwerpunktsetzungen, Strukturreformen oder Ressourcenverteilung möglich.

Der herausragenden Bedeutung des Leitbildes als allgemein akzeptiertem Selbstverständnis der Universität wird durch das – partizipative und konsekutive – Verfahren seiner Entwicklung Rechnung getragen, welches 2002 abgeschlossen sein wird.⁴

Qualitätssicherung kann nur gelingen, wenn sie sich mit der Struktur und der Prozessorganisation der Johannes Gutenberg-Universität deckt. Nachdem sich die Mainzer Universität dafür ausgesprochen hat, die Ressourcenverantwortung auf die Fachbereiche und die zentralen Betriebseinheiten zu delegieren, muss auch die Qualitätssicherung und -entwicklung dezentral wahrgenommen werden. Entsprechend müssen die Einheiten der Universität, beispielsweise Fachbereiche, Institute etc. im Rahmen des Gesamtleitbildes ihre Ziele selbst bestimmen. Das Instrument der Vermittlung zwischen den Ebenen ist die Zielvereinbarung.

Die Hochschulleitung der Universität Mainz hat bislang mit drei Fachbereichen (19 - Chemie/Pharmazie, 21 - Biologie und 22 - Geowissenschaften) und sechs Fächern (Chemie, Pharmazie, Biologie, Anthropologie, Geowissenschaften und Geographie) auf der Basis des vorher durch-

⁴ Das Leitbild-Projekt wird geleitet von Dr. Bernhard Einig; wissenschaftliche Mitarbeiter sind Felicia Lauer, M.A. und Thomas Vogt, M.A.

laufenen Evaluationsverfahrens konkrete Zielvereinbarungen abgeschlossen.

Dabei hat sich die Hochschulleitung für eine anreizorientierte Variante der Zielvereinbarung entschieden, d.h. Maßnahmen, die dem Ausbau von Stärken dienen, aber auch solche, die Schwächen ausgleichen sollen, werden durch eine Anschubfinanzierung gefördert. Die Mittel hierfür werden aus einem Studienstrukturfond bereitgestellt, der 2000 eingerichtet wurde. Fachbereiche und Fächer erhalten damit für eine Laufzeit von max. fünf Jahren Planungssicherheit und Anreize für Veränderungen. Allerdings müssen sie sich dann einer erneuten (Kurz-)Evaluation unterziehen, um Rechenschaft über die Zielerreichung zu geben.

Damit ist der zweite Schritt der Qualitätssicherung angesprochen: Es reicht nicht aus, gut klingende Ziele zu formulieren, es muss auch geprüft werden, ob die Ziele von den Beteiligten erreicht wurden. Hierfür sind Evaluationsverfahren notwendig. Bislang beziehen sie sich jedoch allein auf Forschung und Lehre.

Sie müssen daher ergänzt und im Verfahren modifiziert werden,

1. um die Leistungsfähigkeit von universitären Supportstrukturen (Universitätsbibliothek, ZDV, Fremdsprachenzentrum, Elektronisches Medienzentrum, Allgemeiner Hochschulsport, Verwaltung, Werkstätten etc.) analysieren zu können und
2. sie müssen mit der Kosten- und Leistungsrechnung verbunden werden, z.B. durch Schaffung einer entsprechenden Datenbasis bzw. Gestaltung der Kostenträgerrechnung, damit Produkte mit maßgeblichen Leistungsbereichen übereinstimmen.

Schließlich gilt es, Ziele und ihre Umsetzung intern und insbesondere extern zu kommunizieren, um die Leistungsstärke der Johannes Gutenberg-Universität im Wettbewerb zu dokumentieren. Letzteres ist angesichts der veränderten Rahmenbedingungen eine zwingende Voraussetzung, um die in Zukunft immer wichtiger werdenden Mittel von Privaten akquirieren zu können. Spenden, Stiftungen und Public Private Partnership-Projekte, die drei Felder, auf die sich die Universität Mainz konzentriert, werden nur erfolgreich eingeworben werden können, wenn sich die Universität der Öffentlichkeit, ihren privaten Förderern und Partnern belegbar als innovativ und effizient präsentiert.

Zusammenfassend kann man also feststellen, dass ein solches Konzept der Qualitätssicherung und -entwicklung, mit entsprechend flexiblen Evalua-

tionsverfahren als Controllinginstrumenten, notwendiger Bestandteil der Selbststeuerung der Johannes Gutenberg-Universität Mainz ist.

Die Zukunft: Evaluation und Wettbewerb

Evaluation wurde bisher als Instrument der Selbstreflexion in Studium, Lehre und Forschung sowie der Qualitätssicherung bzw. -entwicklung und damit der Selbststeuerung der Johannes Gutenberg-Universität diskutiert.

In Zukunft werden Evaluationsverfahren für die Universität Mainz zusätzlich im Rahmen des Wettbewerbs zwischen

- a) Professoren (man denke an die aktuelle Dienstrechtsreform der Bundesregierung) und
- b) Studiengängen (hier ist auf das Akkreditierungsverfahren zu verweisen)

eine bedeutende Rolle spielen und spezifisch weiter entwickelt werden müssen.

a) Unter dem Motto „Die Reform sorgt für einen Modernisierungsschub an den Hochschulen“ hat das Bundeskabinett im Sommer 2001 das sog. Professorenbesoldungsreformgesetz verabschiedet, das Anfang 2002 in Kraft treten soll.⁵ Kern der Reform ist die leistungsorientierte Besoldung.

Neben ein Mindestgehalt von 7000 DM in der Gruppe W 2 bzw. 8500 in der Gruppe W 3 tritt ein variabler Gehaltsbestandteil, der sich aus der Bewertung von Leistungen in Lehre, Studienbetreuung, Forschung und Verwaltung zusammensetzt. Eine Obergrenze für diese variablen Gehaltsbestandteile gibt es nicht, die Hochschulen müssen es allerdings aus ihrem Gesamtpersonalbudget finanzieren, welches wiederum Teil eines Landespersonalbudgets ist, und das jährlich nur um max. 2% steigen darf.

Unabhängig von der Bewertung der Besoldungsreform ist es evident, dass die variablen Gehaltsbestandteile den einzelnen auf Dauer nur erhalten bleiben, wenn ihre Leistungen in Lehre und Forschung regelmäßig evaluiert werden.

⁵ vgl. Pressedienst, BMBF vom 30.5.2001, ftp://ftp.bmbf.de/010530_gesetzbmiprobe-soldung.pdf

Die einzelnen Hochschulen sind nun damit konfrontiert, zur Bewertung der Leistungen ihrer Professoren Standards und Verfahren zu entwickeln, die selbstverständlich konsensfähig aber auch justitiabel sein müssen; dies wird eine immense Herausforderung sein.

b) Die Globalisierung der Wirtschaft führt zu einem internationalen Arbeitsmarkt für Akademiker(innen). Daher müssen sich die Studiengänge und -abschlüsse der internationalen Konkurrenz stellen. Diese Einsicht führte dazu, dass die Kultusministerkonferenz die Einführung von Studiengängen mit den international bekannten Abschlüssen Bachelor (B.A.) und Master (M.A.) ermöglichte.⁶ Zur Sicherung der Qualität dieser Studiengänge wollten die Kultusminister von einer Prüfung dieser Studiengänge via Rahmenprüfungsordnung Abstand nehmen; als Ersatz wurde das Akkreditierungsverfahren⁷ implementiert.

Die Akkreditierung stellt eine simple Gutachter-Evaluation fachlich-inhaltlicher Mindeststandards, die Überprüfung der Berufsrelevanz eines akademischen Abschlusses sowie der notwendigen Ressourcen dar.

Durchgeführt wird das Akkreditierungsverfahren auf Antrag einer Hochschule durch eine Akkreditierungsagentur, die selbst durch den Akkreditierungsrat, in dem Repräsentanten der Länder, der Hochschulen, der Studierenden sowie der Wirtschaft vertreten sind, akkreditiert worden ist. Die Agentur benennt für das Verfahren wenigstens drei Fachgutachter aus der Gruppe der Professoren, der Studierenden sowie Vertretern der Wirtschaft. Diese prüfen die eingereichten Unterlagen (Beschreibung des Studienprogramms, Prüfungsordnung, Aussagen über die eingesetzten Ressourcen etc.) und führen eine Begehung vor Ort durch. Auf der Basis ihres Gutachtens entscheidet die Agentur über die Akkreditierung, die auch mit Auflagen erfolgen kann. Nach drei bis fünf Jahren wird eine Reakkreditierung durchgeführt.

Die bisherigen Erfahrungen der Johannes Gutenberg-Universität – sie beziehen sich auf den Masterstudiengang Chemistry of Materials des Fachbereichs Chemie/Pharmazie und den Bachelorstudiengang Informatik des Fachbereichs Mathematik und Informatik – mit der Akkreditierung sind nicht überzeugend. In beiden Fällen wurden sicherlich nicht nur

⁶ Seit 1998 sieht das HRG die Möglichkeit zur Einrichtung von B.A.- und M.A.- Studiengängen vor.

⁷ Vgl. www.akkreditierungsrat.de

Mindeststandards überprüft, sondern Gutachter und Agentur versuchten, den Fächern eigene Inhalte und Ressourcenvorstellungen nahe zu bringen.

Die bisherige Form der Akkreditierung ist auch sonst keinesfalls unumstritten. Dies betrifft sowohl das Verfahren als auch die Kompetenzen der Akteure.

Das Verfahren der Akkreditierung bzw. der Reakkreditierung bezieht sich bislang auf Studiengänge, d.h. auf den Bereich Studium und Lehre. Nachdem B.A.- und M.A.-Studiengänge auch an Fachhochschulen angeboten werden, muss bei der Akkreditierung auch die Forschung berücksichtigt werden, denn gerade die Verbindung von Forschung und Lehre ist konstitutiv für universitäre Studiengänge.

Die Verfahrensdurchführung ist grundsätzlich mit den Kritikpunkten behaftet, die das einfache Evaluationsverfahren in Misskredit gebracht haben: Externe Gutachter, zumal wenn sie Berufsverbände repräsentieren, oktroyieren einem Fach ihre Vorstellungen und eine ernsthafte Diskussion über die Mindestqualität eines Studienganges ist nicht möglich.

Hinzu kommt, dass es zahlreichen Akteuren in den Akkreditierungsagenturen, aber auch im Akkreditierungsrat an Erfahrungen mit Evaluationen, insbesondere mit den methodischen aber auch psychologischen Problemen, mangelt. Von daher ist es erforderlich, dieses Verfahren und die damit verbundenen Strukturen sorgfältig zu überdenken.

Angesichts dieser Erfahrungen sowie der Tatsache, dass die Verfahren gegenwärtig viel zu zeitaufwendig (es gibt zu wenig Agenturen) und zu teuer sind – eine Akkreditierung kostet zwischen 25.000 und 30.000 DM – stellt sich für die Johannes Gutenberg-Universität die Frage, in welcher Form sie sich zukünftig an einer Akkreditierungsagentur, einem Akkreditierungsverbund und/oder einem entsprechenden Evaluationsnetzwerk beteiligen wird. Dass die Universität Mainz in einer Agentur mitarbeiten muss, steht außer Frage, weil nur so ihre Interessen und ihre Evaluationserfahrungen, repräsentiert durch das ZQ, angemessen angebunden werden können.

Mehrere Modelle stehen gegenwärtig zur Debatte, so eine von der Landesregierung Rheinland-Pfalz favorisierte Akkreditierungs- und Evaluationsagentur mit Nordrhein-Westfalen und den Niederlanden, AQUA genannt, oder ein Verbund mit Universitäten aus Hessen, Thüringen und dem Saarland. Hier werden in den nächsten Monaten intensive Beratungen stattfinden, um für die Johannes Gutenberg-Universität Mainz aber

auch für das Land zu einer Lösung kommen zu können, die einer Differenzierung der Studienstruktur förderlich ist.

Evaluation 2000 des Fachbereichs Biologie – ein Rückblick

Harald Paulsen und Jürgen Markl

Der Fachbereich Biologie der Johannes-Gutenberg-Universität hat sich im Jahr 2000 einer internen und externen Evaluation der Forschung und Lehre unterzogen. Es handelte sich im Prinzip um eine freiwillige Begutachtung, allerdings war ein entsprechender Wunsch vom rheinland-pfälzischen Ministerium für Bildung, Wissenschaft und Weiterbildung an den Fachbereich herangetragen worden. Anlass hierzu waren Pläne für ein „Zentrum für Biowissenschaften“ an der Universität Mainz; das Begutachtungsergebnis sollte zur Bewertung der gegenwärtigen Struktur des Fachbereichs sowie als Empfehlung für die Strukturentwicklung im Rahmen eines „Zentrums für Biowissenschaften“ herangezogen werden. Der vorläufige Zeitplan für die Baumaßnahme bestimmte auch den relativ straffen Zeitplan für das Evaluationsverfahren, der als Zeitpunkt für den Abschluss des gesamten Verfahrens Ende des Sommersemesters 2000 vorsah. Dieser Zeitplan konnte im wesentlichen eingehalten werden. Nach vorbereitenden Gesprächen zwischen dem Fachbereich und dem Zentrum für Qualitätssicherung und –entwicklung über den Ablauf des Verfahrens sowie Verhandlungen mit der Universitätsleitung und dem Ministerium über seine Finanzierung in der zweiten Jahreshälfte 1998 wurden 1999 die Vorbereitungen für die interne Evaluation getroffen. Eine wichtige Informationsquelle für diese interne Evaluation waren Gespräche mit Gruppen von Fachbereichsmitgliedern, die von MitarbeiterInnen des Zentrums für Qualitätssicherung und –entwicklung vom Mai 1999 bis Februar 2000 durchgeführt wurden. Der 92-seitige interne Evaluationsbericht wurde vom Fachbereichsrat am 3. Mai 2000 einstimmig verabschiedet. Am 5./6. Juni 2000 folgte dann die Begutachtung des

Fachbereichs durch eine neunköpfige Gutachtergruppe. Der Bericht dieser Gutachtergruppe lag am 23. Oktober vor.

Anliegen dieses Rückblicks ist nicht, Ergebnisse und Empfehlungen aus dem Evaluationsverfahren wiederzugeben. Vielmehr soll zusammengefasst werden, welcher Nutzen für den Fachbereich aus der Evaluation bereits ein knappes Jahr nach Ablauf des Verfahrens sichtbar wird, welches die im Rückblick wichtigsten Voraussetzungen waren für eine erfolgreiche Durchführung der Evaluation und was wir aus jetziger Sicht anders planen und handhaben würden.

Was hat unserem Fachbereich die Evaluation gebracht?

Ein Hauptertrag aus der Evaluation ist ein enormer Informationsgewinn über den eigenen Fachbereich, seine Strukturen und Leistungen. Daten etwa über die Zahl der Stellen und andere Ressourcen, die einzelnen ProfessorInnen zugeordnet sind, über Drittmiteleinwerbungen und Publikationsleistungen, über Unterrichtsbeiträge der einzelnen Institute oder die Entwicklung der Studiendauer verschiedener biologischer Studiengänge in den letzten Semestern waren in irgendeiner Form auch vorher vorhanden, aber nicht in der zusammengefassten, strukturierten und leicht zugänglichen Form, in der sie nach der internen Evaluation vorliegen. Diese Basis an Zahlenmaterial ist ungemein hilfreich bei Überlegungen zur weiteren Strukturentwicklung für den Fachbereich.

Ähnliches wie für dieses Datenmaterial gilt für Ansätze und Ideen zur Strukturverbesserung innerhalb des Fachbereichs, die in großer Zahl ebenfalls vor der Evaluation bereits angedacht worden waren, für deren Verwirklichung die Evaluation jedoch die entscheidende Initialzündung lieferte. Das gilt für so verschiedene Maßnahmen wie die Umwandlung von Dauerstellen in Zeitstellen, die verstärkte Einbeziehung weiterer biologischer Fächer ins Grundstudium, die Verbesserung des Informationsangebots für Studierende durch eine neugestaltete „Homepage“ des Fachbereichs, unter anderem mit Inhaltsangaben einzelner Lehrveranstaltungen und ausführlicher Darstellung der Forschungsaktivitäten in den

biologischen Arbeitsgruppen, und einer von der Fachschaft initiierten und vom Fachbereich organisierten flächendeckenden Evaluation der Lehrveranstaltungen im Grundstudium. Dieser Innovationsschub stärkt das Selbstbewusstsein des Fachbereichs, wie auch das Erlebnis des Fachbereichs als gemeinsame, verbindende und veränderbare Organisation im und nach dem Evaluationsverfahren unter den Fachbereichsangehörigen durchaus konstruktive Kräfte freigesetzt hat.

Die Evaluation hat nicht nur die Innensicht der Fachbereichsangehörigen auf ihren Fachbereich verändert; auch nach außen hin hat das Evaluationsergebnis, besonders die positiven Stellungnahmen der externen Gutachter, das Bild des Fachbereichs verbessert. Wir haben den Eindruck, dass Gesprächspartner innerhalb und auch außerhalb der Universität, die die Mainzer Biologie eher in die Kategorie unmodern bis vorgestrig einzureihen geneigt waren, aufgrund der Begutachtung ihre Ansicht revidiert haben. Wir hoffen natürlich, dass diese veränderte Sichtweise auch ihren Niederschlag darin finden wird, dass die Baumaßnahme eines „Zentrums für Biowissenschaften“ – immerhin der Anlass des ganzen Begutachtungsverfahrens – in einer für den Fachbereich befriedigenden Form und Zeitspanne Gestalt annimmt.

Was waren wichtige Voraussetzungen für die Evaluation?

Die wichtigste Komponente war zweifellos der interne Evaluationsbericht, der mehrere Voraussetzungen erfüllen sollte: Er sollte alle notwendigen Primärinformationen enthalten, diese selbstkritisch und kompetent interpretieren, Lösungsvorschläge für erkannte Probleme bereithalten und schließlich kurzweilig und kurz genug sein, dass er auch gelesen wird. Das Erstellen dieses Berichtes wurde deshalb nicht „nach unten“ delegiert, sondern die beiden Verfasser waren über die letzten fünf Jahre hinweg an den wichtigen Entscheidungsprozessen im Fachbereich persönlich beteiligt. Nur so konnten wir die gesammelten und z.T. recht widersprüchlichen Daten adäquat bewerten. Nur indem wir uns die Arbeit hälftig geteilt haben war sie überhaupt im vorgegebenen Zeitraum zu bewältigen. Wichtig war auch, einerseits die Zuständigkeit klar festzule-

gen (Forschung bzw. Lehre), sich andererseits aber ständig zu koordinieren.

Ebenso wichtig war es, dem Ministerium bzw. der Universitätsleitung möglichst hochrangige Gutachter zu benennen und nicht etwa auf ein Gefälligkeitsgutachten zu bauen. Um den Evaluationsbericht im Fachbereich wirklich konsensfähig zu machen wurde der Fachbereichsrat temporär um mehrere Gäste erweitert mit dem Ziel, auch die nicht im FBR vertretenen Institute bzw. Abteilungen von vorne herein in die Diskussionen einzubeziehen.

Eine unabdingbare Voraussetzung für ein Gelingen des Evaluationsvorhabens war die Bereitschaft aller Fachbereichsangehörigen zur Offenheit. Dazu zählen die offenen Äußerungen und Stellungnahmen in den Evaluationsgesprächen ebenso wie die Bereitschaft der ProfessorInnen im Fachbereich, etwa ihre Drittmittelinwerbungen auch über Fachbereichs- und Universitätsgrenzen hinaus transparent zu machen oder sich der (durchaus angreifbaren) Quantifizierung ihrer Forschungsaktivität aufgrund von „Impactfaktoren“ zu stellen. Diese Bereitschaft, die Karten auf den Tisch zu legen, verhalf dem internen Evaluationsbericht zur Aussagekraft, was wiederum den externen Gutachtern eine Grundlage für ihre Bewertungen lieferte. Wichtig war auch, dass der Bericht im Fachbereichsrat einstimmig verabschiedet wurde, denn dadurch bilden seine Inhalte für alle künftigen Entscheidungen eine feste Grundlage.

Sowohl bei der internen als auch der externen Evaluation haben wir Hilfe bekommen. Der Fachbereich ist für die Unterstützung durch das Zentrum für Qualitätssicherung und –entwicklung außerordentlich dankbar. Dessen MitarbeiterInnen übernahmen nicht nur einen wichtigen Teil der Datenerhebung zur internen Evaluation, nämlich die Organisation, Durchführung und Dokumentation der Gruppengespräche mit Fachbereichsangehörigen. Auch die Vorgabe der Struktur für das ganze Verfahren, die auf Erfahrungen bei der Evaluation anderer Fachbereiche beruhte, und die Hilfe bei der Auswertung der Information, die in der internen Evaluation gesammelt worden war, waren äußerst hilfreich. Selbstverständlich haben wir zusätzlich Information eingeholt von zahlreichen KollegInnen an anderen

Universitäten, die ähnliche Begutachtungsverfahren durchlaufen hatten, um aus deren Erfahrungen zu lernen. Als nützlich erwies sich auch, dass einer von uns (J.M.) als Gutachter an der Evaluation mehrerer anderer biologischer Fakultäten teilgenommen hatte und daher ein solches Verfahren aus Gutachtersicht erlebt hatte.

An wichtiger Unterstützung sind weiterhin zu nennen ein finanzieller Zuschuss vom Ministerium und von der Universität, ohne den die Evaluation die finanzielle Kapazität des Fachbereichs gesprengt hätte, einschließlich der Finanzierung einer wissenschaftlichen Hilfskraft für mehrere Monate. Ohne die Mitwirkung dieser Hilfskraft hätten arbeitsaufwändige Teile des Evaluationsverfahrens, etwa die Befragung von AbsolventInnen des Biologiestudiums aus den letzten 5 Jahren zu ihren Erfahrungen im Studium im Rückblick oder die Auszählung aller Zitiierungen von Veröffentlichungen aus dem Fachbereich im Zeitraum von 6 Jahren, schlicht nicht durchgeführt werden können. Nicht zuletzt wollen wir als wichtige Voraussetzung für unsere Evaluation eine leistungsfähige Fachbereichsverwaltung nennen, die einen großen Teil des Zahlenmaterials über Jahre gesammelt hatte und in elektronischer Form bereithielt oder sehr rasch erstellen konnte. Schließlich erwies es sich noch als äußerst effektiv, zusätzlich zum Evaluationsbericht eine professionell gestaltete, illustrierte Fachbereichsbroschüre zu erstellen.

Was würden wir jetzt anders machen?

An der Grundstruktur des Evaluationsverfahrens würden wir aus jetziger Sicht wenig oder nichts ändern. Die externe Evaluation folgte der Begutachtungsform von größeren Forschungseinrichtungen wie Sonderforschungsbereichen durch die Deutsche Forschungsgemeinschaft, die sich über viele Jahre hinweg bewährt hat. Zu überlegen wäre allenfalls, die Größe der Gutachtergruppe etwas zu verringern – die Zahl von 9 Gutachtern hat sicher die internen Diskussionen dieser Kommission nicht vereinfacht und möglicherweise auch dazu beigetragen, dass das externe Gutachten erst mit einiger Zeitverzögerung erstellt werden konnte. Andererseits halten wir es im Hinblick auf die Akzeptanz der externen Begut-

achtung innerhalb des Fachbereichs für wichtig, dass sich jedes Institut in der Kommission fachkundig vertreten sieht.

Bei der internen Evaluation hat uns das sequenzielle Verfahren des Zentrums für Qualitätssicherung und -entwicklung bei der Durchführung der Gespräche mit Fachbereichsangehörigen überzeugt. Dadurch, dass bei den aufeinanderfolgenden Gesprächsrunden mit Studierenden, DoktorandInnen, nichtwissenschaftlichen und wissenschaftlichen MitarbeiterInnen und ProfessorInnen die Ergebnisse der jeweils vorangegangenen Diskussionen die Gesprächsgrundlage bildeten, richtete sich der Fokus erstaunlich zielgenau auf Problembereiche des Fachbereichs. Allerdings drehten sich die Gespräche dann gelegentlich auch um nicht mehr existente Probleme, etwa wenn Studierende höherer Semester Zustände im Grundstudium beklagten, die seit Jahren nicht mehr bestehen, und dies sich dann durch zahlreiche folgende Gesprächsrunden zog. Wir sehen den großen Wert einer vollkommen unvoreingenommenen Gesprächsleitung durch das Zentrum für Qualitätssicherung und -entwicklung; andererseits geht durch das genannte Problem viel wertvolle Gesprächszeit ins Leere. Daher regen wir an, dass die GesprächsleiterInnen des Zentrums für Qualitätssicherung und -entwicklung bereits im Verlauf der Gruppengespräche Rücksprache mit dem Fach halten, um Gesprächsergebnisse differenzierter bewerten zu können, und sehen uns darin auch mit den VertreterInnen des Zentrums für Qualitätssicherung und -entwicklung völlig einig.

Der mit der internen Evaluation verbundene Arbeitsaufwand war für die aktiv Beteiligten über ca. 6 Monate hinweg immens und auch davor und danach noch erheblich. Vor diesem Hintergrund erscheint uns die künftige Durchführung vergleichbar aufwändiger Aktionen auf einer regelmäßigen Basis, auch wenn dies im Abstand von mehreren Jahren passiert, als völlig unrealistisch; der Fachbereich kann sich eine so massive Bindung mehrerer zentraler Professuren über einen so langen Zeitraum schlicht nicht leisten. Eine Aktualisierung der Bestandsaufnahmen aus der jetzigen Begutachtung in Zeitabständen von ca. 5 Jahren erscheint uns dagegen sinnvoll und wünschenswert.

LITERATURVERZEICHNIS

- ALTRICHTER Herbert; Michael SCHRATZ; Hans PECHAR, Hrsg. (1997):
Hochschulen auf dem Prüfstand. Was bringt Evaluation für die
Entwicklung von Universitäten und Fachhochschulen? (Studien zur
Bildungsforschung und Bildungspolitik; 16). Innsbruck.
- ARGYRIS, Chris; Donald A. SCHÖN (1999): Die lernende Organisation.
Grundlagen, Methode, Praxis (aus d. Amerik.). Stuttgart.
- AXELROD, Robert (1988): Die Evolution der Kooperation (aus dem
Amerik.). München.
- BACHMAIER, Helmut; Ernst Peter FISCHER, Hrsg. (1997): Streit der
Fakultäten. Oder die Idee der Universität (Konstanzer Bibliothek; 24).
Konstanz.
- BAGGEN, Peter (1998): Knowledge Society and the Idea of the University.
In: Peter BAGGEN; Agnes TELLINGS; Wouter van HAAFTEN, Hrsg.:
The University and the Knowledge Society (Philosophy of Higher
Education Series; 1). Bommel, 87-104.
- BAGGEN, Peter; Agnes TELLINGS; Wouter van HAAFTEN, Hrsg. (1998):
The University and the Knowledge Society (Philosophy of Higher
Education Series; 1). Bommel.
- BAMBERG, Sebastian; Harald GUMBL; Peter SCHMIDT (2000): Rational
Choice und theoriegeleitete Evaluationsforschung. Opladen.
- BENGEL, Jürgen; Uwe KOCH (1988): Evaluationsforschung im Gesund-
heitswesen. In: Uwe KOCH; Gabriele LUCIUS-HOENE; Reiner STEGI,
Hrsg.: Handbuch der Rehabilitationspsychologie. Berlin.

-
- BOHM, David (2000): Der Dialog. Das offene Gespräch am Ende der Diskussionen (aus d. Engl.). 2. Aufl. Stuttgart.
- BORTZ, Jürgen; Nicola DÖRING (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3., überarb. Aufl., Berlin.
- BROEKSTRA, Gerrit (1998): An Organisation is a Conversation. In: David GRANT; Tom KEENOY; Cliff OSWICK, Hrsg.: Discourse and Organization. London u.a., 152 – 176.
- BRUENE Stefan, Hrsg. (1998): Erfolgskontrolle in der entwicklungspolitischen Zusammenarbeit. Hamburg.
- CAVELL, Stanley (2001): Nach der Philosophie. Essays. Hrsg. von Ludwig NAGL; Kurt R. FISCHER. 2., erweiterte und überarb. Aufl., Berlin.
- CHEN, Huey-Tsyh (1990): Theory Driven Evaluations. Newbury Park.
- CHEN, Huey-Tsyh; Peter H. ROSSI (1992): Using Theory to Improve Program and Policy Evaluations. New York.
- COOK, Thomas D.; Georg E. MATT (1990): Theorien der Programmevaluation – Ein kurzer Abriß. In: Uwe KOCH; Werner WITTMANN, Hrsg.: Evaluationsforschung – Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin, 15-38.
- CRONBACH, Lee J. (1982): Designing Evaluations of Educational and Social Programs. San Francisco.
- CRONBACH, Lee J. (1980): Toward Reform of Program Evaluation. San Francisco.
- DANIEL, Hans-Dieter; Rudolf FISCH, Hrsg. (1988): Evaluation von Forschung. Methoden, Ergebnisse, Stellungnahmen (Konstanzer Beiträge zur sozialwissenschaftlichen Forschung; 4). Konstanz.
- DAXNER, Michael (1999): Die blockierte Universität. Warum die Wissensgesellschaft eine andere Hochschule braucht. Frankfurt am Main u. New York.

-
- DENZIN, Norman K.; Yvonna S. LINCOLN, Hrsg. (1994): Handbook of Qualitative Research. Thousand Oaks.
- DRUCKER, Peter F. (2000): Die Kunst des Managements (aus d. Amerik.). 2. Aufl., München.
- DRUWE, Ulrich; Volker KUNZ, Hrsg. (1994): Rational Choice in der Politikwissenschaft. Opladen.
- DUCK, Steve (1998): Human Relationships. 3. Aufl., London u. Thousand Oaks.
- EIGEN, Manfred u.a. (1988): Die Idee der Universität. Versuch einer Standortbestimmung. Berlin u.a.
- EINIG, Bernhard; Uwe SCHMIDT (1999): Zur Situation von Studium und Lehre an der Johannes Gutenberg-Universität (Erster Gesamtbericht des Präsidenten). Mainz.
- ENGEL, Johann Jakob u.a. (1990): Gelegentliche Gedanken über Universitäten [zuerst 1800 - 1812]. Leipzig.
- ETZIONI, Amitai (1975): Die aktive Gesellschaft. Eine Theorie gesellschaftlicher und politischer Prozesse (aus d. Amerik.). Opladen.
- EVALUATIONSVERFAHREN,
<http://evanet.his.de/evanet/knowhow/kh.grund/kh.evalarten.html>
- FACHBEREICH CHEMIE, Hrsg. (2000): Evaluation im Fachbereich Chemie an der Johannes Gutenberg-Universität (Mainzer Beiträge zur Hochschulentwicklung; 1). Mainz.
- GRANT, David; Tom KEENOY; Cliff OSWICK, Hrsg. (1998): Discourse and Organization. London u.a.
- GREENE, Jennifer C. (1994): Qualitative Program Evaluation. Practice and Promise. In: Norman K. DENZIN; Yvonna S. LINCOLN, Hrsg.: Handbook of Qualitative Research. Thousand Oaks, 530-544.

-
- GROHMANN, Romano (1996): Das Problem der Evaluation in der Sozialpädagogik. Frankfurt a. M.
- HABERMAS, Jürgen (1992): Erläuterungen zur Diskursethik. Frankfurt/M.
- HABERMAS, Jürgen (1981): Theorie des kommunikativen Handelns. Bd. I und II., Frankfurt am Main.
- HAHNE, Anton (1998): Kommunikation in der Organisation. Opladen u. Wiesbaden.
- HAURIOU, Maurice (1965): Theorie der Institution und zwei andere Aufsätze. Hrsg. Roman SCHNUR. Berlin.
- HENNEN, Manfred (1994): Egoismus und Altruismus in der Sozialtheorie. In: Manfred HENNEN; Michael JÄCKEL, Hrsg.: Privatheit und soziale Verantwortung. Festschrift zum 60. Geburtstag von Friedrich Landwehrmann. München, 285 - 330.
- HENNEN, Manfred (2001): Hat Wissenschaft Macht? In: Bernd GOEBEL und Manfred WETZEL, Hrsg.: Eine moralische Politik? Vittorio Hösles Politische Ethik in der Diskussion. Würzburg: Königshausen und Neumann, 145-180.
- HENNEN, Manfred (1994): Ist Verhalten beeinflussbar? In: Evangelische Akademie Baden, Hrsg.: Wenn die Umwelt krank macht... Über den Zusammenhang von Gesundheit und Umwelt. Karlsruhe, 115 -160.
- HENNEN, Manfred (1998): Leitbilder zur Evaluation der Lehre. In: Hochschulrektorenkonferenz (HRK) Hrsg.: Evaluation und Qualitätssicherung an den Hochschulen in Deutschland - Stand und Perspektiven. Nationales Expertenseminar der Hochschulrektorenkonferenz, Bonn, 29. Mai 1998. Bonn, 61 - 77.
- HENNEN, Manfred (1994): Motivation als Konstrukt einer Sozialtheorie. In: Gerhard RUSCH; Siegfried J. SCHMIDT, Hrsg.: Konstruktivismus und Sozialtheorie. Frankfurt am Main, 133 - 171.
- HENNEN, Manfred (1990): Soziale Motivation und paradoxe Handlungsfolgen. Opladen.

-
- HENNEN Manfred; Michael JÄCKEL, Hrsg. (1994): Privatheit und soziale Verantwortung. Festschrift zum 60. Geburtstag von Friedrich Landwehrmann. München.
- HENNEN, Manfred; Thomas REIN (1994): Bindung und Motivation als implizite Annahmen der ‚Rational Choice‘-Theorie. In: Ulrich DRUWE; Volker KUNZ, Hrsg.: Rational Choice in der Politikwissenschaft. Opladen, 206 - 242.
- HENNEN, Manfred; Elisabeth SPRINGER (1994): Modellprojekt zur Förderung von Studium und Lehre (Schriften der Johannes Gutenberg-Universität; 1). Mainz.
- HIRSCHMANN, Albert O. (1974): Abwanderung und Widerspruch. Reaktionen auf Leistungsabfall bei Unternehmungen, Organisationen und Staaten (aus d. Amerik.). Tübingen.
- HOCHSCHULREKTORENKONFERENZ (HRK), Hrsg. (2000): Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen unter besonderer Berücksichtigung der Lehre (Dokumente und Informationen 1/2000). Bonn.
- HORNBOSTEL, Stefan (1997): Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen.
- HÖSLE, Vittorio (1997): Moral und Politik. Grundlagen einer Politischen Ethik für das 21. Jahrhundert. München.
- HOWALDT, Jürgen; Ralf KOPP, Hrsg. (1998): Sozialwissenschaftliche Organisationsberatung. Auf der Suche nach einem spezifischen Beratungsverständnis. Berlin.
- JASPERS, Karl (1946): Die Idee der Universität. Berlin.
- KÖBLER, Gerhard; Meinhard HEINZE; Wolfgang HROMADKA, Hrsg. (2000): Europas universale rechtsordnungspolitische Aufgabe im Recht des dritten Jahrhunderts. Festschrift für Alfred Söllner zum 70. Geburtstag. München.
- KOCH, Uwe; Gabriele LUCIUS-HOENE; Reiner STEGI, Hrsg. (1988): Handbuch der Rehabilitationspsychologie. Berlin.

-
- KOCH, Uwe; Werner WITTMANN, Hrsg. (1990): Evaluationsforschung. Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin.
- KORNHUBER, Hans Helmut (1988): Mehr Forschungseffizienz durch objektivere Beurteilung von Forschungsleistungen. In: DANIEL, Hans-Dieter; Rudolf FISCH, Hrsg. (1988): Evaluation von Forschung. Methoden, Ergebnisse, Stellungnahmen (Konstanzer Beiträge zur sozialwissenschaftlichen Forschung; 4). Konstanz, 361 – 382.
- KRASNEY, Otto Ernst (2000): Ein Ombudsmann für die Opfer von Massenunglücken? In: Gerhard KÖBLER; Meinhard HEINZE; Wolfgang HROMADKA, Hrsg.: Europas universale rechtsordnungspolitische Aufgabe im Recht des dritten Jahrhunderts. Festschrift für Alfred Söllner zum 70. Geburtstag. München, 575 - 588.
- KROMREY, Helmut (2000): Qualität und Evaluation im System Hochschule. In: Reinhard STOCKMANN, Hrsg.: Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. (Sozialwissenschaftliche Evaluationsforschung; 1). Opladen, 233-258.
- KUNZ, Volker (1997): Theorie rationalen Handelns. Opladen.
- LENK, Hans (1995): Schemaspiele. Über Schemainterpretationen und Interpretationskonstrukte. Frankfurt am Main.
- LEPENIES, Wolf (1989): Gefährliche Wahlverwandtschaften. Essays zur Wissenschaftsgeschichte. Stuttgart.
- LUHMANN, Niklas (2000): Organisation und Entscheidung. Opladen.
- LUHMANN, Niklas (1994): Perspektiven für Hochschulpolitik. In: Ders.: Soziologische Aufklärung 4. Beiträge zur funktionalen Differenzierung der Gesellschaft. Opladen, 216-223.
- MATTELART, Armand; Michèle MATTELART (1998): Theories of Communication (aus d. Franz.). London u.a., 135.

-
- MAYNTZ, Renate u.a.: Differenzierung und Verselbständigung. Zur Entwicklung gesellschaftlicher Teilsysteme. Frankfurt/M. u. New York.
- MINSSEN, Heiner (1998): Soziologie und Organisationsberatung. In: Jürgen HOWALDT; Ralf KOPP, Hrsg.: Sozialwissenschaftliche Organisationsberatung. Auf der Suche nach einem spezifischen Beratungsverständnis. Berlin, 53-72.
- MITTAG, Waldemar; Matthias JERUSALEM (1997): Evaluation von Präventionsprogrammen. In: Rolf SCHWARZER, Hrsg.: Gesundheitspsychologie. Göttingen, 595-613.
- MÜLLER, Klaus (1996): Allgemeine Systemtheorie. Geschichte, Methodologie und sozialwissenschaftliche Heuristik eines Wissenschaftsprogramms. Opladen.
- MÜLLER-BÖLING, Detlef (1997): Evaluation zur Rechenschaftslegung oder Qualitätsverbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen. In: Herbert ALTRICHTER; Michael SCHRATZ; Hans PECHAR, Hrsg.: Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen? (Studien zur Bildungsforschung und Bildungspolitik; 16). Innsbruck, 88-107.
- MÜNCH, Richard (1982): Theorie des Handelns. Zur Rekonstruktion der Beiträge von Talcott Parsons, Emile Durkheim und Max Weber. Frankfurt/M.
- NEWMAN, John Henry (1996): The Idea of a University [zuerst 1852-1873]. Hrsg. v. Frank M. Turner. New Haven u. London.
- NICHOL, Lee (2000): Vorwort. In: David BOHM: Der Dialog (aus d. Engl.). 2. Aufl. Stuttgart.
- PARSONS, Talcott (1996): Das System moderner Gesellschaften (aus d. Amerik.). 4. Aufl., München.
- PARSONS, Talcott; Gerald M. PLATT (1990): Die amerikanische Universität (aus d. Amerik.). Frankfurt am Main.

-
- PATTON, Michael Quinn (1996): *Utilization-Focused Evaluation*. 3. Aufl., London u.a.
- PELIKAN, Jaroslav (1992): *The Idea of the University. A Reexamination*. New Haven u. London.
- PIRKER, Reinhard (1997): *Die Unternehmung als soziale Institution. Eine Kritik der Transaktionskostenerklärung der Firma*. In: Günther ORTMANN, Jörg SYDOW, Klaus TÜRK, Hrsg.: *Theorien der Organisation. Die Rückkehr der Gesellschaft*. Opladen, 67-80.
- READINGS, Bill (1997): *The University in Ruins*. Cambridge u. London.
- RICHTER, Roland (1996): *Qualitätsmanagement im europäischen Hochschulwesen – Drei Varianten über ein Thema*. In: *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie* 16, 165-183.
- RITSERT, Jürgen (2001): *Soziologie des Individuums*. Darmstadt.
- ROHRMOSER, Günter (1994): *Zur Idee und Reformierbarkeit der Universität*. Stuttgart.
- ROSSI, Peter H.; Howard E. FREEMAN (1993): *Evaluation. A Systematic Approach*. 5. Aufl., London u.a.
- ROSSI, Peter H.; Howard E. FREEMAN; Gerhard HOFMANN (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*. Stuttgart.
- ROSSI, Peter H.; Howard E. FREEMAN; Mark W. LIPSEY (1999): *Evaluation. A Systematic Approach*. 6. Aufl., Thousand Oaks u.a.
- RUSCH, Gerhard; Siegfried J. SCHMIDT, Hrsg. (1994): *Konstruktivismus und Sozialtheorie*. Frankfurt am Main.
- SANDERS James R., Hrsg. (2000): *Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“*. 2., durchges. Aufl., Opladen.

-
- SANDERS, James R. (2000): Aspekte der Entwicklung und Verbreitung der Evaluationsstandards. In: James R. SANDERS, Hrsg.: Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. 2., durchges. Aufl., Opladen, 237-241.
- SCHELSKY, Helmut (1971): Einsamkeit und Freiheit. Idee und Gestalt der deutschen Universität und ihrer Reform. 2., um einen „Nachtrag 1970“ erweiterte Aufl., Gütersloh.
- SCHIMANK, UWE (2001): Festgefahrene Gemischtwarenläden – Die deutschen Hochschulen als erfolgreich scheiternde Organisationen. In: STOELTING, Erhard; Uwe SCHIMANK, Hrsg. (2001): Die Krise der Universitäten. (Leviathan; SoH 20). Opladen. 223-242.
- SCHMID, Michael (1998): Soziales Handeln und strukturelle Selektion. Opladen u. Wiesbaden.
- SCHMIDT, Uwe; Simon HÄUSER (2001): Problembereiche und Reformansätze in Studium und Lehre (Mainzer Beiträge zur Hochschulentwicklung; 2). Mainz.
- SCHULZ VON THUN, Friedemann (1998): Miteinander reden. Bd. 3. Reinbek bei Hamburg.
- SCHWARZER, Rolf, Hrsg. (1997): Gesundheitspsychologie. Göttingen.
- SENGE, Peter M. (2001): Die fünfte Disziplin (aus d. Amerik.). 8. Aufl., Stuttgart.
- SPRANGER, Eduard (1973): Das Wesen der deutschen Universität. In Ders.: Hochschule und Gesellschaft. Hrsg. Walter Sachs. Heidelberg, 82-158.
- STICHWEH, Robert (1988): Differenzierung des Wissenschaftssystems. In: Renate MAYNTZ u.a.: Differenzierung und Verselbständigung. Zur Entwicklung gesellschaftlicher Teilsysteme. Frankfurt/M. u. New York, 45-115.

-
- STICHWEH, Rudolf (1991): Der frühmoderne Staat und die europäische Universität. Frankfurt am Main.
- STOCKMANN, Reinhard, Hrsg. (2000): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. (Sozialwissenschaftliche Evaluationsforschung; 1). Opladen.
- STOCKMANN, Reinhard (1998): Kleine Entwicklungsgeschichte der Evaluationsforschung. Nachholende Entwicklung in Deutschland. In: Stefan BRUENE, Hrsg.: Erfolgskontrolle in der entwicklungspolitischen Zusammenarbeit. Hamburg, 27-63.
- STOELTING, Erhard; Uwe SCHIMANK, Hrsg. (2001): Die Krise der Universitäten. (Leviathan; SoH 20). Opladen.
- TYRELL, Hartmann (1976): Probleme einer Theorie der Ausdifferenzierung der privatisierten modernen Kernfamilie. In: Zeitschrift für Soziologie 5, 175-193.
- VEIL, Cornelia (1993): Relationale Kommunikation. Kommunikationsverständnis und -praxis im Wandel. 2. Aufl., München u. Mering.
- VERBUND NORDDEUTSCHER HOCHSCHULEN, Hrsg. (1999): Evaluation von Studium und Lehre im Fach Mathematik 1997/98 (Verbund-Materialien; 3). Hamburg.
- VERBUND NORDDEUTSCHER HOCHSCHULEN, Hrsg. (1999): Evaluation von Studium und Lehre im Fach Geowissenschaften 1997/98 (Verbund-Materialien; 4). Hamburg.
- WEICK, Karl E. (1995): Sensemaking in Organizations. London u.a.
- WEISS, Carol H. (1974): Evaluationsforschung. Methoden zur Einschätzung von sozialen Reformprogrammen. Opladen.
- WIDMER, Thomas (2000): Qualität der Evaluation. In: Reinhard STOCKMANN, Hrsg.: Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Opladen, 77 – 102.

WIDMER, Thomas; Wolfgang BEYWL (2000): Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder. In: James R. SANDERS, Hrsg.: Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. 2., durchges. Aufl., Opladen, 243-257.

WILLKE, Helmut (1998): Systemisches Wissensmanagement. Stuttgart.

WOTTAWA, Heinrich; Heike THIERAU (1998): Lehrbuch Evaluation. 2., vollst. überarb. Aufl., Bern u.a.

Die Autoren:

Dr. Katja AHLSTICH, Zentrum für Qualitätssicherung und –entwicklung, Johannes Gutenberg-Universität Mainz (ZQ).

Prof. Dr. Ulrich DRUWE, Vizepräsident für Studium und Lehre der Johannes Gutenberg-Universität Mainz.

Prof. Dr. Manfred HENNEN, Leiter des Zentrum für Qualitätssicherung und –entwicklung, Johannes Gutenberg-Universität Mainz (ZQ).

Prof. Dr. Jürgen MARKL, Institut für Zoologie, Johannes Gutenberg-Universität Mainz.

Prof. Dr. Harald PAULSEN, Institut für Allgemeine Botanik, Johannes Gutenberg-Universität Mainz.

Uwe SCHMIDT, Dr., Zentrum für Qualitätssicherung und –entwicklung, Johannes Gutenberg-Universität Mainz (ZQ).

Elisabeth SPRINGER, M.A. Zentrum für Qualitätssicherung und –entwicklung, Johannes Gutenberg-Universität Mainz (ZQ).

Tanja URICH-NEITZERT, M.A. Zentrum für Qualitätssicherung und –entwicklung, Johannes Gutenberg-Universität Mainz (ZQ).