

# **Leistungsbewertung und Leistungsanreize in der Hochschullehre**

Eine Untersuchung von Konzepten, Leistungskriterien und  
Bedingungen erfolgreicher Institutionalisierung

Dissertation zur Erlangung des Grades eines Doktors der Philosophie

an der

Philosophischen Fakultät

der

Technischen Universität Dresden

vorgelegt am 4. Mai 2005 von

René Krempkow

geb. am 18.02.1970 in Ludwigslust (Meckl.)

Betreuer: Prof. Dr. Karl Lenz, Technische Universität Dresden

Gutachter: 1. Prof. Dr. Karl Lenz, Technische Universität Dresden  
2. Prof. Dr. André Wolter, Technische Universität Dresden  
3. Prof. Dr. Karl-Heinz Reuband, Universität Düsseldorf

Termin der Verteidigung: 7. Oktober 2005

# Inhalt

1. Einführung in das Thema und Erläuterung der Forschungsfrage .....	5
1.1 Ausgangslage .....	5
1.2 Zielstellung und Schwerpunkte der Arbeit .....	7
1.3 Zum methodischen Vorgehen .....	13
2. Was ist Leistung, was ist Qualität? – Klärung von unterschiedlichen Begriffsverständnissen .....	15
2.1 Zur Klärung der Begriffe Leistungsbewertung, Leistungsanreize und Hochschullehre .....	16
2.1.1 Leistung und Leistungsbewertung .....	16
2.1.2 Leistungsanreize .....	19
2.1.3 (Hochschul-)Lehre .....	19
2.2 Evaluation und Qualitätsbewertung .....	20
2.2.1 Zum Begriff Evaluation .....	20
2.2.2 Die Entwicklung des Qualitätsbegriffs in der „Dienstleistungsgesellschaft“ .....	22
2.3 Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen .....	27
2.3.1 Definition von Lehrqualität auf Hochschulebene .....	29
2.3.2 Definition von Lehrqualität auf Studiengangsebene .....	32
2.3.3 Definition von Lehrqualität auf Lehrveranstaltungsebene .....	33
2.3.4 Zielgruppen für die Qualität der Hochschullehre .....	36
2.3.5 Fazit zur Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen .....	38
3. Konzepte, Instrumente und Modelle zur Qualitätsorientierung an Hochschulen .....	40
3.1 Überblick über Konzepte, Instrumente und Modelle .....	41
3.1.1 Überblick über Konzepte und Versuch einer Systematisierung .....	41
3.1.2 Überblick über Instrumente und Modelle zur Förderung der Qualitätsorientierung .....	43
3.2 Leistungs- und Qualitätsbewertung .....	47
3.2.1 Evaluation .....	48
3.2.2 Berichtssysteme .....	49
3.2.3 Rankings bzw. vergleichende Studienführer .....	51
3.3 Modelle des Qualitätsmanagements und der Qualitätsorientierung .....	53
3.3.1 DIN ISO 9000ff .....	54
3.3.2 Total Quality Management .....	56
3.3.3 Balanced Scorecard .....	60
3.4 Leistungsanreizmodelle .....	63
3.4.1 Nichtmaterielle Anreize .....	64
3.4.2 Leistungsabhängige Besoldungszulagen .....	66
3.4.3 Leistungsorientierte Verteilung finanzieller Mittel bzw. Ressourcen .....	76
3.5 Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen .....	88
3.5.1 Schwächen von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen als Ausgangspunkt .....	88
3.5.2 Gegenüberstellung theoretischer Überlegungen und eingesetzter bzw. diskutierter Indikatoren der leistungsorientierten Mittelverteilung an ausgewählten Hochschulen .....	92
3.5.3 Indikatoren zur Leistungs- bzw. Qualitätsbewertung in mehrdimensionalen Rankings/ vergleichenden Studienführern .....	94
3.5.4 Fazit .....	97

4. Experteninterviews zu Leistungsbewertung und Leistungsanreizen im Hochschulbereich.....	100
4.1 Ziele der Experteninterviews im Forschungsvorhaben.....	101
4.2 Methodisches Vorgehen .....	102
4.3 Durchführung der Experteninterviews.....	108
4.4 Auswertung der Experteninterviews.....	111
4.4.1 Arbeitsschritte der Auswertung.....	111
4.4.2 Thematischer Vergleich .....	112
4.4.3 Soziologische Konzeptualisierung .....	135
4.5 Fazit .....	140
5. Kriterien für Leistungsbewertung und Leistungsanreize .....	142
5.1 Kriterien, Kennziffern und Indikatoren .....	143
5.1.1 Unterscheidung zwischen Kriterien, Kennziffern und Indikatoren.....	143
5.1.2 Anforderungen an Kriterien, Kennziffern und Indikatoren.....	144
5.2 Die Eignung „objektiver“ Kennziffern und Indikatoren.....	147
5.2.1 Objektive Kennziffern und Indikatoren - Ersatz subjektiver Bewertung? .....	147
5.2.2 Studenten- und Studienanfängerzahlen als Orientierungshilfe, aber nicht als <i>Leistungsindikator</i> .....	149
5.2.3 Anzahl (abgelehnter) Bewerber – Indiz für Attraktivität des Studienganges?.....	153
5.2.4 Betreuungsrelation als ein Indikator für Überfüllung .....	154
5.2.5 Durchschnittliche Fachstudiendauer - ein Indikator für Lehrqualität?.....	158
5.2.6 Studierbarkeit – besseres Kriterium als die Studiendauer? .....	162
5.2.7 Abschlußnoten: Kein Kriterium für Studienerfolg, aber Orientierungshilfe.....	164
5.2.8 Erfolgsquote bei Prüfungen: Qualitätskriterium eines Studienganges? .....	167
5.2.9 Absolventenquote als Indikator für „Produktivität“?.....	169
5.2.10 Fazit zur Aussagekraft „objektiver“ Kennziffern und Indikatoren .....	176
5.3 Subjektive Bewertungen als Leistungskriterien?.....	180
5.3.1 Die Bedeutung subjektiver Bewertungen für die Qualitätsorientierung .....	181
5.3.2 Möglichkeiten zur Erhebung subjektiver Bewertungen .....	182
5.4 Die Eignung von Studentenbefragungen zur Qualität der Lehre.....	185
5.4.1 Wie kann Lehrqualität mittels Studentenbefragungen „gemessen“ werden?.....	185
5.4.2 Die Datenqualität von Ergebnissen studentischer Bewertungen der Lehre - oder: Welche Argumente gegen Befragungen halten empirischer Prüfung stand? .....	187
5.4.3 Vergleichbarkeit der Ergebnisse von Studentenbefragungen.....	206
5.4.4 Weitere Anforderungen an Studentenbefragungen als Leistungsindikator.....	213
5.4.5 Fazit zur Eignung von Studentenbefragungen .....	216
5.5 Die Eignung von Absolventenbefragungen zur Messung der Wirkungen der Hochschulausbildung.....	218
5.5.1 Wie können Wirkungen der Hochschulausbildung mittels Absolventenstudien „gemessen“ werden? .....	218
5.5.2 Die Datenqualität der Ergebnisse von Absolventenbefragungen.....	225
5.5.3 Vergleichbarkeit der Ergebnisse von Absolventenbefragungen .....	229
5.5.4 Weitere Anforderungen .....	240
5.5.5 Fazit zur Eignung von Absolventenbefragungen .....	243
5.6 Lehrendenbefragungen als Bestandteil der Qualitätsorientierung .....	245
5.6.1 Die Datenqualität der Ergebnisse von Lehrendenbefragungen .....	245
5.6.2 Vergleichbarkeit .....	259
5.6.3 Weitere Anforderungen .....	260
5.6.4 Fazit zur Eignung von Lehrendenbefragungen .....	262

6. Bedingungen erfolgreicher Institutionalisierung.....	264
6.1 Machtkonstellationen, Interessen und Akzeptanz .....	265
6.1.1 Machtkonstellationen und Interessen als Institutionalisierungsbedingungen .....	267
6.1.2 Akzeptanz als Bedingung erfolgreicher Institutionalisierung .....	272
6.2 Lehrendenbefragungen als Instrument der Akzeptanzuntersuchung .....	277
6.3 Die Eignung vorhandener Lehrendenbefragungen .....	278
6.3.1 Entstehungshintergrund der bundesweiten Lehrendenbefragungen.....	279
6.3.2 Zur Anlage und Repräsentativität der bundesweiten Lehrendenbefragungen .....	281
6.3.3 Ergebnisse der Lehrendenbefragungen zur Arbeitssituation.....	282
6.3.4 Wunsch und Wirklichkeit beim Zeitbudget der Lehrenden an deutschen Hochschulen...	283
6.3.4 Einschätzung hochschulpolitischer Reformvorhaben an deutschen Hochschulen.....	285
6.4 Akzeptanz von Reformvorhaben und Arbeitszufriedenheit bei Lehrenden in Dresden, Frankfurt/ Oder und Zittau-Görlitz .....	286
6.4.1 Anlage und Durchführung der eigenen Lehrendenbefragungen .....	287
6.4.2 Betrachtung von Rücklauf und Repräsentativität der Befragungen .....	288
6.4.3 Akzeptanz hochschulpolitischer Reformvorhaben.....	292
6.4.4 Berufliche Zufriedenheit der Lehrenden am Beispiel der TU Dresden .....	299
6.4.5 Das Zeitbudget und sein Zusammenhang mit der Arbeitszufriedenheit .....	303
6.5 Determinanten der beruflichen Zufriedenheit.....	314
6.5.1 Konzepte zu Determinanten der beruflichen Zufriedenheit .....	314
6.5.2 Empirische Überprüfung der Konzepte.....	317
6.5.3 Subgruppenanalysen.....	317
6.5.4 Bivariate und multivariate Zusammenhangsanalysen .....	321
6.5.5 Fazit zu Determinanten der beruflichen Zufriedenheit .....	329
6.6 Fazit zu Bedingungen erfolgreicher Institutionalisierung .....	330
7. Zusammenfassung .....	334
7.1 Vorstellung und Einordnung von Konzepten, Modellen und Instrumenten .....	336
7.2 Analyse von (potentiellen) Kriterien und Indikatoren .....	339
7.3 Bedingungen erfolgreicher Institutionalisierung .....	347
8. Literaturverzeichnis.....	350
Anhang A:	
Tabellarische Übersicht prägnanter Textpassagen .....	377
Anhang B:	
Leitfaden für Experteninterviews.....	397
Anhang C:	
Hypothesen zur Analyse der Bestimmungsgründe beruflichen Erfolges anhand der Dresdner Absolventenstudien .....	399
Anhang D:	
Ausgaben für zentrale Ergebnisse der Regresionsanalysen zu Determinanten beruflicher Zufriedenheit .....	401
Anhang E:	
Dokumentation der Fragebögen .....	403

# 1. Einführung in das Thema und Erläuterung der Forschungsfrage

Gliederung:

1.1 Ausgangslage

1.2 Zielstellung und Schwerpunkte der Arbeit

1.3 Zur Methode

## 1.1 Ausgangslage

Gut eineinhalb Jahrzehnte sind vergangen, seit das Thema "Qualität der Hochschullehre" in Deutschland auf die Tagesordnung gebracht wurde. Anteil daran hatten Massenmedien<sup>1</sup>, aber auch studentische und gewerkschaftliche<sup>2</sup> Initiativen. Inzwischen zeigt sich, daß dieses Thema längst nicht mehr nur einfach ein Modethema ist, wie anfangs teilweise vermutet wurde. Vielmehr zeichnen sich in der Hochschullandschaft Deutschlands seit einiger Zeit tiefgreifende Veränderungsprozesse ab, die die Hochschulen von "Stätten freier Geistesbildung zu gesellschaftlichen Dienstleistungseinrichtungen" werden lassen (vgl. Wolter 1995).

Immer deutlicher wird, daß angesichts steigender Studentenzahlen bei knapper werdenden finanziellen Mitteln der politische Druck auf öffentliche Rechenschaftslegung der Hochschulen über die Verwendung der gesamtgesellschaftlichen Ressourcen wächst. Mit den bisherigen Steuerungsmitteln wird es jedoch immer schwieriger, den Problemen von Massenhochschulen und veränderten gesellschaftlichen Erwartungen gerecht zu werden. Eine stärker leistungsorientierte Finanzierung von Hochschulen und Fachbereichen wird auch deshalb schon länger diskutiert. Es ist – für die Verhältnisse der Veränderungen im Hochschulbereich – noch nicht allzu lange her, daß die Hochschulrektorenkonferenz den Vorschlag einer Bund-Länder-Arbeitsgruppe aufgriff, das Besoldungsrecht zu reformieren und die bislang altersabhängigen Besoldungszulagen in leistungsabhängige Zulagen umzuwandeln.<sup>3</sup> Inzwischen wurde die Dienst-

---

<sup>1</sup> Zu einem Thema des öffentlichen Interesses und letztlich auch zu einem Thema für die (Sozial-)Wissenschaft wurde die Qualität der Hochschullehre erst durch eine Veröffentlichung des Nachrichtenmagazins "Der Spiegel". Das Magazin publizierte in Nr. 50/ 1989 erstmals ein großangelegtes Hochschul-Ranking für Deutschland. Dieses wurde in mehreren Ausgaben der Zeitschrift „Soziologie“ (1990, 1991) diskutiert.

<sup>2</sup> Beispielsweise die 11. GEW-Sommerschule 1991 stand speziell unter dem Thema "Qualität von Lehre und Studium". Sie fand in Zusammenarbeit mit Hans-Böckler-Stiftung und Friedrich-Ebert-Stiftung statt (vgl. GEW: "Materialien und Dokumente zu Hochschule und Forschung" Nr. 66).

<sup>3</sup> Hierzu legte die Hochschulrektorenkonferenz "Empfehlungen zum Dienst- und Tarif-, Besoldungs- und Vergütungsrecht sowie zur Personalstruktur in den Hochschulen" vor (Entschließung vom 2. 11. 1998). Die Bundesregierung erarbeitete auf der Grundlage des Berichtes der Expertenkommission „Reform des Hochschuldienstrechts“ (bmb+f 2000a) ihr Konzept „Hochschuldienstrecht für das 21. Jahrhundert“ (bmb+f 2000b). Geplant sei "eine eindeutige Bewertung - der Forschung und vor allem auch der Lehre". (...) "Ohne die Studenten wird es aber nicht gehen. Wir brauchen deren Meinung. Und wir werden sie auch einbeziehen." (vgl. Interview von Bundesbildungsministerin Edelgard Bulmahn (SPD) in "Sächsische Zeitung" vom 8. 9. 1999, S. 27) (Vgl. auch "Sächsische Zeitung" vom 2. 6. 1999, S. 27 und "Süddeutsche Zeitung" vom 22. 6. 1999, S. 2)

rechtsreform auf Bundesebene beschlossen und ist in den einzelnen Bundesländern umzusetzen. Es geht also längst nicht mehr um, das „ob“, sondern um das „wie“ bei diesen Veränderungen. Doch über Konzepte und Kriterien herrscht nach wie vor keine Einigkeit. Dabei erscheint diese Diskussion symptomatisch für die Einführung von Leistungsanreizen an Hochschulen. Das gilt offenbar nicht nur für Deutschland, sondern auch für andere europäische Länder.<sup>5</sup> Ohne die Unterschiede in den Bildungs- und insbesondere den Hochschulsystemen verkennen zu wollen, wird bereits seit längerer Zeit von einer „Phase des Übergangs“ hin zu wettbewerbsähnlichen Strukturen, stärkerer Autonomie und Übernahme von Verantwortung durch die Hochschulen in den meisten europäischen Ländern gesprochen (vgl. Goedegebuure u.a. 1993: 419). Eine gewisse Vorbildrolle für die Umgestaltung des deutschen Hochschulsystems haben die Niederlande inne. Nach dem Wahlsieg der „New Labour“ blickte man in der Politik auch auf Großbritannien. Hinzu kommt, daß das angelsächsische System<sup>6</sup> in einigen Punkten (z.B. Einführung von Bachelor- und Master-Studiengängen im Zuge des sogenannten Bologna-Prozesses) auch in Deutschland auf der politischen Ebene als nachahmenswert angesehen wird. Bei der Lektüre einschlägiger wissenschaftlicher Literatur fällt außerdem auf: Die aktuelle Diskussion in der deutschen Hochschulpolitik gleicht in mehreren Punkten jener Diskussion z.B. in der britischen Hochschulpolitik vor etwa 10 Jahren (vgl. Bruce 1999).<sup>7</sup>

So sind die britischen Universitäten wie die deutschen zwar traditionell durch ein relativ großes Maß an Hochschulautonomie gekennzeichnet. Andererseits müssen britische Universitäten und seit einigen Jahren auch die deutschen mit enormen Sparzwängen bei steigenden Studentenzahlen existieren. (vgl. Forbeck 2000, Scott 1995). Und wie in Großbritannien vor einigen Jahren und seitdem auch in vielen anderen OECD-Staaten wird in Deutschland verstärkt eine Verknüpfung zwischen Ergebnissen der Bewertung von Lehr- und Forschungsleistungen und der Vergabe finanzieller Mittel diskutiert (vgl. Leszczensky/ Orr 2003, 2004). An den britischen Hochschulen wurden solche Kopplungsmechanismen nach einer Phase der Konfrontation zwischen Ministerien und Hochschulen bei ihrer Einführung letztlich weitgehend akzeptiert (vgl. Leszczensky/ Orr 2003: 44, Jackson 1997). Aus der Schweiz wird ähnliches berichtet (vgl. Schenker-Wicki 2002). Aber auch andere Länder außerhalb Europas wie Australien haben

---

<sup>5</sup> So sind inzwischen in vielen OECD-Ländern leistungsorientierte Finanzzuweisungsverfahren eingeführt worden, mit denen Anreize zu Leistungssteigerungen gesetzt werden (vgl. Leszczensky u.a. 2004).

<sup>6</sup> Natürlich kann man nicht von *dem* angelsächsischen System im eigentlichen Sinne sprechen, da es deutlich stärker diversifiziert ist als beispielsweise das deutsche Hochschulsystem. Es gibt jedoch wichtige Gemeinsamkeiten (vgl. Berdahl u.a. 1991, Kogan/ Hanney 1999, Miller 1999).

<sup>7</sup> Am intensivsten wurde in Großbritannien wohl der „Dearing-Report“ und die ihm vorausgehenden „white papers“ diskutiert. (vgl. Dearing 1997 bzw. Goedegebuure u.a. 1993: 166 ff.).

bereits längere Zeit Erfahrungen mit Leistungsbewertungen und Leistungsanreizen gesammelt. Einige australische Entwicklungen können sogar als eine Art Re-Import in europäische Hochschulsysteme angesehen werden,<sup>8</sup> denn das australische Hochschulsystem lehnte sich historisch bedingt ursprünglich stark an das britische an.

In Deutschland wurde in den letzten Jahren häufiger die Einbeziehung der Modelle anderer Länder in die nationalen Reformpläne diskutiert. Welche Effekte Kopplungsmechanismen zwischen Qualitätsbewertungen und Vergabe finanzieller Mittel für die Qualität der Lehre und des Studiums jedoch im einzelnen haben (können) und welche Faktoren bzw. Kriterien dabei ausschlaggebend sind, ist für Deutschland bisher kaum untersucht worden. Vor allem die (möglichen) Auswirkungen von Anreizen in der Lehre auf die Arbeitszufriedenheit und das (Arbeits-)Klima der Lehrenden sowie das Zusammenspiel innerhalb der Hochschulen wurden bislang in Deutschland eher wenig erörtert.<sup>9</sup> Für die (Mit-) Gestaltung und Akzeptanz der sich abzeichnenden Veränderungsprozesse könnte es aber von großem Interesse sein, solche Aspekte stärker einzubeziehen.

## 1.2 Zielstellung und Schwerpunkte der Arbeit

Ziel der Arbeit ist es, bei der Umgestaltung des Hochschulwesens bisher gesammelte Erfahrungen zu Leistungsbewertung und Leistungsanreizen in der Hochschullehre und dabei insbesondere bisher wenig beachtete Aspekte in die Diskussion einzubeziehen, um daraus ggf. Möglichkeiten zur Weiterentwicklung abzuleiten. Schwerpunkte der Arbeit sollen sein:

**- erstens ein Überblick über die Konzepte, Modelle und Instrumente**, auf deren Grundlage bzw. mit denen die Veränderungsprozesse zu Leistungsbewertungen und -anreizen in den verschiedenen Ebenen des Hochschulsystems stattfanden bzw. noch stattfinden (Dabei liegt ein besonderes Augenmerk auf dem Aspekt der Qualität von Leistungen, der in diesem Zusammenhang bisher häufig eher nachrangig erscheint),

---

<sup>8</sup> Ein Beispiel wären die „nachgelagerten“ Studiengebühren, die in Schottland nach einer Phase direkt im Studium fälliger Studiengebühren unter Rückgriff auf australische Erfahrungen eingeführt wurden. In Deutschland ist dieses Modell auch unter dem Begriff „Akademikersteuer“ bekannt, da die Absolventen erst zahlen müssen, wenn sie im Beruf stehen und ein entsprechendes (Mindest-)Einkommen erzielen.

<sup>9</sup> Auch in jüngeren Veröffentlichungen zum Thema (z.B. Böhm 2000: 116) wurden lediglich Erfahrungen von amerikanischen und australischen Hochschulen sowie von ehemals öffentlichen Dienstleistern wie „Deutsche Post“ diskutiert, die allerdings nahelegen, den Kriterien und der Mitarbeiterorientierung weit größere Aufmerksamkeit als bisher zu widmen. Von Großbritannien wird ebenfalls berichtet, daß Konkurrenzdruck und Streß im Universitätssektor gestiegen und die Kollegialität gesunken sei (Vgl. Orr 2001: 159).

- **zweitens eine Analyse der Leistungskriterien bzw. Indikatoren**, die bei der Einführung von Leistungsbewertungen und -anreizen Verwendung fanden oder finden sollen (dabei erfolgt eine ausführlichere Diskussion bisher kaum in diesem Zusammenhang diskutierter subjektiver Bewertungen zur Erfassung von Qualität), und

- **drittens eine Diskussion von Bedingungen erfolgreicher Institutionalisierung** von Leistungsbewertungen und -anreizen, um hieraus evtl. Ansatzpunkte für Schlußfolgerungen ableiten zu können (insbesondere wird die Akzeptanz bei den Betroffenen als Grundlage für eine „Veränderung in den Köpfen“ und damit als bisher wenig fokussierte Voraussetzung zur Verwirklichung von Qualitätsorientierung diskutiert).

In diesem Zusammenhang soll auf einige thematische Einschränkungen hingewiesen werden, die zum Verständnis der gewählten Schwerpunkte wichtig sein könnten: So erfolgte eine Einschränkung auf den Aufgabenbereich Lehre an Hochschulen - wie der Titel der Arbeit bereits anzeigte. Hierbei ist klarzustellen, dass dies nur ein zentraler Aufgabenbereich neben der Forschung sein kann. Aber Lehre und insbesondere deren Qualität hat in bisher existierenden Anreizsystemen an Hochschulen einen geringeren Stellenwert als die Forschung. Ein jüngstes Beispiel hierfür kann die Eliteförderung bzw. die Exzellenzförderung des Bundes sein, in der Spitzenforschung im Vordergrund steht, Spitzenlehre jedoch kaum eine Rolle spielt. Zudem ist das Thema Bewertung von Forschungsleistungen eher noch anspruchsvoller, wenn man den Qualitätsaspekt einbeziehen möchte, und erfordert daher eine eigenständige Bearbeitung. So hat sich die DFG kürzlich entschieden, hierzu ein eigenes Institut zu gründen.

Insgesamt war es grundsätzlich auch nicht etwa das Ziel, ein fertiges Leistungsanreizsystem und eine komplette Einführungsstrategie vorzulegen. Dies kann m.E. nur im Zusammenwirken mit betroffenen Hochschulen und Wissenschaftsverwaltungen gelingen und nicht im Alleingang einer einzelnen Person. Das Ziel war daher vielmehr, eine Grundlage für mögliche Weiterentwicklungen zu schaffen und einen Beitrag zur Diskussion über eine stärkere Verankerung von Qualitätsaspekten bei Indikatoren und Leistungsanreizen zu leisten. Diese Grundlage könnte dann in angewandten (Begleit-)Forschungsprojekten genutzt werden, für die von der Wissenschaftsverwaltung dann entsprechende Ressourcen bereitgestellt werden sollten.

### ***zu erstens) Überblick über Konzepte, Modelle und Instrumente***

Der Überblick über und die Einordnung von Konzepten, die hinter den Veränderungsprozessen stehen, soll den ersten Schwerpunkt der Arbeit bilden. Angesichts der Fülle von Konzepten und des von der deutschen Politik definierten Nachholbedarfs an Leistungsanreizen insbesondere in der Lehre<sup>10</sup> erscheint es sinnvoll, sich in dieser Diskussion auf den Aspekt der Lehre zu konzentrieren. Hierbei ist dies – worauf bereits hingewiesen wurde – natürlich in einem breiteren Kontext zu sehen und entsprechend einzuordnen. Eine Vorstellung *aller* möglicherweise relevanten Konzepte und Modelle im Hochschulbereich kann im Rahmen dieses Vorhabens in angemessener Zeit nicht geleistet werden. Deshalb soll eine ausführlichere Vorstellung nur für eine Auswahl erfolgen.

Hierbei wird insbesondere davon ausgegangen, daß im Hochschulbereich Leistung nicht allein auf das Erbringen einer bestimmten Menge (z.B. an Absolventen, wiss. Publikationen) reduziert werden und daher Leistung immer im Zusammenhang mit deren Qualität betrachtet werden soll. Veränderungsprozesse, die in angelsächsisch geprägten Bildungssystemen vorgenommen wurden, erfolgten häufig unter den Stichworten Quality Management (QM) oder Total Quality Management (TQM) und enthalten Leistungsanreize. An solchen Ansätzen und Konzepten, die meist eine Stärkung des Hochschulmanagements befürworten, wird oft eine zu sehr betriebswirtschaftliche Betrachtungsweise kritisiert (vgl. z.B. Ronge 2000: 189). Es mehren sich jedoch die Stimmen, die modifizierte Ansätze des QM auch auf Hochschulen anwendbar halten (vgl. Sims & Sims 1995: 138, Dahlgaard/ Madson 1998, Zink 1998, 1999). Diese Diskussion soll aufgegriffen werden, wobei der Fokus v.a. auf Ansätze gerichtet werden soll, in denen die Mitarbeiterorientierung eine größere Rolle spielt.

Die Notwendigkeit einer solchen Diskussion wird vor dem Hintergrund gesehen, daß sich die Hochschulen entwickelter Industriestaaten generell veränderten Erwartungen und einer stärkeren Verantwortlichkeit für die Verwendung ihrer Ressourcen gegenüber der Gesellschaft stellen müssen. Im Gegensatz zu einigen betriebswirtschaftlich orientierten Ansätzen sollen hier unter Ressourcen nicht nur die aus Steuern erbrachten finanziellen Mittel verstanden werden, sondern auch das Potential an Ideen, Engagement usw., das in Studenten, Mitarbeitern und Professoren ruht. Hierbei wird von der Überzeugung ausgegangen, daß Leistungsanreize in der Lehre so gestaltet sein sollten, daß sie dieses Potential aktivieren. Hier liegt auch der Unterschied zu den Konzepten bisher in Deutschland vorgenommener Evaluationen. Evaluation findet in Deutschland immer noch eher als "Prozeß der Selbstbeschau statt, bei dem Schlußfolgerungen

---

<sup>10</sup> Vgl. z.B. Interview mit Bundesbildungsministerin Bulmahn in: "Sächsische Zeitung" vom 8. 9. 1999, S. 27. (bzw. Fußnote 4 im Abschnitt 1.1 in dieser Arbeit).

aus den Ergebnissen zum Beispiel von Lehrveranstaltungsbewertungen in den meisten Fällen unterbleiben“ (Rindermann 1997: 43)<sup>11</sup>. In Ansätzen des Qualitätsmanagements wird dagegen davon ausgegangen, daß die Leistungsbewertung auch zur Steuerung und zur Setzung von Anreizen zum gewünschten Handeln verwendet wird.<sup>12</sup> Wie dies von Lehrenden in Bezug auf die verschiedenen Handlungsebenen innerhalb (und soweit eine solche Einschätzung möglich ist auch außerhalb) der Hochschule gesehen wird, soll in der Arbeit genauer untersucht werden.<sup>13</sup> Systematisch zu unterscheiden wäre hierbei nach Dozenten- bzw. Lehrveranstaltungsebene, nach Studiengangs-, Fachbereichs- bzw. Fakultätsebene, nach Hochschul- und ggf. Landesebene. Nach dem Überblick über Konzepte für Leistungsanreize und (soweit Erfahrungen hierzu zugänglich sind) den Effekten, die evtl. in anderen Ländern wie z.B. an britischen oder australischen Hochschulen bei ihrer Umsetzung erzielt wurden, wird sich dann der Frage zugewandt, inwieweit solche Konzepte für das deutsche Hochschulsystem adaptionsfähig sind und welche der Erfahrungen für Veränderungen an deutschen Hochschulen nützlich sein können.

#### *zu zweitens) Analyse der Leistungskriterien*

Ob Leistungsanreizsysteme die Ziele erreichen können, die man mit ihnen verknüpft, hängt nicht nur von der Angemessenheit der Konzepte ab, die ihnen zugrunde liegen. In der Diskussion, die seit einiger Zeit an deutschen Hochschulen geführt wird, spielen die Kriterien, mit denen man "messen" will, inwieweit die gewünschte Leistung erreicht wird oder nicht, eine große Rolle. Dies gilt nicht nur für die interne und externe Evaluation, sondern auch für andere Verfahren zur Qualitäts- wie auch der Leistungsbewertung. Vielfach wurde nicht nur diskutiert, welche Kriterien in die Bewertung einfließen sollen, sondern auch, wie valide und zuverlässig diese Kriterien sind - bzw. welche Ansprüche an ihre Zuverlässigkeit zu stellen sind. In der Arbeit soll deshalb nicht nur aus konzeptioneller, sondern auch aus methodischer Sicht und anhand empirischer Analysen diskutiert werden, welche Kriterien unter welchen Bedingungen inwiefern geeignet sind. Dies stellt – da die Diskussion von Leistungsanreizmodellen bisher häufig von Betriebswirtschaftlern v.a. aus dem Controlling-Bereich vorangetrieben wird, nach Ansicht des Verfassers ein bisher relativ wenig bearbeitetes Forschungsfeld dar, zumal z.T. auch noch nicht in ausreichender Menge oder Güte Daten für umfassendere empirische Analysen vorliegen. Hier könnte eine spezifische Kompetenz der empirischen Sozialwissenschaften wie der Soziologie evtl. stärker als bisher in die Diskussion eingebracht werden. So erscheint die Zuverlässigkeit

---

<sup>11</sup> Hierzu findet sich auch aktueller eine ähnliche Einschätzung in HRK (2004b: 105).

<sup>12</sup> Dies soll keineswegs heißen, dass Leistungsbewertung und Leistungsanreize gleichgesetzt werden (können). Allerdings sollte bei der Diskussion um Leistungsbewertungen

einiger Kriterien (wie zum Beispiel Hörerzahl, Anzahl der abgenommenen Prüfungen, Studiendauer usw. - vgl. u.a. Böhm 2000: 116) bei den bisher praktizierten Erhebungsmethoden in Deutschland durchaus fragwürdig. Andere Kriterien, die sich in mehreren empirischen Untersuchungen deutlich zuverlässiger als erwartet erwiesen haben, flossen bisher kaum in die Diskussion der Leistungsindikatoren ein (z.B. Ergebnisse von Absolventenbefragungen, unter bestimmten Bedingungen auch Studentenbefragungen)<sup>14</sup>.

Vor dem Hintergrund der recht großen Rolle, die diese Diskussion in Deutschland bei der Einführung von Leistungsanreizen in der Hochschullehre spielt, erscheint es außerdem interessant einzubeziehen, wie bestimmte Kriterien in anderen Ländern wie den USA, Großbritannien oder Australien bewertet werden, die schon länger Erfahrungen mit diesen haben und ob eventuell vorhandene methodische Mängel von Leistungsindikatoren als so schwerwiegend eingeschätzt werden, daß sie die Akzeptanz und damit letztlich auch die Wirksamkeit von Leistungsanreizsystemen beeinträchtigen können (vgl. z.B. Jackson 1997, Johns/ Taylor 1990).

### ***zu drittens) Analyse von Bedingungen erfolgreicher Institutionalisierung***

Gerade unter den Bedingungen rigider Sparpolitik der öffentlichen Haushalte und der Befürchtung seitens vieler Hochschullehrer in Deutschland, daß Evaluationen lediglich als Vorwand zur Durchsetzung weiterer Sparmaßnahmen verwendet werden könnten, ist die Frage von Bedeutung, unter welchen Bedingungen eine erfolgreiche Institutionalisierung von Leistungsbewertungen und Leistungsanreizen an Hochschulen erfolgen kann (vgl. Leszczensky/ Orr 2004).<sup>16</sup> In Deutschland gibt es hierzu bis auf das LARS-Programm (Pilotprojekt zur Einführung von Leistungsanreizen in der Lehre an den Fachhochschulen Baden-Württembergs)<sup>17</sup> kaum langjährige Erfahrungen. Verfolgt man jedoch die Diskussion in Hochschulen und Politik Deutschlands in den vergangenen 15 Jahren, so scheint es, daß neben solchen Bedingungen wie öffentlichem Interesse an der Thematik auch der Nachweis einer methodisch und „wissenschaftlich soliden“ Vorgehensweise eine der Bedingungen erfolgreicher Institutionalisierung sei. Es ist aber davon auszugehen, daß es weitere Bedingungen gibt.

---

<sup>13</sup> Untersucht werden soll hierzu auch, welche Auswirkungen sich auf das (Arbeits-)Klima und insbes. die Arbeitszufriedenheit an Hochschulen ergeben (vgl. hierzu auch Kopp/Weiß 1995: 118, Böhm 2000: 116 ff.)

<sup>14</sup> Vorschläge zur Einbeziehung von „student ratings“ in ein System leistungsorientierter Besoldung finden sich z.B. von Ulrich Karpen in Enders/ Teichler (1995: 91). Zur Zuverlässigkeit von Studentenbefragungen zur Qualität der Lehre vgl. insbes. Hage 1996, Rindermann 1996; ausführlich diskutiert in Kremmkow 1997, siehe auch Kremmkow/ Heldt 2000, Kremmkow/ Winter 2000 sowie Rindermann 2001.

<sup>16</sup> Zur Situation vor etwa 10 Jahren vgl. die länderübergreifende Synopse von Goedegebuure u.a. (1992: 48).

<sup>17</sup> Siehe hierzu z.B. die Ausführungen von Fischer/ Voss (1998).

So wird vermutet, daß das Vorhandensein oder Nichtvorhandensein bestimmter Machtverhältnisse ebenfalls Bedingungen erfolgreicher Institutionalisierung sind. Einige Autoren nennen hierzu externen Druck, z.B. seitens der Politik, als Erfolgsfaktor (vgl. Schenker-Wicki 1996). Hierbei wird allerdings auch darauf hingewiesen, daß es auf das richtige Maß ankommt.<sup>18</sup> Machtverhältnisse können aber noch an anderer Stelle von Bedeutung sein, beispielsweise bei der Beurteilung der Eignung von studentischen Lehrbewertungen. Hier sind diejenigen Experten, die darüber urteilen, als in der Lehre tätige Professoren meist gleichzeitig auch Betroffene. Bei positiver Beurteilung der Eignung könnte ein Interessenkonflikt auftreten, wenn hierbei die Machtkonstellation der eigenen Statusgruppe als gefährdet angesehen wird (vgl. Rindermann 2001).

Als wesentlicher Faktor für das Funktionieren von Leistungsanreizen wird von mehreren Autoren neben ausreichender Datenqualität, Transparenz und Nachvollziehbarkeit von Kriterien auch die grundsätzliche Akzeptanz bei den Betroffenen genannt, da ansonsten Unterlaufungsstrategien wahrscheinlich sind, in deren Folge die angestrebten Ziele nicht oder nur teilweise erreicht werden (vgl. z.B. Arnhold/ Schreiterer 2002, Tondorf/ Jochmann-Döll 2004). Dies gilt erst recht bei einer relativ starken Rechtsstellung und Existenzsicherung von Systemangehörigen (vgl. Wottawa 2001: 154). Da es hier nicht nur um die Implementation von Verfahren zur Bewertung von Leistungen und Leistungsanreizen geht, sondern eine „Veränderung in den Köpfen“ als notwendig angesehen wird, damit das Ziel der Maßnahmen erreicht werden kann, soll in dieser Arbeit der Aspekt der Akzeptanz im Vordergrund stehen.

Auf längere Sicht können auch die Auswirkungen auf die Arbeitszufriedenheit eine Rolle spielen. Der Zusammenhang von Arbeitszufriedenheit und Akzeptanz wird hierbei darin gesehen, daß Reformmaßnahmen, die dazu führen, daß Lehrende mit ihrer beruflichen Situation insgesamt unzufriedener sind, auch schwerer Akzeptanz finden dürften. Umgekehrt könnten Reformmaßnahmen, die vielleicht zunächst keine große Akzeptanz finden, aber auf längere Sicht zu einer höheren Arbeitszufriedenheit führen, langfristig auch eher Akzeptanz finden. Mit Hilfe von empirischen Analysen zur Akzeptanz und zur Arbeitszufriedenheit soll schließlich auch nach Ansatzpunkten für akzeptanzfördernde Maßnahmen gesucht werden, bzw. nach Möglichkeiten, Leistungsbewertung und Leistungsanreize so zu gestalten, daß sie grundsätzlich Akzeptanz finden können.

---

<sup>18</sup> So führt Pellert (1999: 298) als Hintergrund (warum dies überhaupt möglich war) für einen erfolgreichen Fall von Maßnahmen der Organisationsentwicklung an, daß es eine vorangehende heftige Diskussion über den Weiterbestand der Einrichtung gab. Dadurch sei die „Einsicht in den Nutzen von organisationsentwicklerischen Maßnahmen erhöht“ worden. Im Zuge von Überlegungen zur Übertragbarkeit des Projektes fügt sie hinzu, daß das Interesse an Veränderung v.a. von innen kommen müsse, aber auch „leichter Außendruck hilfreich“ wäre (vgl. Pellert 1999: 304ff.).

### 1.3 Zum methodischen Vorgehen

Im ersten Teil, dem Überblick über Konzepte, auf deren Grundlage die Veränderungsprozesse hinsichtlich der Leistungsbewertung und der Leistungsanreize an den Hochschulen in Deutschland stattfanden bzw. noch stattfinden, soll dies vor allem anhand der vorliegenden wissenschaftlichen Literatur und der Schriftstücke aus der Hochschulpolitik (z.B. sogenannte "white papers") erfolgen. Daneben soll jedoch auch eine Einordnung und ggf. vergleichende Diskussion der Papiere anhand der weiteren Entwicklung und anhand von Experteninterviews<sup>19</sup> vorgenommen werden.

Diese Expertenbefragungen sollen neben der Darstellung und Diskussion empirischer Untersuchungen von Leistungskriterien auch eine Grundlage für den zweiten Teil der Arbeit sein. Da davon ausgegangen wird, daß neben der Zuverlässigkeit und Validität von Leistungskriterien für die Einführung und Kopplung mit Anreizsystemen vor allem die Akzeptanz bei den zu Bewertenden bzw. deren Einschätzung durch Funktionsträger/ Experten von Bedeutung ist, sollen auch diese einbezogen werden. In Experteninterviews werden deshalb insbesondere mit der Einführung von Leistungsbewertungssystemen befaßte Vertreter/innen bzw. Personen der Hochschulleitungen und der Personalvertretungen der Hochschulen, Vertreter/innen der Ministerien, und hochschulübergreifender (Hochschulforschungs-)Institutionen zu Wort kommen. Nur durch diese relativ breite Auswahl kann bei den im Rahmen dieses Vorhabens maximal zu leistenden Interviews ein Mindestmaß an Vielfalt der Perspektiven abgesichert werden. Um die Akzeptanz der Kriterien bei den zu Bewertenden selbst und die Auswirkungen auf das Arbeitsklima einschätzen zu können und evtl. auch für Experten unvermutete neue Informationen zu erlangen<sup>20</sup>, erscheint es aber unumgänglich, neben den Experten auch die Betroffenen direkt zu befragen.<sup>21</sup> Diese Befragung soll, um weitgehende Vergleichbarkeit der Ergebnisse untereinander zu gewährleisten, als standardisierte Befragung mit überwiegend geschlossenen Antwortvorgaben erfolgen. Diese Methode hat sich sowohl bei Mitarbeiterbefragungen (durch Leitungen von Institutionen) als auch bei (gewerkschaftlichen) Befragungen zu strittigen Themen als zweckmäßig herausgestellt, da sie zuverlässigere Aussagen ermöglicht und für Befragte wie Auswertende weniger Zeit in Anspruch nimmt. Für weitergehende Kommentare und Hinweise ist aber ausreichend Raum vorgesehen. Eine solche Erhebung mit

---

<sup>19</sup> Zur Bedeutung und Zielstellung der Expertenbefragungen vgl. Meuser/ Nagel 1994 und 2003.

<sup>20</sup> Hochschullehrerbefragungen (Vgl. Kopp/ Weiß 1995) ergaben z. T. eher unvermutete Ergebnisse.

<sup>21</sup> Eine direkte Befragung wird auch bei Mitarbeiterbefragungen in Unternehmen als unumgänglich angesehen. Die tatsächliche Akzeptanz von Bewertungskriterien wäre hier nicht allein durch Befragung von Arbeitgeber- und Gewerkschaftsvertretern als „Experten“ feststellbar (vgl. Bartel-Lingg 1996).

mehreren hundert Fragebögen war im Rahmen dieser Arbeit nur mittels moderner Scanner-technik sowie mit institutioneller Unterstützung von Hochschulleitungen und -organisationen ohne zusätzliche Mittel durchführbar.<sup>22</sup> Dennoch birgt diese Methode der Befragung trotz relativ gesehen weniger Aufwand und insgesamt höherer Zuverlässigkeit generell die Gefahr einer geringeren Tiefenschärfe in sich. Deshalb werden Experteninterviews und standardisierte schriftliche Befragungen mit dem Ziel einer gegenseitigen Ergänzung der jeweiligen Stärken und Schwächen eingesetzt. Weitere, detailliertere Ausführungen zum methodischen Vorgehen, zur Durchführung und Auswertung der Experteninterviews erfolgen im vierten Kapitel dieser Arbeit. Ähnlich gilt dies für die standardisierte schriftliche Befragung der zu Bewertenden, die im fünften und sechsten Kapitel ausführlicher thematisiert wird.

---

<sup>22</sup> Zu Erfahrungsberichten mit der entsprechenden Technik vgl. Krempkow/ Heldt 1999, 2000.

## **2. Was ist Leistung, was ist Qualität? – Klärung von unterschiedlichen Begriffsverständnissen**

*Die Begriffe Leistungsbewertung und Leistungsanreize werden in den letzten Jahren in zahlreichen hochschulpolitischen Reden, vielen Zeitungsartikeln und Medienbeiträgen zum Bereich Hochschule und Bildung und auch im Alltag an den Hochschulen häufig verwendet. Dabei wird in Diskussionen häufig ebenfalls deutlich, daß damit von den Diskussionsparteien vielfach nicht dasselbe gemeint ist. Einige unterschiedliche Begriffsverständnisse herauszuarbeiten und den Begriff für die nachfolgenden Ausführungen zu klären, ist das Anliegen dieses Abschnittes. Dabei wird in diesem Zusammenhang auch auf die Begriffe Evaluation und Qualität einzugehen sein, kurz deren jüngere Anwendungsgeschichte zu umreißen und diese Begriffe für verschiedene Ebenen, Ziele und Zielgruppen, für die sie Anwendung finden, zu differenzieren sein.*

Gliederung:

- 2.1 Begriffsklärungen: Leistungsbewertung, Leistungsanreize und Hochschullehre
  - 2.1.1 Leistungsbewertung
  - 2.1.2 Leistungsanreize
  - 2.1.3 (Hochschul-)Lehre
- 2.2 Evaluation und Qualitätsbewertung
  - 2.2.1 Zum Begriff Evaluation
  - 2.2.2 Die Entwicklung des Qualitätsbegriffs in der „Dienstleistungsgesellschaft“
- 2.3 Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen
  - 2.3.1 Definition von Lehrqualität auf Hochschulebene
  - 2.3.2 Definition von Lehrqualität auf Studiengangsebene
  - 2.3.3 Definition von Lehrqualität auf Lehrveranstaltungsebene
  - 2.3.4 Zielgruppen für die Qualität der Hochschullehre
  - 2.3.5 Fazit zur Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen

## 2.1 Zur Klärung der Begriffe Leistungsbewertung, Leistungsanreize und Hochschullehre

### 2.1.1 Leistung und Leistungsbewertung

Die Leistung und Leistungsbewertung sowie Qualität und Qualitätsbewertung werden häufig in denselben Zusammenhängen verwendet. „Es ist von Leistung oder Qualität die Rede – selbst dann, wenn es um Effektivität oder Effizienz im Hinblick auf sehr unterschiedliche Ziele geht“ konstatiert Teichler (2003: 5). Eine einheitliche Begriffswelt gäbe es nicht und ein Konsens, was Leistung und was Qualität ist, sei schwer zu erzielen.

Dennoch sollte, um zu verdeutlichen, um welche Ziele es im Zusammenhang mit den hier verwendeten Begriffen geht, eine Klärung dessen versucht werden, was mit Leistung und was mit Qualität gemeint ist. Denn nur wenn die Ziele klar sind, kann eine Leistungsbewertung oder Evaluation erfolgreich sein – darauf wird in der Forschung zu Hochschulen immer wieder hingewiesen. Zunächst sollen daher die Begriffe Leistung und Leistungsbewertung, in nachfolgenden Abschnitten auch Qualität und die Bewertung von Qualität für diese Arbeit geklärt werden.

Allgemeinsprachlich wird nach Meyers Lexikon unter *Leistung* der „Grad einer körperlichen und geistigen Beanspruchung sowie deren Ergebnis“ verstanden (vgl. Strych/ Weiß 1998) oder nach Brockhaus Deutsches Wörterbuch das „Ergebnis einer Anstrengung, eine ausgeführte Arbeit oder Tat“ (vgl. Wahrig 2004). Während erstere Definition auch den Aspekt der Belastung enthält, zielt die zweite stärker auf das Ergebnis. Je nach Betonung kann aber allgemein auch eine besondere Leistung im Sinne einer *besonders guten* Leistung gemeint sein. Schon in der Umgangssprache kann sich die Bedeutung des Begriffes Leistung also deutlich unterscheiden. In der Wissenschaft gibt es darüber, was als Leistung verstanden werden kann, erst recht sehr verschiedene Auffassungen. An dieser Stelle sollen nur drei Beispiele für Begriffsdefinitionen genannt werden. Diese mögen illustrieren, daß manche Missverständnisse oder Konflikte bereits aus einem anderen Grundverständnis des Begriffes Leistung resultieren können.

In den Naturwissenschaften wird Leistung als Arbeit pro Zeiteinheit; oder auch als Produkt aus Kraft und Geschwindigkeit definiert. In Meyers Lexikon findet sich außerdem als zusätzliche Definition für komplexere Fälle: „Ist die während gleicher Zeitintervalle geleistete Arbeit nicht konstant, so wird die Leistung zu jedem Zeitpunkt durch die zeitliche Ableitung der Arbeit gegeben“ (vgl. Strych/ Weiß 1998).

Nach einer rechtswissenschaftlichen Definition wird Leistung dagegen definiert als jede rechtliche Handlung, die eine Vermögensverschiebung in Geld, Sachen, Diensten u. a. bewirkt. Im Palandt (2000) (BGB § 362, Abs. 1) heißt es hierzu: „Unter Leistung ist nicht die Leistungs-

handlung, sondern der Leistungserfolg zu verstehen“. Daher sei derjenige, der die Leistung zu erbringen hat, „u. U. zur Wiederholung der Leistungshandlung verpflichtet“ (vgl. Palandt 2000). In der Betriebswirtschaft kann Leistung gelten als „die Menge (Leistungseinheiten, Output) oder der Wert (Geldgröße, Ertrag) der innerhalb eines Zeitraumes im betrieblichen Produktionsprozeß hervorgebrachten Sachgüter (Leistungsergebnis) oder der bereitgestellten Dienstleistungen“ (vgl. Strysch/ Weiß 1998). Im speziellen betriebswirtschaftlichen Bereich Controlling öffentlicher Verwaltungen, dessen Erfahrungen bei der Einführung z.B. von Kosten- und Leistungsrechnung an Hochschulen bzw. in der Steuerungsdebatte an Hochschulen zugrunde gelegt werden, wird Leistung auch verstanden als „das Ergebnis (Produkt/ Kostenträger), das zur Erfüllung einer vorgegebenen öffentlichen Aufgabe dienen soll“ (vgl. Seeger u.a. 1998: 269). Andere Autoren versuchen gar nicht erst, eine Definition von Leistung zu liefern, sondern zeichnen lediglich nach, welche Veränderungen sich bei den Indikatoren vollziehen, die als Leistungsgrößen zugrunde gelegt werden (vgl. Bahnmüller 2001: 429).

Während für die physikalische Definition „nur“ ein Maßsystem und ein geeignetes Messgerät notwendig erscheint, wird bei der rechtswissenschaftlichen Definition der Vertragsaspekt einer (wirtschaftlichen) Leistung betont, wobei außerdem der Leistungserfolg hervorgehoben wird. Die rechtswissenschaftliche Definition kann bereits das Element einer Bewertung z.B. in einem Marktverhältnis enthalten. Hier gerät der Begriff der Leistung schon in die Nähe des zusammengesetzten Begriffes Leistungsbewertung. Der Leistungsbegriff aus dem Controlling ist noch enger finanzwirtschaftlich definiert, zielt aber ebenfalls auf das Ergebnis und erfasst Elemente einer Bewertung.

Für die nachfolgenden Ausführungen kommen zur klareren Abgrenzung dieser Begriffe von Leistung nur Definitionen in Betracht, die explizit Bezug auf eine Zeiteinheit nehmen (z.B. die allgemeine naturwissenschaftliche oder die allgemeine betriebswirtschaftliche). Dies erscheint notwendig, da angesichts der heutigen Innovationsgeschwindigkeit die Entwicklung z.B. eines Studienangebotes in den nächsten 3 Jahren sehr sinnvoll, in beispielsweise 10 oder 15 Jahren aber bereits nicht mehr angemessen und ineffektiv sein kann. Deshalb soll der Zeitaspekt Berücksichtigung finden. Leistung könnte demnach vorerst gelten als „die (von den Hochschulen bzw. ihrem Personal) geleistete Arbeit, gemessen an ihrer Menge (Leistungseinheiten, Output) pro Zeiteinheit“. Diese Begriffsdefinition erfasst aber lediglich Quantitäten. Sie wäre also z.B. durch die Erfassung der „Stückzahl“ an Absolventen innerhalb von 5 Jahren bereits erfüllt, ohne dabei deren Qualität zu berücksichtigen. Eine solche Leistungsdefinition und eine entsprechende Erfassung von Leistung würde aber die Gefahr einer Fehlsteuerung in sich bergen (vgl. z.B. Seeger u.a. 1998: 211).

Hier wird daher dafür plädiert, bei der Betrachtung von Leistung (bzw. der von den Hochschulen/ ihrem Personal) geleisteten Arbeit nicht nur die Menge pro Zeiteinheit, sondern immer auch deren Qualität einzubeziehen. Dies erscheint für eine realitätsnahe Information über die Leistungen der Hochschulen notwendig, erst recht für deren Steuerung. Leistung kann deshalb gelten als „die (von den Hochschulen bzw. ihrem Personal) geleistete Arbeit, gemessen an ihrer Menge (Leistungseinheiten, Output) pro Zeiteinheit *und Qualität*“. Dies entspricht dem Verständnis, dass Qualität an Hochschulen durchaus unabhängig von der Menge (z.B. an Absolventen, wiss. Publikationen) pro Zeiteinheit betrachtet werden kann. Leistung ohne Berücksichtigung von Qualität wird nach diesem Verständnis nicht als sinnvoll angesehen, da sie wie o.g. die Gefahr der Fehlsteuerung enthält.

Wenn im Leistungsbegriff zudem explizit Elemente der Bewertung enthalten sind, soll nachfolgend von *Leistungsbewertung* die Rede sein. Damit ist nun auch der Begriff der *Bewertung* zu definieren. In den Wirtschaftswissenschaften wird Bewertung z.B. definiert als „betriebswirtschaftliche Regel, nach der Vermögensgegenständen Geldbeträge zugeordnet werden. Die Höhe des zugeordneten Geldbetrags (*Wertansatz*) richtet sich nach dem Zweck der Bewertung. Die wichtigsten Anwendungsgebiete der Bewertung sind die Buchhaltung und Bilanzierung, die Kostenrechnung und die Investitionsrechnung.“ (vgl. Wahrig 2004). Ähnlich findet sich dies auch bei Strych/ Weiß (1998).

Nachfolgend soll einer solchen betriebswirtschaftlichen Definition nicht gefolgt werden, da sie nicht explizit den Qualitätsaspekt enthält. Vielmehr soll der Begriff Bewerten stärker im sozialwissenschaftlichen Sinne verwendet werden. In diesem Sinne wird häufig für den Begriff Bewerten bzw. Bewertung auch der Begriff Evaluieren bzw. Evaluation verwendet. Da dieser Begriff vor allem für Qualitätsbewertungen im Hochschulbereich üblich geworden ist, die Begriffe Leistung und Qualität hierbei jedoch ebenfalls häufig im Zusammenhang und z.T. auch nicht klar voneinander abgegrenzt verwendet werden, soll der Begriff Evaluation im nachfolgenden größeren Abschnitt noch näher erläutert werden. Zunächst sind aber noch die Begriffe (Leistungs-)Anreize und (Hochschul-)Lehre zu klären, da sich die vorliegende Arbeit auf diesen Themenbereich konzentriert.

### **2.1.2 Leistungsanreize**

Unter einem Anreizsystem, also einem System, das Anreize enthält, sind nach Meyers Lexikon (vgl. Strysch/ Weiß 1998) allgemeinsprachlich „Gratifikationsmaßnahmen, um Arbeitnehmer zum Verbleiben im Unternehmen oder zur Steigerung der Arbeitsleistung zu bewegen“ zu verstehen. Abgesehen von der Einschränkung auf Arbeitnehmer, die auf Beamte und/ oder häufig freiberuflich und damit im Prinzip als Selbständige arbeitende Privatdozenten weniger zutrifft, soll dieser allgemeinen Begriffsdefinition zunächst gefolgt werden, da sie für die direkt folgenden Abschnitte zum Verständnis genügen dürfte. Auf verschiedene Leistungsanreizmodelle wird im Zusammenhang mit der Kopplung von Leistungsbewertung und Leistungsanreizen (zu einer Grundlegung vgl. Ziegele/ Handel 2004) noch detaillierter eingegangen.

### **2.1.3 (Hochschul-)Lehre**

Allgemeinsprachlich wird nach Meyers Lexikon unter Lehre eine Lehrmeinung oder ein wissenschaftliches System verstanden (vgl. Strysch/ Weiß 1998). Nachfolgend soll dies insbesondere in Abgrenzung zum Aufgaben- bzw. Leistungsbereich Forschung genauer definiert werden, um deutlich zu machen, auf welchen Bereich sich diese Arbeit bezieht.

Das Hochschulrahmengesetz (HRG) stellt die beiden Aufgabenbereiche Lehre und Forschung in den Vordergrund. Demnach soll die Lehre (und das Studium) „den Studenten auf ein berufliches Tätigkeitsfeld vorbereiten und ihm die dafür erforderlichen fachlichen Kenntnisse, Fähigkeiten und Methoden dem jeweiligen Studiengang entsprechend so vermitteln, daß er zu wissenschaftlicher oder künstlerischer Arbeit und zu verantwortlichem Handeln in einem freiheitlichen, demokratischen und sozialen Rechtsstaat befähigt wird“ (vgl. HRG § 7, Stand vom 27.12.2004). Damit enthält das Aufgabenverständnis von Lehre auch den Grundgedanken der Förderung demokratischer Tugenden und fachunabhängiger Kompetenzen. Die Forschung „dient der Gewinnung wissenschaftlicher Erkenntnisse sowie der wissenschaftlichen Grundlegung und Weiterentwicklung von Lehre und Studium“ (vgl. HRG §22). Von Hochschulforschern wird neben den Bereichen Forschung und Lehre außerdem beispielsweise zwischen den Aufgaben- bzw. Leistungsbereichen Nachwuchsförderung, interne und externe Dienstleistungen, Weiterbildung<sup>24</sup> usw. unterschieden (vgl. Pasternack 2004: 15, ähnlich: Enders/ Teichler 1995: 18).

---

<sup>24</sup> Nur kurz erwähnt werden soll, daß in den vergangenen Jahren (wieder) stärker die Weiterbildung an Hochschulen in den Fokus der Aufmerksamkeit geriet. Der als privatwirtschaftlich einzuordnende Weiterbildungssektor kann, auch wenn er an Hochschulen angesiedelt ist, nach dem hier verwendeten Begriffverständnis nicht ohne weiteres einbezogen werden. So weist z.B. Sauter (1995: 135) darauf hin, daß sich der (private) Weiterbildungsbereich von anderen Teilen des Bildungssystems wesentlich unterscheidet.

Eine Diskussion detaillierterer Aufgliederungen soll nicht erfolgen, da auch sie die Unterscheidung zwischen Lehre/ Forschung enthalten und es hier v.a. auf diese Abgrenzung ankommt. Wichtig zum Verständnis dieser Arbeit ist, daß der Aufgabenbereich der Lehre in Anlehnung an die Definition des HRG eher weit gefaßt werden soll. Unter den Begriff (Hochschul-)Lehre zusammengefaßt werden damit nicht nur die Lehrveranstaltungen, sondern auch z.B. deren Abstimmung aufeinander im gesamten Studienprogramm und dessen Rahmenbedingungen (siehe ausführlicher hierzu auch nachfolgende Ausführungen zum Begriff *Qualität* der Lehre). Ein breites Verständnis vom Aufgabenbereich Lehre schließt außerdem solch traditionell im Aufgabenverständnis der Hochschulen verankerte bereichsübergreifende Ziele wie kognitive und methodische Rationalität und Förderung der intellektuellen Neugierde ein, wenngleich diese hier nicht im Vordergrund stehen können.

## **2.2 Evaluation und Qualitätsbewertung**

### **2.2.1 Zum Begriff Evaluation**

Alltagssprachlich wird an den Hochschulen mit „Evaluation“ meist eine Fülle von Methoden und Maßnahmen zur Feststellung wie auch Verbesserung der Lehrqualität umschrieben. Allerdings hat sich das Alltagsverständnis im Laufe der letzten Jahre ausgeweitet, so daß dieser Begriff inzwischen nicht mehr nur für Wissenschaftseinrichtungen, sondern z.B. auch schon einmal für Handytarife verwendet wird

Seit Anfang der 1990er Jahre sei dieser Begriff zu einem „Modewort“ avanciert, das sich durch „inflationären, oft unangemessenen Gebrauch“ auszeichnet, meinen denn auch einige Autoren in wissenschaftlichen Publikationen (vgl. Kellermann 1992: 142, ähnlich in Müller-Böling 1996: 1, Spiel 2001: 9). Andererseits fordern Parlament, Bundesrechnungshof und auch Nichtregierungsorganisationen in den letzten Jahren immer stärker die Durchführung von Evaluationen und auch in Programmen der EU wird die Durchführung von Evaluationen zunehmend zum Standard (vgl. Stockmann 2002a: 2 ff.). Allerdings wird dieser Begriff in der Politik z.T. mit sehr unterschiedlichen Bedeutungen versehen. Auch deshalb erscheint es notwendig, dies genauer zu definieren. Allgemeinsprachlich verstanden meint der Begriff zunächst einmal nur „Bewertung; Beurteilung“ (Duden 2000), oder – wenn man fremdsprachliche Wurzeln unterstellt – französisch *évaluer*: abschätzen, berechnen; englisch *evaluate*: zahlenmäßig bestimmen, auswerten) Auch Schmitz (1995: 40) nennt eine relativ allgemeine Definition als „sach- und fachgerechte *Bewertung* eines Sachverhalts“. Angesichts der Vielfalt der Maßnahmen und Methoden, die mit diesem Begriff bezeichnet werden, erscheint dies zwar im weiteren Sinne

zutreffend, jedoch noch zu ungenau. Werden doch mit Evaluation sowohl die Bewertung wissenschaftlicher Einrichtungen durch Expertengremien, als auch in ökonomischen Kontexten Effizienzmessungen, sowie in Zusammenhang mit der Qualitätsdiskussion zur universitären Lehre Studentenbefragungen zusammengefaßt (vgl. Kromrey 1995: 313-316). Dabei beschreiben alle diese mit dem Begriff Evaluation bezeichneten Maßnahmen höchst unterschiedliche Ziele und Methoden. Inzwischen gebe es eine schwer eingrenzbare Vielfalt von Begriffsverwendungen und dahinter stehenden Denk- und Handlungskonzepten (vgl. Kromrey in Spiel 2001: 21).

Kromrey verwies daher auf eine Definition des Arbeitskreises „Wirkungsanalysen und Erfolgskontrolle in der Raumordnung“ (1984: 30). Dort heißt es: „Evaluation kann als die Aufgabe bezeichnet werden, Planungen, Programme, Projekte und Maßnahmen zu bewerten.“ und fügte hinzu, daß Evaluation nicht von „irgend jemand“ durchgeführt wird, sondern von Personen, die in besonderer Weise dazu befähigt erscheinen. Außerdem betont er, daß Evaluation „systematisch“, nach „zuvor präzise festgelegten“ und „explizit auf den zu bewertenden Sachverhalt bezogenen Kriterien“ durchgeführt werden muß. Ähnlich definierten Wottawa & Thierau (1990: 9) den Begriff Evaluation als „systematische Beurteilung eines Programmes oder Produktes“. Nur leicht abgewandelt findet sich diese Definition auch im „Handbuch der Evaluationsstandards“ (vgl. Beywl 2000: 25): Danach ist Evaluation „die systematische Untersuchung der Verwendbarkeit oder Güte eines Gegenstandes. (...) Zu den Evaluationsgegenständen gehören Programme, Projekte...“ usw. In ihrer Neuauflage des Lehrbuches Evaluation formulieren Wottawa & Thierau (1998: 14) dagegen: „In Anbetracht der bestehenden Definitionsvielfalt ist es zweckmäßiger, statt einem weiteren Definitionsversuch die allgemeinen Kennzeichen wissenschaftlicher Evaluation herauszuarbeiten.“ Verkürzt dargestellt, gehört zum durch die Wortwurzel gegebenen Konsens, daß erstens „solche Tätigkeiten etwas mit Bewerten zu tun haben“, daß sie zweitens „ziel- und zweckorientiert“ sind (um etwas zu überprüfen, verbessern oder zu entscheiden) und daß sie drittens „dem aktuellen Stand wissenschaftlicher Techniken und Forschungsmethoden angepasst sein sollten“. An solchen relativ allgemeinen Beschreibungen wird jedoch kritisiert, daß sie praktisch für alles zutreffen und deren Gemeinsamkeit allenfalls als Leerformel ausdrückbar ist (vgl. Spiel 2001: 21ff.).

Wottawa und Thierau (1998: 13) erläutern allerdings auch die Unterscheidung in die Begriffe Evaluation und Evaluationsforschung, wobei sie nur letzterem die „explizite Verwendung wissenschaftlicher Forschungsmethoden und -techniken“ zuschreiben. Bortz/ Döring (2002: 102) nennen eine ähnliche Definition von Evaluationsforschung, ohne allerdings eine Definition

von Evaluation zu formulieren. Wottawa/ Thierau schätzen hierzu jedoch ein, daß sich dieser Systematisierungsversuch in Evaluation und Evaluationsforschung nicht durchsetzen konnte.<sup>25</sup>

Der Wissenschaftsrat (1996: 16) unterscheidet schließlich zwischen Evaluation als der „*Messung* von Qualität“ (nach objektiven bzw. objektivierbaren Indikatoren oder standardisierten Leistungsparametern) einerseits und der „*Bewertung* von Qualität“ (aufgrund subjektiv geprägter Expertenurteile<sup>26</sup>) andererseits.

Wenn nachfolgend fachwissenschaftlich von Evaluation die Rede ist, soll dies in Anlehnung an die die Definition des Wissenschaftsrates und Kromrey geschehen, der diese Überlegungen in einer relativ einfachen, prägnanten Weise zusammenfaßte: als „methodisch kontrollierte Bewertung (...) von Qualität“.<sup>27</sup> Hierbei soll im Anschluß an die vorhergehenden Überlegungen zum Thema Leistungsbewertung der subjektive Aspekt als immanent aufgefaßt und daher dieser betont werden. Es wird bewußt nicht explizit in einen subjektiven Aspekt und einen objektiven Aspekt gemäß der Definition des Wissenschaftsrates unterschieden. Da unter methodisch kontrolliert im wissenschaftlichen Kontext nach Stockmann (2002b: 219) verstanden wird, daß „empirische Methoden zur Informationsgewinnung und systematische Verfahren zur Informationsbewertung anhand offen gelegter Kriterien verwendet (werden), die eine intersubjektive Nachprüfbarkeit möglich machen“, kann die Bewertung als hinreichend objektiviert gelten. Um Begriffsverwirrungen vorzubeugen, wird im Folgenden der Begriff „Evaluation“ nur noch in seiner fachwissenschaftlichen Bedeutung verwandt.

Eine Erläuterung der Evaluation als Instrument der Qualitätsorientierung erfolgt im entsprechenden Abschnitt im dritten Kapitel dieser Arbeit.

### **2.2.2 Die Entwicklung des Qualitätsbegriffs in der „Dienstleistungsgesellschaft“**

Um einer Definition des häufig als sehr diffus beschriebenen Qualitätsbegriffes näher zu kommen, soll dessen jüngste historische Entwicklung in groben Zügen nachgezeichnet werden.<sup>28</sup>

---

<sup>25</sup> Weitere Systematisierungsversuche nach den theoretischen Konzeptionen von Evaluation nennen Lames (2000: 9 ff.), die jedoch ebenfalls nicht allgemein anerkannt sind (vgl. Bank 2000: 50) und daher nicht weiter diskutiert werden sollen.

<sup>26</sup> Ob Studierende als „Experten“ zu sehen sind, wird z.T. strittig diskutiert. Insgesamt sprächen jedoch mehr Argumente für als gegen die Einbeziehung der Studierenden „als Experten für die Lehrbeurteilung“. (vgl. Pasternack 2004:74).

<sup>27</sup> Ursprünglich sprach Kromrey (1995: 313ff.) von Evaluation als „methodisch kontrollierte Bewertung im Sinne der *Messung* von Qualität“. Später vermied Kromrey die Formulierung *im Sinne der Messung* und erweiterte seine Definition in Evaluation als „methodisch kontrollierte, verwertungs- und bewertungsorientierte Form des Sammelns und Auswertens von Informationen“ (vgl. Kromrey 2001: 112).

<sup>28</sup> Bartel-Lingg (1996: 27) weist darauf hin, daß der Begriff lateinischen Ursprungs (*qualitas*) und älter als 2000 Jahre ist. Die Diskussion über den Begriffsinhalt sei mindestens ebenso alt. Auf den historischen Hintergrund des Begriffs kann hier jedoch nicht näher eingegangen werden.

Die Ausgangslage ist – wie bereits in der Einleitung dargestellt wurde, daß Anfang der 1990er Jahre die jüngste Debatte um die Qualität an den Hochschulen nicht unerheblich von studentischen Initiativen und Medien angestoßen wurde. Zudem kamen aufgrund der Finanzkrise der öffentlichen Haushalte die Hochschulen in den letzten Jahren mehr und mehr in Legitimationszwänge. Die Hochschulen mußten sich stärker als Dienstleister der Gesellschaft verstehen – so oder ähnlich lauten Forderungen aus Politik und Medien. Zumindest der Anlaß für die Qualitätsdebatte kann also in studentischer Unzufriedenheit und steigendem Legitimationsdruck gesehen werden. Die Frageentsteht jedoch, ob dies alleinige Ursachen für den seit Jahren anhaltende Entwicklungsprozeß sein können und inwieweit insbesondere der in den letzten Jahren zunehmende Transfer von Konzepten aus der Wirtschaft hiermit erklärt werden kann.

Ohne hierauf an dieser Stelle ausführlicher eingehen zu wollen - es gibt unbestreitbar deutliche Unterschiede zwischen der Situation an den Hochschulen und der Situation bei Einführung von Qualitätsmanagementsystemen in großen Unternehmen (vgl. HRK 1998a: 74) - wird an den Hochschulen seit mehreren Jahren die Übertragbarkeit solcher Qualitätsmanagementsysteme diskutiert. Hierfür muß es, bei allen zu beachtenden Systemunterschieden, Gründe geben. Wolter (1999: 26) spricht im Zusammenhang mit Schlüsselbegriffen der zweiten `großen` Novellierung des Hochschulrahmengesetzes 1998 (wie Wettbewerbsfähigkeit, Effektivität und Effizienz sowie leistungsorientierte Steuerung) davon, daß deren `Philosophie` weitgehend dem „neuen betriebswirtschaftlichen Paradigma der Hochschulsteuerung“ entspricht. Einerseits können die Gründe also in einer Annäherung der für den Hochschulbereich maßgeblichen von der Öffentlichkeit formulierten Erwartungen an die in der Wirtschaft gesehen werden. Andererseits ist es aber auch möglich, daß sich die in der Wirtschaft existierenden Ideen (bzw. in diesem Fall die Vorstellungen davon, was Qualität ist) verändert haben, so daß sie leichter für den Hochschulbereich adaptierbar wurden.

In der Tat hat sich die Definition des Qualitätsbegriffs auch in der Wirtschaft verändert (vgl. HRK 2001a: 31). Zwar gibt es selbst in der Betriebswirtschaftslehre, die den Qualitätsbegriff im Rahmen ihrer Qualitätsmanagementkonzepte unablässig nutzt, keine griffige Spezifikation (vgl. Stockmann 2002b: 209). Dennoch lassen sich einige wesentliche Aspekte festhalten: Im Mittelpunkt steht also tendenziell weniger das Produkt und stattdessen stärker die Kundenzufriedenheit mit dem Produkt. Unternehmen konkurrieren zunehmend nicht mehr nur über Produktpreis und Produktqualität, sondern immer stärker auch über den Service rund um das Produkt, also über die Dienstleistungsqualität. (Malorny 1997: 96-105). Dies hat Auswirkungen

auch auf Teile des Dienstleistungsbereichs, die – wie z.B. das Bildungswesen – nicht primär eine wirtschaftliche Absicht verfolgen (vgl. Gnahs u.a. 1995: 1).<sup>29</sup>

Diese Entwicklung wurde auch in einer international als Standard bzw. Norm akzeptierten Begriffsdefinition berücksichtigt. Demnach ist Qualität „die Gesamtheit von Merkmalen (und Merkmalsausprägungen) bezüglich ihrer Eignung, festgelegte und vorausgesetzte Erfordernisse zu erfüllen“ (vgl. DIN EN ISO 8402). In dieser Begriffsauffassung sei zwar noch der ursprüngliche Inhalt der einwandfreien Funktion und Freiheit von Fehlern enthalten, doch nicht abstrakte Normen stellen den Qualitätsmaßstab dar, sondern „die Erfüllung von Bedürfnissen und Erfordernissen, die durch den Kunden festgelegt werden“ (Vgl. Stockmann 2002b: 210). Stockmann beschreibt dies auch als Gedanken des „fitnes for use“, also der Nützlichkeit und Verwendbarkeit eines Produktes, der zunehmend in den Mittelpunkt des Qualitätsverständnisses rückte. In den letzten Jahren hat die Qualitätsdiskussion in Unternehmen außerdem stark Unternehmensabläufe und Prozesse als Voraussetzung für erfolgreiches Bestehen am Markt fokussiert.

Auch Hochschulleitungen versprechen sich von der Einführung von Elementen des Qualitätsmanagements unter anderem eine bessere Steuerbarkeit von Entscheidungsprozessen in der Hochschule und kürzere Reaktionszeiten auf Veränderungen der Rahmenbedingungen, z.B. bei der Konzeption von Studienangeboten. Dies sind auch Erwartungen, die im Zusammenhang mit Stichworten wie Globalisierung, Differenzierung, Konzentration und Wettbewerb sowie Autonomie von der Politik an die Hochschulen herangetragen werden (vgl. HRK 2000c: 7ff.).

Da diese bessere Steuerbarkeit von Unternehmensprozessen mit dem Ziel, schneller auf sich verändernde Rahmenbedingungen (z.B. „Kundenbedürfnisse“) reagieren zu können, auch eines der Ziele bei der Einführung von Qualitätsmanagementsystemen wie Total Quality Management (TQM) in der Wirtschaft war (vgl. ausführlicher hierzu entsprechenden Abschnitt im dritten Kapitel), wurden auf Tagungen der Hochschulrektorenkonferenz mehrfach auch Fachleute aus der Wirtschaft eingeladen. Diese verfügten demnach sowohl über Hochschulerfahrungen, als auch über Erfahrungen in der Wirtschaft (vgl. HRK 1998a: 21f., HRK 1999a: 144f., HRK 1999b: 27f., HRK 1999e: 19f. usw.). Begriffe aus der Wirtschaft fanden beispielsweise auf diesem Weg Eingang in die Qualitätsdiskussionen an Hochschulen. Aber auch indirekt über die Qualitätsdiskussion in Schule und Weiterbildung gibt es Austauschprozesse über die z.T.

---

<sup>29</sup> An anderer Stelle wird die Qualitätsdebatte im Bildungswesen auch als „nur eine Facette eines ökonomischen Megatrends“ bezeichnet. Diesem könne sich das Bildungswesen um so weniger entziehen, je größer seine volkswirtschaftliche Bedeutung wird (z.B. durch Bindung öffentlicher oder privater Mittel oder in seiner Funktion als zentrale gesellschaftliche Qualifizierungseinrichtung) – (vgl. Gnahs u.a. 1995: 8).

getrennt voneinander verlaufenden Diskussionen über die Qualität von Bildung (vgl. Olbertz/Otto 2001).<sup>30</sup>

Um die Wandlung des Qualitätsbegriffs, wie er auch in den Hochschulen diskutiert wird, besser nachvollziehen zu können, erscheint es deshalb sinnvoll, zunächst die Entwicklung des Qualitätsbegriffs in der Wirtschaft nachzuvollziehen. Laut Barth (vgl. HRK 3/1999: 27f.) kann man die Entwicklung des Qualitätsbegriffs in der Wirtschaft kurz in drei Phasen darstellen:

Die erste Phase liegt danach etwa in den 1950er Jahren, wo die Qualitätsprüfung durch stochastische Qualitätsproben erfolgte. Sehr bald kam man jedoch zu der Erkenntnis, daß Fehlervermeidung durch Prozeßoptimierung billiger ist als nachträgliche Kontrolle und Fehlerbeseitigung.

Dies führte zur zweiten Phase, die etwa in 1970er Jahren begann. Stellvertretend für diese Idee steht die DIN ISO 9000ff.. Auch heutzutage wird in Deutschland mit dem Qualitätsgedanken häufig die DIN ISO 9000ff. (vgl. hierzu ausführlicher die Erläuterungen im Abschnitt zu Instrumenten des Qualitätsmanagements) assoziiert, die aber v.a. (Mindest-)Standards definiert und sich nicht als kontinuierlicher Prozeß versteht.

Die dritte Phase, für die meist der Begriff TQM (vgl. auch hierzu ausführlicher die Erläuterungen im Abschnitt zu Instrumenten des Qualitätsmanagements) steht, begann in Deutschland in vielen Unternehmen erst vor einigen Jahren, während in anderen Staaten bereits häufiger von TQM als ständige Aufgabe die Rede ist. Hierbei geht es im Unterschied zur DIN ISO 9000ff. stärker um eine kontinuierliche Verbesserung.

Allerdings gebe es heutzutage im Qualitätsmanagement „ein fast babylonisches Sprachgewirr“ (vgl. HRK 3/99: 27), so daß mit diesen drei Phasen die Entwicklung nur grob zu umschreiben ist. Weitere Begriffe aus der Wirtschaft wie z.B. KVP<sup>31</sup>, EFQM<sup>32</sup>, BSC<sup>33</sup> usw. halten inzwischen in einige Öffentliche Verwaltungen ebenso wie in Hochschulen Einzug. Man solle sich aber von den vielen Begriffen und Modellen nicht verwirren lassen, denn „Es gibt unterschiedliche Wege; welchen man wählt, ist letztlich egal, man muß sich nur konsequent daran halten.“ so Barth zusammenfassend hierzu (vgl. HRK 3/99: 29).

Ein weiterer Versuch der Systematisierung insbesondere für öffentliche Verwaltungen und Hochschulen findet sich bei Schenker-Wicki (1996: 49f.). Sie verwendet für Qualitäts-

---

<sup>30</sup> In der Medizin wurde dies ebenfalls schon zu früheren Zeitpunkten diskutiert (vgl. Bargel 1994: 56ff.).

<sup>31</sup> Kontinuierlicher Verbesserungsprozeß (kann als eine Variante des TQM aufgefasst werden)

<sup>32</sup> European Foundation for Quality Management (vgl. Abschnitt zu Modellen des Qualitätsmanagements im dritten Kapitel dieser Arbeit)

<sup>33</sup> Balanced Scorecard (vgl. Abschnitt zu Modellen des Qualitätsmanagements im dritten Kapitel dieser Arbeit)

prüfungsmaßnahmen in diesem Bereich die Oberbegriffe Revision, Controlling, Evaluation; wobei sie diese mehr oder minder als aufeinander aufbauende Systeme versteht.

An deutschen Hochschulen wird Evaluation häufig getrennt z.B. von Controlling bzw. Kosten- und Leistungsrechnung (KLR) gesehen, obwohl die Leistungsrechnung mit ihrem Anspruch der Bewertung von Hochschulleistungen ähnliche Ziele verfolgt wie unter den Begriff Evaluation gefaßte Maßnahmen der Qualitätsbewertung.<sup>34</sup> Bayer (2000) geht deshalb noch einen Schritt weiter als Schenker-Wicki (1996), indem er kennzahlengestützte Mittelverteilungssysteme, Evaluationen, Rankings bzw. Studienführer (wie z.B. die vom CHE verantworteten) und Zielvereinbarungen zwischen Hochschulen und Ministerien als Maßnahmen mit ähnlichen methodischen Wurzeln und Zielen bezeichnet (vgl. Bayer 2000: 30).

Deutlich gemacht werden sollte mit diesen Ausführungen, daß es hilfreich sein könnte, auch andere als hochschulinterne Perspektiven einzunehmen, da trotz verschiedener Begrifflichkeiten z.T. Parallelen und ähnliche Ziele existieren, die oft aufgrund unterschiedlicher Fachterminologien nicht deutlich werden. Selbst Stawicki, der die Übertragbarkeit von Qualitätsmanagementsystemen relativ kritisch sieht, räumt ein: „Die umfassendste Erfahrung mit der Erzeugung und Verbesserung von Qualität liegt derzeit sicherlich im Bereich der Wirtschaft vor. Deshalb ist es naheliegend, auf den dort gemachten Erfahrungen aufzubauen (...). Auch wenn `normgemäße` Qualitätsmanagementsysteme `not invented here` sind, so ist doch der (kluge und abwägende) Blick über den Zaun (...) hilfreich und sinnvoll.“ (vgl. HRK 1998a:70).

Nachfolgend sollen daher auch Erfahrungen aus anderen Disziplinen einbezogen werden. Schließlich könnten Schwächen der einen Disziplin in Bezug auf ein bestimmtes Ziel Stärken der anderen sein und umgekehrt:

Deutlich wird dies z.B. bei der Diskussion von Modellen zur leistungsorientierten Mittelvergabe oder auch der Kosten- und Leistungsrechnung, für die häufig eine stärkere Einbeziehung von qualitativen Aspekten (oder besser: Qualitätsaspekten) gefordert wird, aber auch für Modelle des Qualitätsmanagements, wozu beispielsweise die Nutzung subjektiver Bewertungen in Form von Befragungen vorgeschlagen wird (vgl. Seeger u.a. 1998: 208, Ederer 1999: 2, Zurwehme 2000: 36ff.). Hier könnte die Ergänzung von betriebswirtschaftlichen Ansätzen durch Erfahrungen und

---

<sup>34</sup> Häufig werden Kosten-/ Leistungsrechnung und Controlling im Zusammenhang genannt. Während Controlling als Steuerungsaufgabe bzw. Instrument der Führung an den Hochschulen teilweise als der Qualitätssicherung übergeordnet gesehen wird (vgl. z.B. HRK 2001a: 25), wird die Kosten- und Leistungsrechnung hier wie auch in anderen Publikationen zum Thema Qualitätsorientierung an Hochschulen nicht als Modell des QM eingeordnet (vgl. z.B. HRK 2003b, Pasternack 2004). Auf die Kosten- und Leistungsrechnung wird in dieser Arbeit nicht weiter eingegangen, da an Hochschulen bisher nur erste Schritte in dieser Richtung gegangen wurden, hierbei v.a. der Aspekt der Kostenrechnung betont wurde, für diese Arbeit aber nur der Aspekt der Leistungsrechnung interessant wäre (vgl. z.B. Rammelt 2003). Eine ausführlichere Erläuterung zu theoretischen Grundlagen findet sich z.B. in Zimmermann u.a. (2003), zu praktischen Erfahrungen in Seeger u.a. (1998).

bewährte Verfahren zur Bewertung von qualitativen Aspekten aus der Evaluation sinnvoll sein. Auf diese Weise könnten Konzepte aus der Wirtschaft evtl. mit Wissensbeständen aus der Soziologie verbunden werden. Soziologen, die im Bereich der Evaluation arbeiten, müßten dazu allerdings die betriebswirtschaftlichen Ansätze kennen und einordnen können. (Buschor 2002: 82).

Die Verbindung von Ansätzen aus verschiedenen Fachgebieten ist für die Evaluationsforschung an sich nichts ungewöhnliches, da sie generell als eine sehr interdisziplinär angelegte Forschungsrichtung angesehen wird (vgl. Schenker-Wicki '96, Rindermann '96).<sup>35</sup> So beschäftigen sich mit Evaluationen beispielsweise Erziehungswissenschaftler, Hochschuldidaktiker, Psychologen, Soziologen, Rechts- und Verwaltungswissenschaftler.

Eine relativ große Interdisziplinarität mit ihren verschiedenen Fachsprachen und unterschiedlichen Perspektiven macht aber neben einer klaren Definition der Begrifflichkeiten erst recht Klärungen zur Qualitätsbestimmung anhand der verfolgten Ziele notwendig. Deshalb soll nachfolgend versucht werden, den Qualitätsbegriff anhand von Zielklärungen auf verschiedenen Ebenen einzugrenzen.

## **2.3 Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen**

In vielen Veröffentlichungen zum Thema wird das Problem unklarer Ziele von Leistungsbewertungen bzw. Evaluationsmaßnahmen im Hochschulbereich angesprochen. Fragen, die demnach zuerst zu stellen sind, lauten: Qualität von was? (Qualitätsgegenstand) und schließlich: Qualität für wen? (Qualitätszielgruppe).

Bei der Frage: Qualität von was? wäre, da in dieser Arbeit der Bereich der Lehre fokussiert wird, z.B. zu unterscheiden einerseits nach Lehrveranstaltungen (wie es implizit z.B. im SächsHG und anderen Hochschulgesetzen formuliert ist und als Bestandteil vieler Evaluationen Anwendung findet) und andererseits nach Qualität der Hochschullehre im Sinne einer Qualität des Studiums bzw. im Sinne von Ausbildungsqualität. Hierbei wird das Studium als Ganzes betrachtet. Eine angemessene Bewertungsebene wäre hier vorrangig die des Studienganges.

Bezogen auf die Evaluation bzw. Bewertung von Qualität findet sich dies in ähnlicher Form wieder. So unterscheidet Wolter (1995) zwischen Mikro-, Meso- und Makroebene. Zwar gibt es

---

<sup>35</sup> Dies gilt im Übrigen auch für neuere Ansätze in den Wirtschaftswissenschaften: So formulieren Lichka u.a. (2002: 918), die erfolgreiche Umsetzung des Ansatzes der Balanced Scorecard (vgl. auch Abschnitt zu Instrumenten des Qualitätsmanagements in dieser Arbeit) stellt „hohe interdisziplinäre Anforderungen.“ „Und zwar nicht nur an die Betriebswirtschaftslehre und die Wirtschaftsinformatik, sondern beispielsweise auch an Fachgebiete wie Soziologie und Psychologie, wenn es um die Inkorporierung innovativer (und strategieorientierter) Human Resource-Ansätze geht.“

weiter ausdifferenzierte Modelle, diese enthalten jedoch die von Wolter beschriebenen Ebenen (vgl. HIS 1992: 121). Pasternack (2004: 19) nennt ebenfalls die von Wolter unterschiedenen Ebenen, die bei der Zielklärung von Qualitätsbewertung und insbesondere bei vergleichender Betrachtung notwendig seien.

Als Mikroebene wird von Wolter die einzelne Lehrveranstaltung bezeichnet. Zahlreiche Evaluationsinstrumente beziehen sich auf diese Ebene einer Mikropolitik der Verbesserung einzelner Lehrveranstaltungen. „Lehr- und Studienqualität umfassen jedoch wesentlich mehr als die Summe einzelner Lehrveranstaltungen. Eine sinnvolle Evaluation der Hochschullehre muß sich im Interesse einer ständigen Qualitätsverbesserung der Lehre primär auf die Ebene der Studiengänge, ihrer Stärken und Schwächen, auf die Strukturen der Studienorganisation und des akademischen Lehrbetriebes konzentrieren, statt auf einzelne Personen oder Lehrveranstaltungen“, betont Wolter (1995: 76).

Mit Mesoebene wird die Ebene einzelner Studiengänge bezeichnet, einschließlich deren Studienprogramm, Lehrveranstaltungsangebot und institutioneller Rahmenbedingungen wie die Studienordnungen, aber auch die Ausstattungssituation der Fachbereiche, die diesen Studiengang ausbilden.

Mit Makroebene schließlich wird die Ebene der Hochschule selbst, aber darüber hinausgehend auch die von den Ministerien und Parlamenten vorgegebenen institutionellen Rahmenbedingungen wie Verordnungen, Gesetze sowie Mittelzuweisungen bezeichnet.

Ähnlich dieser drei Ebenen strukturieren Becher/ Kogan (1992) auch das britische Hochschulsystem, allerdings trennen sie die oben als Makroebene bezeichnete Ebene noch weiter auf (vgl. Orr 2001: 207): Demnach gebe es als Grundstruktur das Individuum (Lehrstuhlinhaber bzw. Lehrveranstaltungsleiter), die Basiseinheit (die Studiengänge verwaltende Einheit, die meist der Fakultät noch untergeordnet ist), die Institution (Hochschule) und die zentrale Instanz (z.B. Bildungsministerium, aber auch Finanzierungsgremien, Qualitätssicherungsagentur und Hochschulrektorenkonferenz).

Die Überlegungen, Evaluation bzw. Qualitätsbewertung auf verschiedenen Ebenen auszu-differenzieren, finden sich schließlich im Zusammenhang mit der Kommunikation von Evaluationsergebnissen auch in der „Informationspyramide“ nach Barz u.a. (1997). Grundlage seiner Überlegungen ist ebenfalls, daß vielerorts unklar blieb, für wen (welche Zielgruppe) die Lehrberichte anzufertigen sind und was mit den Ergebnissen geschehen soll. Hier sei deshalb eine klarere Zielbestimmung und eine Klärung notwendig, welche Ergebnisse genau für wen bestimmt sein sollen, so Barz weiter. Sinnvoll erschien es ihm daher, hier nach dem Modell einer „Informationspyramide“ vorzugehen, nach der auf Studiengangs- und Fachbereichsebene

detailliert auf Stärken und Schwächen eingegangen und an ihrer Verbesserung gearbeitet wird. Mit jeder höheren Hierarchieebene (Hochschulleitung, Ministerien) sollte sich jedoch mehr auf die Kerninformationen und verallgemeinerbare Resultate beschränkt werden (Vgl. Barz u.a. 1997: 62). Er schlug daher eine Begriffstrennung in „Evaluationsberichte“ (umfassende, selbstkritische Bestandsaufnahme) und „Lehrberichte“ (für die Außendarstellung bestimmt) vor. Inzwischen hat der Begriff Lehrbericht auch Eingang in zahlreiche Hochschulgesetze gefunden, wobei in der Tat häufig eine Bereitstellung dieser Berichte für die Öffentlichkeit gefordert wird (vgl. z.B. §6 SächsLehrberVO und SächsHG vom 11. Juni 1999). Für die Evaluationsberichte ist dies weniger klar. Allerdings veröffentlichen einige Evaluationsagenturen auch ihre Evaluationsberichte (vgl. HRK 2004: 133 ff.).

Wenn nun die Frage beantwortet werden soll, „Qualität von was?“, so ist festzustellen, daß dies für verschiedene Ebenen des Hochschulsystems differenziert werden sollte und daß es auch von den jeweiligen Beteiligten abhängt. Pasternack (2004:14) formuliert in diesem Zusammenhang (allerdings ohne daß dies dort im Detail umgesetzt wurde): „Eine konkrete Qualitätsbestimmung, welche die Definitionsprobleme erfolgreich bearbeitet, lässt sich dadurch gewinnen, daß sie situations- sowie interessen- und motivationsabhängig vorgenommen wird. Das heißt, sie soll von der konkreten Situation, sowie ihren konkreten Akteuren und Erwartungen ausgehen.“ Ähnlich argumentiert auch Orr (2001: 106ff.): Qualität und ihre Bewertung seien multi-dimensional und hängen deshalb immer vom betrachteten Gegenstand ab.

Nachfolgend soll also zunächst versucht werden, den Begriff der Qualität der Hochschullehre auf verschiedenen Ebenen zu klären. Anschließend wird er für verschiedene Interessen- bzw. Zielgruppen diskutiert.

### **2.3.1 Definition von Lehrqualität auf Hochschulebene**

Die Qualität der Hochschullehre ist auf dieser relativ abstrakten Ebene wohl am schwersten zu erfassen, da, wie bereits am Begriff Leistung dargestellt, diese je nach Fach sehr verschieden aufgefasst werden kann. Hinzu kommt das Problem der Vergleichbarkeit selbst bei prinzipiell gleichen Qualitätskriterien zwischen verschiedenen Fächern bzw. Studiengängen. Dies zeigten in den letzten Jahren z.B. im Bereich der Forschung die Diskussionen um die Anzahl der Publikationen als Kriterium. Zuvor wurde im Zusammenhang mit Ranglisten bzw. sogenannten Rankings auch problematisiert, Qualitätsaussagen über die Hochschulen als Ganzes zu treffen. Inzwischen ist die Ansicht weit verbreitet, daß Qualitätsaussagen nur auf der Basis der Fächer bzw. im Vergleich derselben Fächer verschiedener Hochschulen untereinander sinnvoll sind.

Versuche, auf dieser Basis ermittelte Qualitätsaussagen doch wieder in Form von Punktwerten bzw. „Scores“ zusammenzufassen, werden eher selten unternommen und sind umstritten.<sup>36</sup>

Eine andere Herangehensweise wäre, neben der Qualität der Fächer die Qualität der Hochschule als fachübergreifende Institution zu betrachten, z.B. bezüglich ihrer Voraussetzungen für die Qualität der Fächer. Dies ist z.B. bei der institutionellen Akkreditierung der Fall, wie sie in den USA seit längerem praktiziert wird (zu einem Überblick vgl. Holtkamp/ Schnitzer 1992: 56) und seit einiger Zeit auch für Deutschland diskutiert wird (vgl. HRK 2003b: 17).<sup>37</sup> Da aufgrund der geringen Akzeptanz dieses Ansatzes Qualitätsaussagen auf der Ebene einer Hochschule als Zusammenfassung verschiedener Fächer nicht sinnvoll erscheinen, soll nachfolgend eher dem Ansatz gefolgt werden, die Qualität der Hochschule als fachübergreifende Institution zu betrachten. Eine umfassende Definition kann hier aufgrund der gerade auf dieser Ebene in Deutschland noch nicht allzu fortgeschrittenen Entwicklung, die gleichzeitig derzeit eine große Dynamik entfaltet, aber (noch) nicht erfolgen. Daher sollen nur einige Elemente genannt werden, die nach Auffassung des Verfassers auf der Hochschulebene Qualität umschreiben.

Letztlich sollte sich eine hohe Qualität der Lehre auf Hochschulebene dadurch auszeichnen, daß Qualität nicht nur in einzelnen Lehrveranstaltungen oder einzelnen Studiengängen bzw. -programmen vorhanden ist. Vielmehr sollten sie nicht nur einzeln für sich stehen, sondern verbundene Teile eines Ganzen sein, welche Synergiewirkungen aufweisen und sich gegenseitig zu Weiterentwicklungen anregen. Pasternack (2004: 7) beschreibt dies als die einen Gegenstand ganzheitlich durchformende Güte auch als „Qualität zweiter Ordnung“, die sich im Gegensatz zur „Qualität erster Ordnung“ (bei der es um diverse Einzeleigenschaften gehe) nicht beschreibend standardisieren lasse.<sup>38</sup> Im Gegensatz zu dieser Auffassung wurde im britischen Hochschulwesen seit einiger Zeit ein Wandel vom „subject review“ hin zum verstärkten Einsatz von „institutional review“ auf der Ebene der Hochschule diskutiert (vgl. Frederiks 2001: 55ff.) und teilweise umgesetzt (vgl. Grieshop 2004: 551). Im „institutional review“ geht es v.a. um die Beurteilung des Qualitätsmanagements einer Hochschule. Hierbei wird davon ausgegangen, daß die Hochschule bei einem funktionierenden Qualitätsmanagement selbst für die Qualität ihrer Studiengänge sorgt und daher „subject review“ ersetzen kann.<sup>39</sup>

---

<sup>36</sup> Beispiele hierfür sind das Spiegel-Ranking 1993 oder als jüngeres Beispiel auch das Focus-Ranking 2004 (siehe Stellungnahme des CHE (2004) zum Focus -Ranking).

<sup>37</sup> So wurden in „Forschung und Lehre“, der Zeitschrift des DHV, durchaus kontroverse Diskussionsbeiträge zum Thema Akkreditierung veröffentlicht (vgl. z.B. Nr. 10/ 2000: 508ff., Nr. 10/ 2004: 546ff.).

<sup>38</sup> In diesem Zusammenhang ist zu beachten, daß Pasternack hier zwischen Quantitäten und Qualitäten unterscheidet, wobei er erstere für meßbar hält, letztere nicht. Dies ergibt ein deutlich anderes Begriffsverständnis, als es die überwiegende Mehrheit der Evaluationsforscher vertritt (vgl. Abschnitt 2.2.2).

<sup>39</sup> Mit dem „institutional review“ findet zwar keine flächendeckende Qualitätsprüfung der Studiengänge mehr statt, die externen Gutachter wählen jedoch aufgrund der Selbstevaluationsberichte Fach- und Themenbereiche zur

An den deutschen Hochschulen existieren bisher häufig unverbundene Einzelmaßnahmen der Qualitätssicherung und -verbesserung, wie Studentenbefragungen und die Lehrberichterstattung, Absolventenbefragungen, Lehrenden- bzw. Mitarbeiterbefragungen zu bestimmten Themen, Benchmarkings und in den letzten Jahren zunehmend die Aktivitäten zur Akkreditierung von Studiengängen. Diese sollten, um eine „Qualität zweiter Ordnung“ erreichen zu können, nicht nur auf der Ebene der Studienprogramme und auf der Ebene der Fakultäten zusammengeführt werden, sondern auch auf der Ebene der Hochschule. Pasternack systematisiert dies für sein Qualitätsstrukturmodell wie folgt: Die Einzelmaßnahmen zur Verbesserung der Qualität erster Ordnung bezeichnet er als „Single-issue-Ansatz“, die Maßnahmen zur Verbesserung der Qualität zweiter Ordnung als „Systemveränderungsansatz“. So wie bei der Zielbestimmung in Qualitäten erster und zweiter Ordnung unterschieden wurde, sollte auch die Bewertung getrennt erfolgen, da sie je spezifische Bewertungsverfahren erfordern (vgl. Pasternack 2004: 12). Als Beispiel wird hierzu erläutert: Das Qualitätsmerkmal „hochschuleigene Kinderbetreuungsmöglichkeiten“ lasse sich als Qualität erster Ordnung relativ einfach prüfen, während es als zur Bewertung, ob eine Hochschule das Prädikat „frauenfreundlich“ verdient, gänzlich unzulänglich sei. Es gibt sicherlich viele weitere Beispiele. Exemplarisch läßt sich die Problematik auch für die Förderung der Fremdsprachenkompetenz der Studierenden als Qualität erster Ordnung und die Bewertung der Internationalität einer Hochschule als Qualität zweiter Ordnung formulieren.

Während ein Single-issue-Ansatz eine punktgenaue Intervention erlaubt und leichter realisierbar ist, gestaltet sich ein Systemveränderungsansatz für die gesamte Hochschule u.a. aufgrund der verschiedenen Fächerkulturen und z.T. aufgrund der Konkurrenzsituation der Fakultäten bzw. Fachbereiche erfahrungsgemäß schwierig, könnte aber – wenn er erfolgreich ist - besonders nachhaltig wirken und das Qualitätsprofil einer Hochschule ausmachen. Die HRK (2003b: 17) spricht in diesem Zusammenhang in ihrem „Sachstandsbericht“ von der Erkenntnis aus den bisherigen Erfahrungen, daß Verfahren zur Qualitätssicherung „nur begrenzt erfolgreich“ sind, wenn sie die einzelnen Bereiche der Hochschule isoliert betrachten. Zum anderen erfordert aber auch die Entstehung des Wettbewerbs zwischen den Hochschulen und das in diesem Zusammenhang an Bedeutung gewinnende Gebot der Transparenz einen solchen „die gesamte Institution umfassenden Ansatz“.

---

Begutachtung aus, bei denen sie entweder eine beispielhafte oder eine bedenkliche Praxis erwarten (vgl. Leszcensky u.a. 2004: 64). Integraler Bestandteil ist außerdem eine formale Stellungnahme der Studierenden zum hochschul-internen System der Qualitätsbewertung.

### **2.3.2 Definition von Lehrqualität auf Studiengangsebene**

Nicht so sehr auf die Ebene einer einzelnen Lehrveranstaltung bezogen, sondern eher auf der Ebene der Ausbildungsprogramme insgesamt hat der Wissenschaftsrat die Frage beantwortet, was unter Qualität der Lehre zu verstehen sei.

Der Wissenschaftsrat (1996: 14) hat folgende „fünf Ausprägungen“ von Qualität unterschieden, wobei angefügt wird, daß „nicht alle Qualitätsdimensionen [...] für die verschiedenen Ausrichtungen von Evaluationsmaßnahmen in gleichem Maße relevant sein“ können. Vielmehr seien „die Ziele und Funktionen eines Evaluationsverfahrens mit den jeweiligen Qualitätsdimensionen abzustimmen“.

So führt der Wissenschaftsrat zur ersten Ausprägung von Qualität aus, daß sich „Qualität [...] insbesondere auf das Ausbildungsprofil der Absolventen eines Studienganges, ihre Kenntnisse, Fähigkeiten und Qualifikationen, die als Ergebnisse der Ausbildung bewertet oder beurteilt werden, beziehen.“ Die Beurteilung von Qualität erfolge in diesem Fall „auf der Grundlage von (implizit oder explizit) formulierten Vorstellungen, Standards oder Normen, die Qualitätsmessung erlauben und ein Urteil darüber ermöglichen, ob ein Absolventenprofil als „gut“ oder „schlecht“ einzuschätzen ist“ (Wissenschaftsrat 1996: 14).

Als zweite Ausprägung wird „Qualität als Resultat eines hohen Maßes an Konsistenz und Kohärenz eines (Ausbildungs-)Prozesses“ genannt (Wissenschaftsrat 1996: 14). Dieser Aspekt hebe, so der Wissenschaftsrat weiter, „auf die Organisation des Lehr- und Studienbetriebs, dessen Schlüssigkeit und Stimmigkeit im Hinblick auf gesetzte und zu erreichende Ausbildungsziele ab“.

In der dritten Ausprägung von Qualität wird diese als „Kongruenz zwischen der von einem Fachbereich getragenen Ausbildungspraxis einerseits und den Erwartungen und Ansprüchen andererseits, die an eine wissenschaftlich fundierte Hochschulausbildung herangetragen werden“ beschrieben. Im Blickfeld stehen hier „die Studienziele und angestrebten Ausbildungsprofile“. Ein zentrales Augenmerk gilt dabei der Verwendbarkeit der Studienabschlüsse auf dem Arbeitsmarkt.

Mit „Qualität als Ausdruck eines günstigen Verhältnisses zwischen eingesetzten Mitteln und dem damit erzielten Ergebnis“ wird die vierte Ausprägung von Qualität beschrieben (Wissenschaftsrat 1996: 15). Da sich diese Ausprägung von Qualität an Effizienzgesichtspunkten orientiert, „sind dabei Aspekte der Finanzierung von Studium und Lehre“ ebenso von Bedeutung sowie die materielle und personelle „Ausstattung von Fachbereichen und Hochschulen“ (Wissenschaftsrat 1996: 15). In dieser Ausprägung führt ein möglichst effizienter

Mitteinsatz im Studienbetrieb bzw. zur Realisierung von Ausbildungszielen zu Qualität (Wissenschaftsrat,1996: 15).

Als fünfte Ausprägung von Qualität wird der „Prozeß der Qualifizierung zwischen Beginn und Abschluß eines Studiums“ (Wissenschaftsrat 1996: 15) beschrieben. Im Vordergrund dieser Ausprägung von Qualität stehen somit „Aspekte von Bildung und Ausbildung, Persönlichkeitsentwicklung und Wissenserwerb“.

Die fünf Ausprägungen der Lehrqualität sollen wie folgt kurz zusammengefaßt werden:

- 1) Absolventenprofil, das Ergebnis der Ausbildung sein soll (Qualität der Ausbildungsziele)
- 2) Schlüssigkeit und Stimmigkeit des Ausbildungsprozesses im Hinblick auf Ausbildungsziele (Zieladäquanz der Ausbildung)
- 3) Verwendbarkeit der Studienabschlüsse auf dem Arbeitsmarkt (Arbeitsmarktadäquanz)
- 4) Verhältnis eingesetzter Mittel zu damit erzieltm Ergebnis (Effizienz)
- 5) Prozeß der Qualifizierung zwischen Beginn und Abschluß eines Studiums (Prozessqualität)

Quelle: Wissenschaftsrat (1996), Zusammenfassung: Krempkow

Auf der Studiengangsebene gäbe es noch eine Reihe weiterer Definitionsmöglichkeiten, die sich u. a. auch aus Evaluationen im Bereich der (Hochschul-)Weiterbildung und dem Bereich der beruflichen Bildung ableiten und entsprechend angepasst verwenden ließen (vgl. z.B. Zurwehne 2000). Da die Definition des Wissenschaftsrates jedoch relativ anerkannt ist und zudem keine Ausschließlichkeit beansprucht, sondern vielmehr explizit Erweiterungen und Spezifikationen zuläßt, soll nachfolgend diese Definition für die Studiengangsebene als Ausgangsbasis genutzt werden.

### **2.3.3 Definition von Lehrqualität auf Lehrveranstaltungsebene**

Die zweite, aus Perspektive der Lehrveranstaltungsebene mögliche Antwort auf die Frage: Qualität von was? muß zunächst weniger klar ausfallen.

Zuerst ist hierzu festzustellen, daß es nach wie vor keine allgemein anerkannte Definition, gibt, was Qualität der Lehre auszeichnet, wenn man sie als Qualität einer Lehrveranstaltung versteht, obwohl es bereits früh Definitionsversuche gab (vgl. Webler 1991: 243). Weblers Ausführungen „Es gibt bisher keinen Konsens über Kriterien guter Lehre an deutschen Hochschulen“ sind nach wie vor aktuell. Dies schlägt sich bis heute auch in den relativ unterschiedlichen Ländergesetzgebungen und Lehrberichtsverordnungen nieder, die doch alle das gleiche Ziel

verfolgen: Einen Beitrag zur Stärkung der Lehre an den deutschen Universitäten zu leisten (Vgl. Hage 1996: 1).

Dennoch sieht sich Webler (1991) stellvertretend für die „professionelle Hochschuldidaktik prinzipiell in der Lage, trotz einiger Schwierigkeiten anzugeben, worin gute akademische Lehre besteht“: Allgemein formuliert heißt seine Definition guter Lehre: „Wenn die Studierenden auf ihrem Weg zu den Zielen des Studiums allgemein und den Zielen des Studienabschnitts speziell in hervorragender Weise gefördert werden“.

Diese Definition schließt, wenn man sie genauer betrachtet, sowohl aktive, als auch passive Aspekte auf Seiten der Studierenden ein: Gefördert zu werden bedeutet nicht, lediglich Lehrveranstaltungen zu „konsumieren“, sondern auch, selbst etwas zu tun, bei dem man unterstützt wird. Die Definition bezieht sich dabei bewußt nicht auf konkrete Lehrinhalte. Zwar bezweifeln viele, daß sich die Frage, wann ein Lehrangebot gut ist, überhaupt fächerunspezifisch lösen läßt. (Vgl. Webler 1991: 243). Jedoch stützt Webler sich bei seinen Ausführungen gerade deshalb ausschließlich auf Informationsquellen, die „unter minimalem Bezug auf die transportierte Information“ (d.h. die Lehrinhalte) Anhaltspunkte für gelingende Lehr- und Lernprozesse geben. Es sind dies Anleihen aus der Lernpsychologie, Lernbiologie, Motivationspsychologie, Kleingruppenforschung, Kreativitätstheorie, Kommunikationstheorie und ähnlichen Gebieten, die in die hochschuldidaktische Lern- und Lehrforschung eingeflossen sind. Diese Informationsquellen seien dabei in jedem Falle wesentlich systematischer als die „Alltagstheorien“, die manchen ministeriellen Bestimmungen als Grundlage dienen, begründete Webler seinen Standpunkt.

Auch der Begriff Didaktik, im Zusammenhang mit der Qualität der Lehre oft verwendet, bedarf in diesem Zusammenhang einer Begriffsklärung. Webler versteht dabei unter Didaktik allgemein „die Verknüpfung der zielabhängigen Auswahl von Inhalten mit Methodenentscheidungen“. Webler legt bei dieser Definition großen Wert auf die Ziele, die mit der Lehre verbunden sind, da ihm diese in der hochschulpolitischen Diskussion nur allzu oft aus dem Blick geraten würden. Deshalb heißt es bei Webler unter diesem Blickwinkel zum Thema Lehre: „Gute Lehre ist dann gut, wenn sie, gemessen an den Studienzielen, *möglichst optimale Lehr- und Lernprozesse organisiert*“ (Webler 1991: 243 ff.). Hier wird klar, daß Webler, abweichend von seiner allgemeinen Definition „guter Lehre“, seinen Kriterienkatalog vor allem nach hochschuldidaktischen Gesichtspunkten zusammenstellte. Das wird schon an der Gliederung deutlich: Im ersten Abschnitt dieses Kataloges wurden Merkmale guter Veranstaltungsplanung und Veranstaltungskonzepte aufgeführt, im zweiten Abschnitt Kriterien für geeignetes Lehrverhalten der Lehrenden (Webler 1991: 246-248). Zwar wird in diesem Kriterienkatalog explizit darauf

hingewiesen, daß dieser lediglich eine Diskussionsgrundlage darstellen soll und keineswegs Anspruch auf Vollständigkeit erhebt. Dennoch wird sich bis heute häufig auf den Kriterienkatalog von Webler bezogen, wenn es *allgemein* darum geht, was unter guter Lehre zu verstehen sei.

Eine weitere, stark am Wissenszuwachs orientierte Definition von „guter Lehre“ lieferte Fritz Süllwold (1997), emeritierter Professor für Psychologie an der Universität Frankfurt/Main. „Gut“ ist Lehre demnach, wenn „Wissen und Problemverständnis der Studierenden auf dem betreffenden Gebiet signifikant erhöht“ werden. Andere Lehrende wiederum legen den Schwerpunkt viel weniger auf die Vermittlung von *Wissen*, sondern vielmehr darauf, Studierende für die zu vermittelnde Thematik zu interessieren und zu selbständiger Beschäftigung damit zu motivieren. Es ließen sich noch eine Reihe weiterer Standpunkte nennen. Zur Verdeutlichung der Breite, in der sich die Diskussion um und die Vorstellungen von „guter Lehre“ bewegt, sollte dies jedoch genügen.

Als eine Art Zwischenbilanz der Diskussion läßt sich aufgrund der vorgelegten empirischen Befunde jedoch festhalten: In eine Definition von „guter Lehre“ sollten sowohl didaktische Aspekte, als auch Aspekte der Studienbedingungen eingehen. Denn was nützt es, Studierende beispielsweise zur eigenständigen Beschäftigung mit einem Thema zu motivieren, wenn die nach besten didaktischen Kriterien empfohlene Literatur aufgrund der Bibliothekssituation nicht nutzbar ist oder aufgrund der Überfrachtung des Studienganges das Zeitbudget der Studierenden nicht dazu ausreicht?

In eine Definition von „guter Lehre“ sollte deshalb nicht nur die (didaktische) Gestaltung von Lehrveranstaltungen eingehen. Vielmehr sollte das Konzept „guter Lehre“ um die Lehr-, Lern- und Studienbedingungen erweitert werden.

Qualität der Lehre besteht damit nicht nur darin, daß die Studierenden auf ihrem Weg zu den Zielen des Studiums allgemein und den Zielen des Studienabschnitts speziell (durch nach hochschuldidaktischen Prinzipien gestaltete Lehrveranstaltungen) in hervorragender Weise gefördert werden. Qualität der Lehre besteht auch darin, neben den Lehrveranstaltungen die (selbständige) Aneignung von Wissen und Fähigkeiten entsprechend den Lehr- und Lernzielen zu ermöglichen und hierfür möglichst günstige (Studien-)Bedingungen zu schaffen.

### 2.3.4 Zielgruppen für die Qualität der Hochschullehre

Nachdem versucht wurde, die Frage Qualität von was? zu beantworten und dies bereits auf mehreren Ebenen ausdifferenziert wurde, soll dies noch einmal aus einer anderen Perspektive betrachtet werden und die Frage: Qualität für wen? unter dem Aspekt der Zielgruppenorientierung diskutiert werden.<sup>40</sup> Bayer (2004: 130) weist in Analogie zum Qualitätsmanagement von Dienstleistungen darauf hin, daß die Qualität einer Hochschulleistung nur in Abhängigkeit von definierten Zielgruppen gemessen werden kann. Durch die Diskussion der Zielgruppen sollen auch noch einmal Probleme etwas herausgearbeitet werden, die in der Praxis der Qualitätssicherung auch im Bildungsbereich häufig eine Rolle spielen (vgl. Zurwehme 2000: 49).

Dies ist insbesondere vor dem Hintergrund von Bedeutung, daß in den letzten Jahren in verschiedenen Studien die Zielgruppenfrage noch stärker als zuvor diskutiert wird und dabei auch innerhalb bisher als relativ homogen angenommener Zielgruppen relevante Unterschiede ausgemacht werden. So sieht Bayer (2000) auf Studiengangsebene größere Unterschiede innerhalb der Zielgruppe der Unternehmen und der potentiellen Absolventen je nach spezifischen Interessenlagen und Berufsfeldern (vgl. auch Bayer 2001).

Kromrey (2000) sieht relevante Unterschiede auch innerhalb der Zielgruppe der Studierenden auf Ebene einer Lehrveranstaltung je nach Studienmotivation und Interessenlagen. Kromrey bezieht sich hierbei auf ältere Publikationen, in denen er den Standpunkt vertrat: „Gute Lehre ist ohne Bezugnahme auf eine spezielle Zielgruppe weder definierbar noch realisierbar“ (Kromrey 1993: 46). Dies betreffe sowohl die Abhängigkeit der Vorstellungen guter Lehre von den zu vermittelnden Lehrinhalten, als auch von der jeweiligen studentischen „Zielgruppe“. Kromrey illustrierte dies folgendermaßen: „Sollen z.B. die für das erfolgreiche Studium eines Faches von den Lehrenden für notwendig gehaltenen, bei den Studierenden jedoch überwiegend unbeliebten Basiskenntnisse und Basisfertigkeiten vermittelt werden und muß zudem davon ausgegangen werden, daß es nennenswerte Lücken in den studentischen Vorkenntnissen gibt, dann ist Lehre ganz anders zu konzipieren, als wenn Studierende im Hauptstudium in einem sie interessierenden Teilgebiet zum kreativen Umgang mit wissenschaftlichen Erkenntnissen oder gar zum Finden neuer Erkenntnisse angeleitet werden sollen. Was für den ersten Fall gute Lehre ist, wäre für den zweiten Fall extrem schlechte Lehre - und umgekehrt.“ Kromrey (1993, 1995, 1996) begründete dies ausführlich anhand mehrerer empirischer Untersuchungen). Auf diese

---

<sup>40</sup> Der Begriff „Zielgruppe“ erscheint hier angemessener als der Begriff „Kunde“ (unter dem z.T. ähnliche Probleme im Zusammenhang mit der Übertragbarkeit von Qualitätsmanagementmodellen auf den Hochschulbereich diskutiert werden), da der Begriff „Kunde“ impliziert, daß Qualität möglichst vollständig auf die Wünsche des „Kunden“

wird jedoch erst an späterer Stelle im Abschnitt zur Vergleichbarkeit der Ergebnisse von Studentenerhebungen eingegangen.

Zunächst soll dies zur Begründung genügen und daher sind nachfolgend die (möglichen) Zielgruppen auf den verschiedenen Ebenen des Hochschulsystems zu benennen: Auf Lehrveranstaltungsebene besteht die Zielgruppe in erster Linie aus den Studierenden, die die Lehrveranstaltung nachfragen. Diese Zielgruppe ist ggf., wenn sich deren Motivation und damit deren Lernziele deutlich voneinander unterscheiden, noch einmal in verschiedene (Unter-)Zielgruppen aufzuspalten. Eine andere Zielgruppe auf Lehrveranstaltungsebene sind aber auch die Lehrenden, für die durch eine Verbesserung der Lehrqualität eine bessere Motivation der Studierenden und (für sich selbst) eine größere Arbeitszufriedenheit zu erlangen möglich wäre. Für wiss. Mitarbeiter könnten sich langfristig, wenn die hochschulpolitischen Ziele einer Aufwertung der Lehre gegenüber der Forschung an den Hochschulen umgesetzt werden, evtl. bessere Karrierechancen auch durch Ausweis guter Lehrqualifikation ergeben. Insbesondere bei letzterem konkurrieren aber derzeit verschiedene Ziele miteinander, da sich wissenschaftlicher Nachwuchs derzeit bessere Jobchancen v.a. über den Ausweis der Forschungsqualifikation sichern kann und die Lehrqualifikation zwar auch gefordert ist, jedoch eine deutlich geringere Bedeutung hat. Dies muß bei der Klärung der Ziele berücksichtigt werden.

Auf Studiengangsebene bzw. auch für die Ebene der Fachbereiche können 1.) einstellende Unternehmen und wiss. Einrichtungen als „Abnehmer“ der Absolventen und 2.) wiederum Studierende gesehen werden, die das Ausbildungsprogramm nachfragen und damit potentielle Absolventen sind. Außerdem sind in der Praxis, solange sie als Geldgeber fungieren, grundsätzlich auch Ministerien und Parlament sehr wichtige Zielgruppen (vgl. HRK 1998b: 75). Hinzu kämen auf Ebene der Hochschule als Zielgruppe auch die Stadt oder die Region. Schließlich haben diese bei Hochschulen mit ausgewiesener guter Ausbildungsqualität bessere Chancen auf Ansiedlung von Unternehmen. Die Hochschulen hätten dann ggf. bessere Kooperationsmöglichkeiten mit der (regionalen) Wirtschaft. Schlußendlich ist es jedoch die Gesellschaft als Ganzes, die gemäß dem gesetzlichen Auftrag der Hochschulen eine hohe Ausbildungsqualität bei möglichst effektivem Mitteleinsatz erwartet bzw. „nachfragt“ (vgl. HRK 1999: 41ff).

Will man Qualität für diese Zielgruppen erreichen bzw. diesen Zielgruppen vermitteln, daß man für die jeweils anderen Zielgruppen eine hohe Qualität erreicht hat, so wären zunächst die Ansprüche der jeweiligen Zielgruppen zu kennen, die Qualität – soweit möglich und sinnvoll – an diesen Ansprüchen zu orientieren, um schließlich glaubwürdig vermitteln zu können, daß die

---

auszurichten sei (vgl. Abschnitt zu Qualitätsmanagementmodellen in dieser Arbeit), während der Begriff „Zielgruppe“ durchaus deren Mitwirkung ermöglicht.

erreichte Qualität diesen Ansprüchen genügt. Genau in einer solchen Zielgruppenorientierung stecken aber auch Probleme. Dies zeigt sich beispielsweise, wenn sich diese Ansprüche der Zielgruppen (teilweise) widersprechen. Neben dem eingangs erwähnten Zielkonflikt für wissenschaftliche Mitarbeiter, gleichzeitig gute Lehre zu bieten und nur über Forschungsleistungen ihre Karrierechancen sichern zu können, gibt es weitere Zielkonflikte; ein weiteres Beispiel könnte das Interesse mehrerer Städte oder Regionen am Erhalt ihrer jeweiligen Studienangebote bei möglichst hoher Qualität und das Interesse des Ministeriums bzw. Parlamentes an einer möglichst effizienten Mittelverwendung sein. In solchen Fällen müssen in einem mehr oder weniger politischen Aushandlungsprozeß die Ziele geklärt oder zumindest gewichtet werden, da es wenig zweckmäßig wäre zu versuchen, allen gleichermaßen gerecht zu werden. Hierfür gibt es in den letzten Jahren Ansätze, wie das Qualitätsmanagementmodell der EFQM oder die sogenannte Balanced Scorecard (BSC). Auf solche Modelle ist jedoch in einem gesonderten Abschnitt detaillierter einzugehen.

### **2.3.5 Fazit zur Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen**

Anhand der Diskussion einschlägiger Publikationen wurde herausgearbeitet, daß für eine Eingrenzung des Qualitätsbegriffs vor allem die Fragen „Qualität von was?“ und „Qualität für wen?“ zu beantworten sind. Die erste Frage bezieht sich auf den Qualitätsgegenstand, die zweite auf die Qualitätszielgruppe.

Der Gegenstand der Qualität von Hochschullehre ist hierbei für die verschiedenen Ebenen des Hochschulsystems differenziert zu betrachten. In dieser Arbeit wurde versucht, Lehrqualität für die Ebene der gesamten Hochschule (Hochschulebene), die Studiengangsebene und die Lehrveranstaltungsebene zu definieren. Während für die Hochschulebene nach wie vor Schwierigkeiten zu konstatieren sind, eine Definition zu finden, die auch operationalisierbar ist, kann für die Studiengangsebene die Definition des Wissenschaftsrates (1996) als Ausgangsbasis genutzt werden. Für die Lehrveranstaltungsebene existiert in der Literatur eine Vielzahl von Definitionen. In dieser Arbeit wird eine Definition favorisiert, in die nicht nur die (didaktische) Gestaltung von Lehrveranstaltungen eingeht, sondern die um die Lehr-, Lern- und Studienbedingungen erweitert wurde.

Zielgruppen für die Qualität der Hochschullehre können ebenfalls sehr vielfältig sein. An erster Stelle sind selbstverständlich die Studierenden zu nennen, die allerdings keine homogenen Interessen haben müssen. Vielmehr können sie je nach Motivationslage und/ oder Vorkenntnissen in unterschiedliche Zielgruppen aufgeteilt werden. Eine ebenso wichtige Zielgruppe sind aber auch die Lehrenden selbst, für die durch eine bessere Lehrqualität z.B. eine höhere

Motivation und Arbeitszufriedenheit zu erlangen möglich wäre. Weitere Zielgruppen sind (auf Studiengangsebene betrachtet) Unternehmen, die potentiell die Absolventen eines Studienganges einstellen oder auch Abiturienten, die ein Studienprogramm nachfragen (sollen). Aber auch Ministerien und Parlament können wichtige Zielgruppen sein. Aus der Vielfalt der Zielgruppen können konfligierende Interessen resultieren, die in einem mehr oder weniger politischen Aushandlungsprozeß geklärt oder gewichtet werden müssen, da in einem solchen Fall nicht alle Ziele gleichermaßen verfolgt werden können. Hierbei könnte auf Modelle zurückgegriffen werden, die in nachfolgendem Kapitel genauer erläutert werden.

### **3. Konzepte, Instrumente und Modelle zur Qualitätsorientierung an Hochschulen**

*Ziel dieses Abschnittes ist es, zunächst einen Überblick über die verschiedenen Konzepte zur Qualitätsorientierung an Hochschulen zu gewinnen und diese in eine Ordnung zu bringen. Außerdem sollen Hauptentwicklungslinien skizziert werden, was anhand einer Einteilung in drei Phasen geschieht. Lediglich für ausgewählte Instrumente und Modelle, die nach den Beschreibungen in der vorhandenen Literatur als die am häufigsten verwendeten bzw. bedeutendsten beschrieben werden, sollen anschließend exemplarisch deren Funktionsweisen sowie deren Vor- und Nachteile als Instrumente zur Förderung der Qualitätsorientierung erläutert werden. Ziel ist auch dabei vor allem, eine grundsätzliche Einordnung zu ermöglichen. Es kann hier nicht das Ziel sein, diese im Detail in ihren verschiedenen Ausprägungen und spezifischen Weiterentwicklungen zu diskutieren; hierzu muß auf die vorhandene Literatur verwiesen werden. Abschließend sollen Möglichkeiten der Weiterentwicklung und gegenseitigen Ergänzung einiger Instrumente und Modelle diskutiert werden.*

Gliederung:

- 3.1 Überblick über Konzepte, Instrumente und Modelle
  - 3.1.1 Überblick über Konzepte und Versuch einer Systematisierung
  - 3.1.2 Überblick über Instrumente und Modelle zur Förderung der Qualitätsorientierung
- 3.2 Leistungs- und Qualitätsbewertung
  - 3.2.1 Evaluation
  - 3.2.2 Berichtssysteme
  - 3.2.3 Rankings bzw. vergleichende Studienführer
- 3.3 Modelle des Qualitätsmanagements
  - 3.3.1 DIN ISO 9000ff
  - 3.3.2 Total Quality Management
  - 3.3.3 Balanced Scorecard
- 3.4 Leistungsanreizmodelle
  - 3.4.1 Nichtmaterielle Anreize
  - 3.4.2 Leistungsabhängige Besoldungszulagen
  - 3.4.3 Leistungsorientierte Verteilung finanzieller Mittel bzw. Ressourcen
- 3.5 Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen
  - 3.5.1 Schwächen von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen als Ausgangspunkt
  - 3.5.2 Gegenüberstellung theoretischer Überlegungen und eingesetzter bzw. diskutierter Indikatoren der leistungsorientierten Mittelverteilung an ausgewählten Hochschulen
  - 3.5.3 Indikatoren zur Leistungs- bzw. Qualitätsbewertung in mehrdimensionalen Rankings/ vergleichenden Studienführern
  - 3.5.4 Fazit

## 3.1 Überblick über Konzepte, Instrumente und Modelle

### 3.1.1 Überblick über Konzepte und Versuch einer Systematisierung

In verschiedenen Publikationen zum Thema Qualität an Hochschulen ist von Konzepten der Qualitätsorientierung, Qualitätssicherung<sup>41</sup>, der Qualitätsentwicklung<sup>42</sup> und des Qualitätsmanagements<sup>43</sup> die Rede. Diese Begriffe werden jedoch in unterschiedlichen Bedeutungen verwendet. Daher soll auch hier zunächst eine Klärung dieser Begriffe erfolgen. Im Alltagssprachgebrauch an den Hochschulen wird häufig der Begriff Qualitätssicherung als eine Art „Sammelbegriff“ für sämtliche Maßnahmen verwendet, die Qualität sichern oder verbessern sollen. Teilweise wird Qualitätssicherung aber auch als spezielle Form von Qualitätsmanagement aufgefasst (vgl. z.B. König 2000: 33). An anderer Stelle wird „Qualitätssicherung bewusst als Teil eines umfassenden Modells“ zur „Steuerung und Kontrolle (Controlling im eigentlichen Sinne)“ konzipiert (vgl. HRK 2001a: 25). Da die Hochschulqualitätsdebatte begrifflich insgesamt durch Unübersichtlichkeit gekennzeichnet ist, wird nachfolgend versucht, zwischen diesen Begriffen klarer zu unterscheiden und diese auch in eine Ordnung zu bringen.

Das derzeit wohl umfassendste Strukturmodell für die Qualitätsorientierung an Hochschulen liefert Pasternak (2004: 8 ff. – vgl. nachfolgende Abbildung). Da dieses Strukturmodell viele der verwendeten Begriffe integriert und bis auf die vierte Ebene („Konzepte, Quasikonzepte und Instrumente“ – darauf wird noch zurückzukommen sein) als stimmig erscheint, soll dieses Strukturmodell sozusagen als eine Art Ausgangsbasis für die angestrebte Systematisierung genutzt werden, die aber noch konstruktiv weiterentwickelt werden sollte. Im Unterschied zum Alltagssprachgebrauch wird von Pasternack „Qualitätsorientierung“ als allgemeiner Oberbegriff verwendet. Darunter wird „eine inhaltliche wie praktische Ausrichtung von Denken und Handeln auf Qualität“ verstanden (vgl. Pasternack 2004: 9).

---

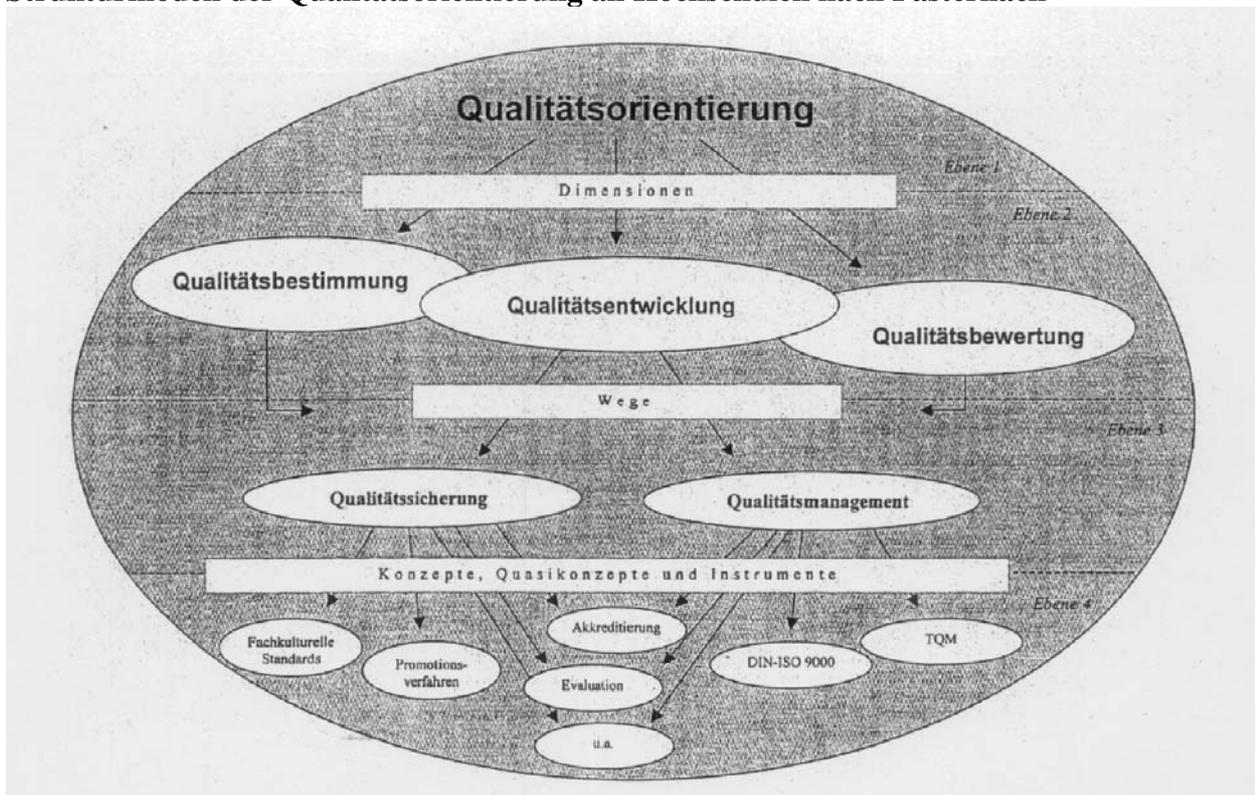
<sup>41</sup> Beispielsweise wies der damalige Sächsische Wissenschaftsminister Meyer auf das divergierende Begriffsverständnis von Qualitätssicherung hin (vgl. HRK 1999b: 17). Unabhängig davon erfreut sich der Begriff großer Beliebtheit: In Tagungstiteln der Hochschulrektorenkonferenz wird der Begriff Qualitätssicherung allein in den Jahren 1998 bis 2000 mindestens fünfmal verwendet (vgl. HRK 1998b, 1999b, 1999d, 2000a, 2000d). Der Stifterverband für die Deutsche Wissenschaft verwendete den Titel „Qualitätssicherung an Hochschulen“ für eine Veröffentlichung zu Empfehlungen zur Durchführung mehrstufiger Evaluationsverfahren in Studium und Lehre.

<sup>42</sup> So führte der Stifterverband für die Deutsche Wissenschaft am 9. Januar 1997 im Wissenschaftszentrum Bonn ein Symposium zum Thema „Qualitätsentwicklung in einem differenzierten Hochschulsystem“ durch, wobei sich viele Beiträge mit Evaluationsverfahren beschäftigten. Andere Autoren verstehen unter Qualitätsentwicklung dagegen: „Qualitätsentwicklung umfasst nicht nur die Diagnose der Lehrqualität, sondern auch die darauf aufbauende ‚Therapie‘ von didaktischen Defiziten bzw. die Weiterentwicklung der kommunikativen Fähigkeiten“ (vgl. z.B. Winter 2000: 190). Der Begriff Qualitätsentwicklung wurde ebenfalls mehrfach in Tagungstiteln der Hochschulrektorenkonferenz verwendet (vgl. z.B. HRK 1999e, 2001a).

<sup>43</sup> Auch der Begriff Qualitätsmanagement taucht in mehreren Tagungstiteln der Hochschulrektorenkonferenz mit durchaus unterschiedlichen Akzentsetzungen auf (vgl. z.B. HRK 1998a, 2000b, 2001b).

Die Qualitätsorientierung bildet in diesem Strukturmodell die erste Ebene und wird auf der zweiten Ebene innerhalb dreier Dimensionen umgesetzt: Qualitätsbestimmung (Was ist das Ziel?), Qualitätsentwicklung (Wie wird das Ziel erreicht?) und Qualitätsbewertung (Inwieweit ist das Ziel erreicht worden?). Die Qualitätssicherung ist *neben* dem Qualitätsmanagement auf der dritten Ebene angesiedelt. Beide bilden gleichwertige, aber verschiedenartige Wege der Umsetzung von Qualitätsorientierung. Mit dieser Unterscheidung soll zugleich die hochschulspezifische Qualitätssicherung vom nichthochschulspezifischen Qualitätsmanagement abgesetzt werden, wobei im folgenden Instrumente des Qualitätsmanagements genannt werden, die als in Teilen adaptionsfähig eingestuft werden. Begründet wird die Ansiedelung von Qualitätssicherung *neben* Qualitätsmanagement damit, daß an Hochschulen Qualitätsentwicklung durchaus auch vorkommen kann, ohne daß diese als Qualitätsmanagement realisiert wird.

### Strukturmodell der Qualitätsorientierung an Hochschulen nach Pasternack



Quelle: Pasternack (2004: 9)

Im einzelnen werden unter Qualitätsmanagement v.a. betriebswirtschaftlich inspirierte Konzepte verstanden, die darauf zielen, auf bestimmten Wegen die jeweilige Prozeßqualität zu verbessern, um eine standardisierte Produktqualität zu sichern, zu erreichen bzw. zu überbieten. Qualitätssicherung wurde definiert als „Handlungen, die der Erhaltung und Entwicklung von Qualität dienen, ohne daß dem ein Qualitätsmanagement-Konzept zugrunde liegt (sondern z.B. fach-

kulturelle Standards)“ (vgl. Pasternack 2004: 10). Es wird außerdem noch eine vierte Ebene angeführt, die die Konzepte und Instrumente beinhaltet, die zur Umsetzung der Qualitätsorientierung eingesetzt werden. Hierbei werden z.B. Promotionsverfahren der Qualitätssicherung zugeordnet, ISO 9000 und TQM werden dem Qualitätsmanagement zugeteilt. Evaluation, Akkreditierung sowie „u.a.“ sind in der Mitte zwischen Qualitätssicherung und Qualitätsmanagement angeordnet und durch Pfeile beiden Wegen der Umsetzung von Qualitätsorientierung zugeteilt.

Bis auf die vierte Ebene soll dem Strukturmodell<sup>44</sup> von Pasternack gefolgt werden. Die Zuordnungen der vierten Ebene erscheinen für diese Arbeit vor dem Hintergrund der hier in dem vorhergehenden Abschnitt zur Klärung des Begriffes Evaluation vorgenommenen Definition nicht vertretbar, da Evaluation grundsätzlich nicht als den betriebswirtschaftlich inspirierten Konzepten zuzuordnen angesehen wird. Demzufolge müßte sie „Handlungen, die der Erhaltung und Entwicklung von Qualität dienen, ohne daß dem ein Qualitätsmanagement-Konzept zugrunde liegt“ zugehören und wäre ein Konzept der Qualitätssicherung. Da dies der Zuordnung von Pasternack widerspricht, wird nachfolgend der Versuch einer anderen Strukturierung vorgenommen, wobei die Begriffe Qualitätssicherung und Qualitätsmanagement dort, wo sie verwendet werden, annähernd in der von Pasternack formulierten Bedeutung verwendet werden.

### **3.1.2 Überblick über Instrumente und Modelle zur Förderung der Qualitätsorientierung**

Wie bereits erwähnt, ist es hier *nicht* das Ziel, allumfassend „das (alte und neue) Instrumentenarsenal zu mustern“, wie es z.B. im Summary von Pasternack (2004: 1) formuliert wird. Daher wird hier auch nicht der Versuch unternommen, möglichst viele Instrumente in ihren verschiedenen Ausprägungen und spezifischen Weiterentwicklungen zu diskutieren.<sup>45</sup> Anstelle dessen soll nur eine Auswahl besonders häufig verwendeter bzw. für diesen Forschungszusammenhang als besonders wichtig erachteter Instrumente und Modelle vorgestellt und eingeordnet werden. Dabei spielt eine Rolle, welcher Anspruch bzw. welches Verständnis von Qualität den jeweiligen Instrumenten und Modellen zugrunde liegt. So werden hier insbesondere Anreizmodelle sehr viel ausführlicher als z.B. bei Pasternack (2004) vorgestellt und diskutiert, da mit diesen der Anspruch einer Leistungssteigerung bei zu haltender bzw. zu steigender

---

<sup>44</sup> Aufbauend auf dem Strukturmodell wird von Pasternack noch ein Prozeßmodell entwickelt, in das er eine Reihe von Instrumentenbeispielen einordnet. Da es hier v.a. um einen Überblick und die Einordnung des Begriffes Qualitätssicherung geht, soll dieses relativ detaillierte Prozeßmodell hier nicht wiedergegeben, sondern auf die umfassenden Erläuterungen bei Pasternack (2004: 11 ff.) lediglich verwiesen werden.

<sup>45</sup> Pasternack (2004) stellt insgesamt 46 verschiedene Instrumente und Verfahren vor, die in anderen Publikationen durchaus auch einander zugeordnet werden (z.B. Qualitätspreise dem TQM, da sie in diesem Zusammenhang am häufigsten verwendet werden). In dieser Arbeit werden bewusst nur insgesamt neun ausgewählte Instrumente und Modelle vorgestellt und bezüglich ihrer Eignung zur Förderung der Qualitätsentwicklung diskutiert.

Qualität verbunden wird (vgl. Minssen/ Wilkesmann 2003: 106ff., Schröder 2003: 43ff.).<sup>46</sup> Teilweise werden (formelgebundene) leistungsorientierte Mittelvergabe und Zielvereinbarungen auch als komplementäre Instrumente zur Erreichung desselben Ziels genannt (vgl. Ziegele 2003b: 3).

Bevor jedoch im einzelnen näher auf bestimmte Modelle und eine Strukturierung dieser Modelle eingegangen wird, soll noch kurz auf die Begrifflichkeiten „Instrumente“ und „Modelle“ eingegangen werden, da diese in der Literatur durchaus mit unterschiedlichen Bedeutungen verwendet werden und auch bei Pasternack nicht genauer definiert wurden.<sup>47</sup> In dieser Arbeit soll unter *Modellen* das verstanden werden, was mittels vereinfachender (also im wörtlichen Sinne *modellhafter*) Annahmen z.B. (gedachte) Wirkungsverläufe nachzeichnet und/ oder in Form eines übersichtlichen Erfassungsschemas darstellt.<sup>48</sup> Solche Modelle müssen nicht unbedingt bis in Details ausformuliert, erprobt und/ oder in der Praxis bewährt sein. (Sie können sich auch noch in der Erprobungs- bzw. *Modellprojektphase* befinden.) Damit sollten sowohl Qualitätsmanagementmodelle, als auch Leistungsanreizmodelle erfaßt werden. Unter *Instrumenten* sollen im Gegensatz dazu konkretere Maßnahmen verstanden werden, die bereits erprobt und/ oder bewährt sind. Nach diesem Verständnis wären Evaluationsmaßnahmen und Lehrberichte den Instrumenten zur Förderung der Qualitätsorientierung zuzuordnen. Letztlich bleibt es aber immer eine Frage des Ziels der Systematisierung, ob diese Zuordnung angemessen ist.

Für die Strukturierung der nachfolgenden Vorstellung von Modellen und Instrumenten zur Förderung der Qualitätsorientierung soll einem Eindruck des Verfassers gefolgt werden, der sich in der praktischen Tätigkeit im Bereich der Qualitätssicherung an Hochschulen entwickelte und

---

<sup>46</sup> Dieser Anspruch an Modelle leistungsorientierter Mittelvergabe wird nicht nur von außen (z.B. seitens der Politik) an die Hochschulen herangetragen, sondern z. T. auch von diesen selbst in ihrer Beschreibung von Modellen zur leistungsorientierten Mittelvergabe formuliert (vgl. Universität Oldenburg 2000, Universität Dortmund 2000: 1).

<sup>47</sup> Beispielsweise spricht Pasternack (2004: 4) von *Instrumenten* des Qualitätsmanagements, ordnet jedoch das *Modell EFQM* und *Leistungsanreizmodelle* diesen hinzu, während Evaluation nicht wie von anderen Autoren den *Instrumenten* der Qualitätsentwicklung, sondern den Verfahren der Qualitätsbewertung zugeordnet wird. Pasternack (2004: 18) begründet dies allerdings und fügte hinzu, dies geschehe ausschließlich im Rahmen einer policy-analytischen Betrachtung.

<sup>48</sup> Nach Stachowiak (1973: 129) ist ein Modell alltagssprachlich ein Abbild von etwas, ein Vorbild für etwas oder die Repräsentation eines bestimmten Originals. Als die drei Hauptmerkmale des allgemeinen Modellbegriffs formuliert Stachowiak (1973: 131ff.) das Abbildungsmerkmal (s.o), das Verkürzungsmerkmal und das pragmatische Merkmal. Das Verkürzungsmerkmal beschreibt, daß Modelle im Allgemeinen nicht alle Attribute des durch sie repräsentierten Originals erfassen, sondern nur solche, die den jeweiligen Modellerschaffern und/ oder Modellbenutzern relevant scheinen. Das pragmatische Merkmal meint, daß Modelle ihren Originalen nicht per se zugeordnet sind. Sie erfüllen ihre Ersetzungsfunktion a) für *bestimmte* – erkennende und/ oder handelnde, modellbenutzende – *Subjekte*, b) innerhalb *bestimmter Zeitintervalle* und c) unter Einschränkung auf *bestimmte gedankliche oder tatsächliche Operationen*. Modelle sind damit nicht nur Modelle von etwas oder für etwas, sondern auch für bestimmte Personen zu einer bestimmten Zeit und einem bestimmten Zweck (zur Anwendung im Bildungsbereich vgl. auch Stachowiak 1980: 29, Forberg 2002: 7ff.).

bei der weiteren Sichtung der Literatur zum Thema verfestigte<sup>49</sup>: Bestimmte Instrumente wurden besonders in der ersten Zeit der Qualitätsdebatte thematisiert, während andere erst in der darauf folgenden Zeit stärker in die Diskussion einfließen. Daher soll versucht werden, diese zeitliche Abfolge nachzuvollziehen. Für die Strukturierung der Entwicklung lassen sich also – dem Eindruck folgend, der durch die Titel von Publikationen, Tagungen usw. über die im Vordergrund stehenden Themen in den jeweiligen Zeitabschnitten entstand – die Hauptentwicklungslinien grob in drei Phasen<sup>50</sup> einteilen:

Die *erste Phase* begann im wesentlichen Ende der 1980er Jahre<sup>51</sup>. Diese Phase könnte man beschreiben als v.a. durch die grundsätzliche Thematisierung und Entwicklung des Aspektes der *Leistungs- bzw. Qualitätsbewertung* an den Hochschulen bzw. für die Hochschulen gekennzeichnet. Hierbei wurde jedoch festgestellt, daß bestimmte Instrumente wie die Lehrveranstaltungsevaluation, aber auch Lehrberichte nicht die erhofften Effekte zeigten und daher ein umfassenderer Ansatz gefordert (vgl. z.B. Barz 1997: 17, Rindermann 1996 sowie 2004: 83-96). Die *zweite Phase* könnte man daher beschreiben als stärker gekennzeichnet durch Aspekte der Entwicklung von umfassenderen *Qualitätsmanagementmodellen*, also einer Verstärkung des Einflusses betriebswirtschaftlich inspirierter Konzepte (seit etwa Mitte/ Ende der 1990er Jahre)<sup>52</sup>. Wolter (1999: 26) spricht in diesem Zusammenhang von einem „neuen betriebswirtschaftlichen Paradigma der Hochschulsteuerung“ (siehe auch weiter oben im Abschnitt zur Entwicklung des Qualitätsbegriffs in der „Dienstleistungsgesellschaft“) Die Abgrenzung der Phasen soll daher auch mit diesem (beginnenden) Paradigmenwechsel begründet werden.

---

<sup>49</sup> Bei Pasternack (2004: 4ff.) erfolgte die Strukturierung allerdings nach einer anderen Systematik: Er gliederte einerseits in Instrumente der Qualitätsentwicklung (und darunter in traditionelle Qualitätssicherungsinstrumente sowie adaptionsfähige Qualitätsmanagementinstrumente) und andererseits in Verfahren der Qualitätsbewertung (und darunter in Bewertungsprogramme sowie -technologien). Durch diese Gliederungen tauchen aber z.B. Leistungsanreizmodelle unter adaptionsfähigen Qualitätsmanagementinstrumenten auf, während leistungsorientierte Besoldung unter Bewertungstechnologien eingeordnet wird. Neben der begrifflichen Unklarheit von Modellen und Instrumenten erscheint diese Einordnung schwer nachvollziehbar. Die Besoldungsreform verfolgt v.a. das Ziel einer stärkeren Anreizwirkung, und der Aspekt der Einführung von Leistungszulagen wird als „der Dreh- und Angelpunkt der Besoldungsreform gesehen“ (vgl. z.B. Arnold/ Kandel 2004: 17).

<sup>50</sup> Die einzelnen Instrumente lassen sich nicht immer klar einer Phase zuordnen, da die Entwicklung sowohl in den einzelnen Bundesländern als auch an den einzelnen Hochschulen sehr unterschiedlich verläuft. Dennoch nahm auch Pasternack (2004: 11 ff.) eine grobe Strukturierung (dort in Qualitätsbestimmung, Qualitätsentwicklung und Qualitätsbewertung) vor und merkte hierzu an, daß die Zuordnung nicht immer eindeutig möglich ist. Deshalb, und weil die Diskussion zeitlich nachgelagert einbezogener Instrumente dadurch leichter verständlich wird, soll trotz der genannten Schwierigkeit zur Erleichterung der Übersicht eine grobe Einteilung erfolgen.

<sup>51</sup> So formulieren mehrere Autoren, daß Ende der 1980er Jahre, also etwa zeitgleich mit der Veröffentlichung des Spiegel-Rankings, eine öffentliche Debatte über die Qualität der Hochschullehre und systematische Versuche ihrer Bewertung begannen (vgl. HIS 1992, Barz 1997: 15).

<sup>52</sup> Ende der 1990er Jahre fanden mehrere Tagungen der Hochschulrektorenkonferenz speziell zum Thema „Qualitätsmanagement in der Lehre“ statt (vgl. HRK 1998a, HRK 2000b). Auch Pasternack (2001: 267) thematisiert die zunehmende Integration betriebswirtschaftlicher Ansätze und die entsprechende Erweiterung des Reformvokabulars in diesem Zeitraum.

*Die dritte Phase* lässt sich beschreiben als charakterisiert durch eine verstärkte Diskussion von Leistungsanreizmodellen. Diese Diskussion begann etwa Ende der 1990er Jahre und manifestierte sich mit dem Inkrafttreten der Professorenbesoldungsreform im Februar 2002. Spätestens seit diesem Zeitpunkt ging es de facto nicht mehr darum ob, sondern wie Leistungsanreize an den Hochschulen umgesetzt werden.

Die nachfolgende Übersicht über Instrumente bzw. Modelle der Qualitätsorientierung erhebt – darauf soll nochmals hingewiesen werden – keinen Anspruch auf Vollständigkeit. Dies dürfte aufgrund der inzwischen erkennbaren Vielfalt und wegen der gewünschten Profilbildung an den Hochschulen, die sich ebenfalls in der Qualitätssicherung niederschlägt, auch immer schwerer möglich sein. Vielmehr soll die Übersicht v.a. das Verständnis der wesentlichen Entwicklungen angesichts der Vielfalt der Instrumente erleichtern.

**Erste Phase: Verstärkte Entwicklung der Leistungs- bzw. Qualitätsbewertung**

- Evaluation (formativ und summativ, intern und extern)
- Berichtssysteme (z.B. Lehrberichte)
- Rankings (z.B. zur bzw. einschließlich der Bewertung von Studienbedingungen)

**Zweite Phase: Verstärkte Einbeziehung von Modellen des Qualitätsmanagements**

- DIN ISO 9000 ff. als Form der Zertifizierung
- Total Quality Management (TQM) und EFQM als spezielle Form von TQM
- Balanced Scorecard (BSC)

**Dritte Phase: Verstärkte Diskussion von Leistungsanreizmodellen**

- nichtmaterielle Anreize
- leistungsorientierte Besoldungszulagen
- leistungsorientierte Vergabe finanzieller Mittel bzw. Ressourcen

## 3.2 Leistungs- und Qualitätsbewertung

Die der ersten Phase zugeordneten Begriffe Evaluation, Berichtssysteme und Rankings sind als Bezeichnungen für Instrumente zur Qualitätsbewertung, wie sie im Hochschulalltag verwendet werden, in der Praxis nicht immer klar voneinander abgegrenzt. So wird eine interne Evaluation, für die ein Ergebnisbericht erstellt wird und der auch als Evaluationsbericht bezeichnet werden kann (vgl. Barz u.a. 1997: 62), z.T. als Berichtssystem eingeordnet. Umgekehrt wird mancherorts die Erstellung eines Lehrberichtes als Evaluation bezeichnet. (Hochschul-)Rankings werden z.T. auch als summative Evaluation eingeordnet, da sie Qualität nicht vorrangig verbessern, sondern zunächst vor allem (Unterschiede in der) Qualität feststellen wollen.

Weiter vorn in dieser Arbeit wurde – wie auch zuvor von anderen Autoren – die Bedeutung der Zielangemessenheit bei Instrumenten der Qualitätssicherung betont. Um zu einer Systematisierung der Instrumente zu kommen, soll daher nachfolgend von der vorrangig verfolgten Zielstellung abhängig gemacht werden, wie die Instrumente eingeordnet werden.<sup>53</sup> Dabei können die folgenden Ziele verfolgt werden:

- Evaluation verfolgt immer auch das Ziel, etwas zu verbessern.
- Berichtssysteme verfolgen v.a. das Ziel, Rechenschaft abzulegen.
- Rankings sollen v.a. bestimmten Zielgruppen vergleichende Entscheidungshilfen liefern.

Auf diese Weise kann die Erstellung eines Evaluationsberichtes, der als „Nebenprodukt“ einer Evaluation entsteht und *nicht* der Rechenschaftspflicht z.B. gegenüber dem Ministerium dient, als Evaluationsinstrument eingeordnet werden. Ein als Berichtssystem bezeichnetes Lehrberichtsverfahren, was vor allem der Verbesserung der Studiensituation dienen soll, wäre dann keines, sondern eine Evaluation.<sup>54</sup> Dagegen ist nach dieser Systematik die Erstellung eines Lehrberichtes, der v.a. der Rechenschaftslegung dient, den Berichtssystemen zuzuordnen. Den Rankings werden demzufolge nur solche Instrumente zugeordnet, die die Zielgruppenangemessenheit beachten, wie es das CHE z.B. durch die Möglichkeit der persönlichen Gewichtung von Kriterien im Jahr 2000 zunächst mit einer CD-ROM und seit 2001 mit seiner Internetveröffentlichung umsetzte.

---

<sup>53</sup> In der Praxis wird es vermutlich Instrumente geben, bei denen keine eindeutige Zuordnung möglich ist. Hier geht es jedoch um eine Systematisierung der Ansätze und um das Herausarbeiten von Problemen verschiedener Ansätze.

<sup>54</sup> So wurden Lehrberichte – jedenfalls in dieser Phase der Entwicklung Anfang bis Mitte der 1990er Jahre – in einer Bestandsaufnahme zur Evaluation an deutschen Hochschulen v.a. als Form der Selbstevaluation eingestuft (vgl. Müller-Böling 1996: 5 ff.).

Bezüglich der Instrumente bzw. Verfahren der Qualitätsbewertung wird z.T. noch zwischen Bewertungsprogrammen (vgl. Pasternack 2004: 69ff.) und -technologien (2004: 89ff.) unterscheiden. Da es jedoch Überschneidungen gibt (Rankings z.B. werden unter beidem genannt, bei ersterem jedoch als komplexe Rankings) und die doppelte Zuordnung nicht genauer begründet wurde, wird eine solche Unterscheidung nicht vorgenommen. Hier wird – um das Beispiel Rankings aufzugreifen – generell davon ausgegangen, daß sich nur komplexe Rankings mit mehreren Indikatoren, die möglichst verschiedene Dimensionen von Qualität (insbesondere objektiv/ subjektiv) erfassen, als Instrument der Qualitätsbewertung eignen.<sup>55 56</sup> Im Abschnitt 3.5 am Ende dieses Kapitels werden daher in die Überlegungen zu Möglichkeiten der Weiterentwicklung bzw. Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen auch nur mehrdimensionale Rankings/ Studienführer einbezogen.

### 3.2.1 Evaluation

Was unter dem Begriff *Evaluation* verstanden werden soll, wurde bereits weiter vorn im Kapitel zu Begriffsklärungen erläutert. An dieser Stelle soll lediglich noch einmal daran erinnert werden, daß im Zusammenhang mit der Definition des Begriffes häufig als Ziel bzw. Kernelement von Evaluation formuliert wurde, daß sie nicht allein der Erforschung eines Gegenstandes dienen soll, sondern auch dazu, etwas zu verbessern oder zu entscheiden (vgl. Wottawa & Thierau 1998: 14).<sup>57</sup> Dieses Ziel war für die Abgrenzung zu Berichtssystemen leitend.

In diesem Abschnitt soll Evaluation als Instrument der Qualitätssicherung betrachtet werden. Entsprechend den Zielstellungen der Evaluation kann unterschieden werden in *formative* und *summative* Evaluation. Diese Unterscheidung wird z.T. unterschiedlich akzentuiert: „*Summative* Evaluation liefert Resultate über den Zustand der Lehre, *formative* Evaluation möchte durch die Erhebung den Zustand der Lehre verbessern“ (Vgl. Rindermann 1994: 40). Bortz/ Döring (2002: 112 ff.) verstehen unter summativer Evaluation eine „Hypothesenprüfung“, die „typischerweise vorgenommen (wird), nachdem die Maßnahme abgeschlossen ist“. Formative Evaluationen, „die vor allem bei der Entwicklung und Implementierung neuer Maßnahmen eingesetzt werden, sind im Unterschied zur summativen Evaluation meistens erkundend angelegt“. Die Gemeinsamkeit

---

<sup>55</sup> Pasternack (2004: 102) weist in Bezug auf Rankings als Bewertungstechnologie auch selbst darauf hin, daß solche, die „allein mit einem einzigen Indikator arbeiten, hinsichtlich ihrer Aussagekraft mit größter Vorsicht zu behandeln sind“ und daher möglichst auf komplexe Rankings orientiert werden sollte.

<sup>56</sup> Außerdem nennt Pasternack (2004: 69) als das traditionell bedeutsamste Qualitätsbewertungsprogramm die „systematische Gutachtung“. Da diese jedoch v.a. zur Einwerbung von Drittmitteln im Bereich der Forschung angewendet wird und hier der Bereich der Lehre im Vordergrund stehen soll, findet dies keine Berücksichtigung.

<sup>57</sup> Andere Autoren (vgl. z.B. Stockmann 2002b: 227) formulieren etwas zurückhaltender, daß im am weitesten verbreiteten Paradigma Evaluation „ein spezielles Erkenntnis- und Verwertungsinteresse hat, bei dem der Nutzen der Evaluationsergebnisse für die ‚Praxis‘ im Vordergrund steht“ (vgl. auch Kromrey 2001: 113).

liegt aber darin, daß es bei der summativen Evaluation v.a. um die Feststellung von Ergebnissen geht, während die formative Evaluation auf die Identifizierung von Wirkverläufen bzw. deren Verbesserung abzielt. In der Praxis sind beide Arten der Evaluation nicht immer scharf voneinander abzugrenzen (vgl. Pasternack 2004: 77).

Methodisch wird Evaluation außerdem unterschieden in *interne* Evaluation (innerhalb eines Fachbereiches bzw. einer Hochschule stattfindend) und *externe* Evaluation. Hierbei werden in Anlehnung an das niederländische Vorbild die umfangreiche Selbstevaluation in eigener Regie eines Faches und die Fremdevaluation durch auswärtige Experten (insbesondere durch sogenannte Peers) verbunden (vgl. HIS 1995a, Barz u.a. 1997: 18, HIS 1998a). Ihre Kombination gilt als ein Weg, mit der Schwierigkeit von Qualitätsbewertungen umzugehen: Die Doppelperspektive, so die Annahme, erhöhe die Chance, Qualitäten identifizieren zu können (vgl. Pasternack 2004: 76). In den letzten Jahren ist zunehmend von einem zwei- bis dreistufigen Modell die Rede. Als dritte Stufe, die die interne und externe Evaluation ergänzen soll, wird dabei von der HRK die Zielvereinbarung zwischen Fachbereich und Hochschulleitung über Maßnahmen zur Behebung von erkannten Defiziten bezeichnet (vgl. HRK 2001b: 10). In den letzten Jahren wird diese dritte Stufe auch als „Follow up“ bezeichnet (vgl. HRK 2002: 83ff.), wobei hierunter nicht nur Zielvereinbarungen, sondern der gesamte Prozeß einschließlich der Umsetzung von Empfehlungen aus Evaluationen verstanden wird (vgl. HRK 2004: 97ff.).

Evaluation könnte im weiteren auch nach den Bereichen unterschieden werden, in denen sie Anwendung findet (Lehre, Forschung und Verwaltung) oder nach ihren Gegenständen (Lehrveranstaltungsevaluation und Studienfachevaluation). In dieser Arbeit steht jedoch der Bereich der Lehre im Vordergrund. Unterschiede, die sich aus der Unterscheidung in Lehrveranstaltungsevaluation und Studienfachevaluation ergeben, wurden bereits im Zusammenhang mit der Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen bei der Diskussion von Lehrqualität auf Studiengangsebene und Lehrveranstaltungsebene angesprochen. Was im einzelnen bei der Durchführung von Evaluationen bestimmter Gegenstände zu beachten wäre, kann an dieser Stelle nicht erläutert werden. Hierzu muß auf die umfassende Literatur zu diesem Themenbereich verwiesen werden, zu der sich einige Hinweise im Abschnitt 2.3 finden.

### **3.2.2 Berichtssysteme**

*Berichtssysteme* zeichnen sich meist dadurch aus, daß sie von außen auferlegt, aber von innen zu realisieren sind (Rechenschaftspflicht). Neben dem jährlichen Haushaltsabschluß existieren eine Vielzahl von inhaltlich orientierten, d.h. die Erfüllung des Leistungsauftrages der Hochschule thematisierenden Berichtsformen. Zum Berichtssystem wird derartige dann, wenn mehrere

Formen und Techniken miteinander kombiniert werden, um die jeweils spezifischen Vorteile zu kumulieren bzw. die spezifischen Nachteile zu neutralisieren (vgl. Pasternack 2004: 72).<sup>58</sup> Dies können z.B. Berichte des Rektoratskollegiums, Lehrberichte, Forschungsberichte usw. sein. Die Einordnung einzelner Berichtssysteme, z.B. der Lehrberichte, ist schwierig, da diese sich teilweise sehr stark voneinander unterscheiden, worauf bereits Hage (1996: 35) in einem Bericht für das Bundesministerium für Bildung und Forschung (bmb+f) hinwies und was Barz u.a. (1997: 22 ff.) noch detaillierter herausarbeiteten. Diese Situation hat sich in den letzten Jahren nicht verändert. So kommt die HRK in ihrem Sachstandsbericht 2003 zu der Einschätzung, daß die Lehrberichte von Land zu Land so stark voneinander abweichen, daß trotz der gemeinsamen Bezeichnung kaum von einem einheitlichen oder gar eigenständigen Instrument der Qualitätssicherung gesprochen werden könne (vgl. HRK 7/2003: 13). Aber auch wenn die Ausgestaltung und Nutzung von Lehrberichten sehr unterschiedlich ausfallen, spielen solche Berichte als institutionalisierter „Reflektionsprozeß“ an Hochschulen durchaus eine Rolle (vgl. z.B. HRK 2000d: 50). Während es in den Anfangsjahren v.a. darum ging, welche Indikatoren solche Lehrberichte für eine umfassende Darstellung der Situation der Lehre enthalten sollten (vgl. Webler u.a. 1993, Barz u.a. 1997: 17), machte in jüngerer Zeit das CHE Vorschläge, in welchem Umfang der Staat (maximal) welche Informationen von seinen Hochschulen verlangen sollte (vgl. Müller/ Ziegele 2003: 8 ff., Ziegele 2003: 5 ff.), „ohne in die alten Muster staatlicher Aufsicht zurückzufallen“, wie Pasternack (2004: 72) hierzu formulierte. Hierbei wurde von Ziegele (2003: 10 ff.) die Idee einer Informationspyramide (vgl. Barz u.a. 1997: 62) aufgegriffen und weiter differenziert, wonach es eine zunehmende Aggregation der Informationen vom „Fuß“ der Pyramide (Institut, Lehrstuhl) über Fachbereichsleitung, Hochschulleitung, Ministerium, Landesregierung und Parlament bis hin zur Öffentlichkeit an der Spitze der Pyramide geben sollte.<sup>59</sup> Diese Idee einer differenzierten Informationspyramide ermöglicht zwar nicht die ursprünglich von Barz u.a. vorgeschlagene relativ klare Trennung der Instrumente in Evaluationsbericht (intern, die v.a. zur Verbesserung der Situation am Fachbereich dienen sollen) und Lehrbericht (zur Rechenschaftslegung für höhere Ebenen) und erschwert damit den

---

<sup>58</sup> Der Begriff Berichtssystem wurde auch bereits etliche Jahre zuvor von Reißert in ähnlichem Zusammenhang verwendet. Dort wird als Ziel neben der modellhaften Entwicklung des Berichtssystems, der Herstellung von Transparenz, der „Früherkennung“ von Problemen und der Schaffung einer soliden Informationsbasis für Entscheidungen aber auch noch die Intensivierung der Selbstevaluation genannt (vgl. Gralki u.a. 1993: 110). Genau letzteres Ziel konfliktierte jedoch, wie sich in den Folgejahren herausstellte, mit anderen, da eine kritische Selbstreflektion schwer z.B. mit der Herstellung von Transparenz und der Schaffung einer soliden Informationsbasis für Entscheidungen vereinbar war, wenn dies z.B. Entscheidungen über Mitteleinsparungen sein konnten.

<sup>59</sup> Auch Bieri (vgl. HRK 2001a: 29) griff diese Idee der Informationspyramide auf und bezeichnete sie als „Stufengerechtigkeit“, adaptierte sie allerdings nur für Informationen innerhalb einer Hochschule (ETH Zürich). Als neues Element führte er ein, daß sich die Informationen auf den unteren Ebenen am Pyramidenfuß stärker an Effizienz, auf den oberen Ebenen an der Pyramidenspitze stärker auf die Effektivität konzentrieren sollten.

hier verfolgten Versuch der Systematisierung. Die von Ziegele vorgeschlagene stärkere Differenzierung wird aber vermutlich den unterschiedlichen Informationsbedürfnissen der verschiedenen von ihm genannten Ebenen eher gerecht. Außerdem wurde neben der Neustrukturierung der Berichtsarchitektur (vgl. Ziegele 2003: 22 ff.) gefordert, die Berichte zu Zielvereinbarungen und die internen Berichte stärker aufeinander abzustimmen, um den Gesamtaufwand zu minimieren und konsistente Steuerungsinstrumentarien zu schaffen (vgl. Müller/ Ziegele 2003: 4).

Ohne die Entwicklungen verkennen zu wollen, die sich aus der Umgestaltung der Hochschulen nach dem sogenannten New Public Management (vgl. z.B. Buschor 2002) für Berichtssysteme ergeben können, sollen nachfolgend die Ansätze zur Weiterentwicklung solcher Berichtssysteme nicht weiter erörtert werden, sondern eine Konzentration auf das Instrument der Lehrberichte erfolgen, da hierzu bereits Erfahrungen vorliegen und diese für viele Hochschulen bei aller geäußerten Kritik (v.a. an vom Gesetzgeber definierten Vorgaben) durchaus eine wichtige Grundlage sind, während die von Ziegele u.a. geforderte Neustrukturierung der Berichtssysteme noch in den Anfängen steckt.

### **3.2.3 Rankings bzw. vergleichende Studienführer**

*Rankings* sind eine typische Variante der vergleichenden bzw. komparativen Bewertung (vgl. Pasternack 2004: 101) Hochschulrankings wie z.B. die des Nachrichtenmagazin „Der Spiegel“ (1989) haben erstmals die Studienbedingungen an deutschen Hochschulen für eine breitere Öffentlichkeit thematisiert. Allerdings können solche ersten Veröffentlichungen aufgrund ihrer methodischen Schwächen<sup>60</sup> noch nicht als Qualitätssicherungsinstrumente angesehen werden. Inzwischen wird jedoch vielerorts deren Weiterentwicklung registriert. So wird z.B. eingeschätzt: „Nach Überwindung der Anfangsschwierigkeiten in Bezug auf die Validität haben Rankings inzwischen eine kaum noch angreifbare Grundlage und bilden für die Wahl des Studienortes zunehmend eine Entscheidungshilfe“ (vgl. Behrens 2001: 130). Rankings finden unabhängig von der Zielgruppe, für die sie erstellt werden, in den Hochschulen große

---

<sup>60</sup> Da die erste Spiegel-Untersuchung (1989/90) sehr ausführlich in soziologischen Veröffentlichungen (z.B. in der Zeitschrift „Soziologie“, Ausgaben 2/1990 und 1/1991) diskutiert wurden, dürften die Fehler und Schwächen der Herangehensweise bei dieser Untersuchung in Fachkreisen hinlänglich bekannt sein. Dennoch soll hier in aller Kürze auf die unserer Ansicht nach wesentlichsten Kritikpunkte eingegangen werden. An erster Stelle wäre die Anzahl und die Auswahl der 12 bzw. 18 an jedem Fachbereich zu befragenden Studenten zu nennen. Die Spiegel-Autoren sprechen zwar von einer Zufallsauswahl, davon kann jedoch keine Rede sein. Vielmehr handelt es sich um eine eher willkürliche Auswahl, bei sehr viel gutem Willen um eine Quotenauswahl. (Die Interviewer wurden angewiesen, an mindestens drei verschiedenen Orten - Mensa, Hörsaal, Bibliothek - eine möglichst gleich große Anzahl an Befragten auszuwählen; vgl. Spiegel-special 3/1993.) Bei einer solchen Auswahl der zu Befragenden ist damit zu rechnen, daß mehr oder weniger große systematische Verzerrungen auftreten können.

Beachtung. Sie werden daher, so Behrens weiter, „von nicht unerheblicher Bedeutung für das Profil und auch für den Erfolg von Hochschulen“.

Einzelne, ursprünglich stärker im eigentlichen Sinne als „Rangliste“ konzipierte Hochschulrankings (wie das des Centrum für Hochschulentwicklung – CHE, das sich aber bewusst als Studienführer und nicht als Ranking bezeichnet) werden inzwischen von mehreren Autoren als Qualitätssicherungsinstrument diskutiert, wenn sie bestimmte Anforderungen erfüllen (vgl. z.B. Bayer 2000: 6, Teichler 2003: 5, Pasternack 2004: 102). Als Anforderungen werden hier z.B. genannt, daß Rankings zielgruppenspezifisch ausgewertet und komplex sein sollten, also nicht allein mit einem einzigen Indikator arbeiten dürften.<sup>61</sup> Bayer (2004: 135) weist zudem darauf hin, daß an Rankings insgesamt höhere Anforderungen gestellt würden (und zu stellen sind) als an Evaluationen, da im Unterschied zu einem Hochschul-Ranking bei einer Evaluation die Meßgrößen und die Interpretation der Messung (z.B. bei formativen Evaluationen) nicht immer exakt definiert werden müsse. Bei einer Evaluation würden die Meßgrößen lediglich Richtlinien zur Beurteilung darstellen. Welche Meßgrößen und mit welcher Bedeutung sie tatsächlich einfließen, entscheiden die Evaluatoren.

Ihre Wirkung entfalten solche Rankings nach Bayer (2000: 4) nicht nur direkt über die Leserschaft<sup>62</sup>, sondern auch indirekt über ihren Einsatz als Marketinginstrument durch die Hochschulen<sup>63</sup> und die Diskussion der Ergebnisse in den Hochschulgremien. Teilweise werden Rankings bzw. die Nutzung der Ergebnisse von Rankings nicht nur als Qualitätssicherungsinstrument und zur „Herstellung von Markttransparenz“ (vgl. Müller-Böling 2004: 548), sondern auch für die interne Steuerung von Hochschulen ins Gespräch gebracht.<sup>64</sup> <sup>65</sup> Je mehr solche Rankings zuverlässige Ergebnisse erbringen und damit höhere Akzeptanz erlangen, werden sie –

---

<sup>61</sup> Forschungsrankings, die nur mit einem Indikator (z.B. Drittmittel) arbeiten, sind demnach keine komplexen Rankings. Andere Publikationen, wie die Rankings in den Nachrichtenmagazinen Spiegel und Focus oder der seit 2000 jährlich erscheinende Studienführer Sachsen (vgl. Krempkow/ König 2003), sind zwar komplex, aber nicht zielgruppenspezifisch ausgewertet, da es nur ein und dieselbe Form von Veröffentlichung für alle Leser gibt.

<sup>62</sup> Bayer (2000: 4) nannte hierzu den steigenden Anteil derjenigen Studierenden, denen als Grund für die Wahl ihrer Hochschule wichtig war, daß ihr Studienfach in einer Hochschulrangliste/ Ranking gute Bewertungen erhielt sowie eine Wirkungsstudie zum Spiegel-Ranking 1999, nach der in allen ZVS-Studiengängen ein enger Zusammenhang besteht zwischen Platzierung nach Studierendenurteil und der Zu- bzw. Abnahme der Bewerberzahl. Nach Leszcensky u.a. (2004: 9) wird auch international die Hochschulwahl durch die Einschätzung der Lehrqualität geprägt.

<sup>63</sup> So warben die TU Chemnitz und die TU Bergakademie Freiberg (Fakultät Wirtschaftswissenschaften) – neben wahrscheinlich vielen anderen Hochschulen, die gut abschnitten – mit Spitzenplätzen in CHE-Rankings.

<sup>64</sup> Bereits vor Veröffentlichung der ersten Spiegel-Rangliste ging es Frackmann (1988: 139ff.) darum, die Bedeutung der vergleichenden Informationen von Rankings nicht nur auf die „Markierung von Differenzen“ beschränkt zu sehen, sondern diese auch als Aufgabe der einzelnen Hochschulen hervorzuheben. Demnach seien die Differenzierungen in Leistungsstärkere und Leistungsschwächere für das Gesamtsystem uninteressant, sie würden „erst interessant und wichtig für Einzelentscheidungen der Universitätsleitung, für die Studienwahlentscheidungen eines Studenten, für die Finanzierungsentscheidung eines Drittmittelgebers oder staatlicher Instanzen“.

<sup>65</sup> Teilweise haben Hochschulen bereits bei der Vorstellung und Diskussion der Ergebnisse des CHE-Studienführers (z.B. 2001 in Berlin) geäußert, die Ergebnisse des CHE-Rankings für ihre interne Steuerung nutzen zu wollen. Hierzu können vom CHE Detailauswertungen für die eigene Hochschule angefordert werden.

zumal wenn sie eine gewisse Marktführerstellung einnehmen – attraktiver als Informationsgrundlage für Studienanfänger und Presse- bzw. Marketingabteilungen von Hochschulen oder Fachbereiche, sowie auch in der ‘größeren’ Hochschulpolitik. Dies gilt es zu beachten, da anderenfalls ein u. U. an Bedeutung gewinnender Wirkungsmechanismus<sup>66</sup> aus dem Blick geraten könnte. Daher wurden neben den Instrumenten Evaluation und Berichtssystemen hier auch Rankings mit aufgeführt.

### **3.3 Modelle des Qualitätsmanagements und der Qualitätsorientierung**

Nachfolgend sollen zum als zweite Phase eingeordneten Prozeß der Entwicklung von Qualitätsmanagement an Hochschulen wiederum relativ kurze Erläuterungen für einen Überblick über die am häufigsten diskutierten Modelle erfolgen. Hierbei soll ebenfalls kein Anspruch auf Vollständigkeit erhoben werden, da die Literatur zum Thema ohnehin „nahezu unübersichtlich“ ist (vgl. Stockmann 2002b: 213). Die wichtigsten Modelle des Qualitätsmanagement (QM) lassen sich zwei Gruppen zuordnen: Einerseits dem Standard nach den DIN ISO 9000ff., andererseits dem Total Quality Management (TQM). Diese beiden werden auch nachfolgend im Vordergrund stehen. Zusätzlich wird die Balanced Scorecard (BSC) wie bei Pasternack (2004: 64) in diesem Zusammenhang diskutiert, wobei ihre Möglichkeiten zur Förderung der Qualitätsorientierung im Vordergrund stehen sollen. Eine Einordnung der BSC wäre durchaus auch anders möglich und würde von Wirtschaftswissenschaftlern aus dem Bereich Controlling wahrscheinlich anders erfolgen, da dort stärker der Aspekt der Steuerung betont wird. Steuerungsmodelle stehen jedoch hier nicht im Vordergrund des Interesses, deshalb entspräche dies nicht der gewählten Systematik. Insbesondere wenn die BSC zusätzlich als Indikatoren-system mit monetären Leistungsanreizen verbunden wird, würde aber eine Zuordnung als Leistungsanreizmodell in Frage kommen.<sup>67</sup> Bislang ist dies im Hochschulbereich jedoch kaum geschehen. Deshalb erfolgt die Diskussion der BSC hier im Zusammenhang mit QM-Modellen.

---

<sup>66</sup> Bereits heute wird ein Teil der staatlichen Mittelzuweisung abhängig von der Anzahl der Studierenden (häufig derjenigen in der Regelstudienzeit) vergeben. Über den o.g. Zusammenhang von Ranking-Ergebnissen und Anzahl der Studieninteressierten wirken Ergebnisse von Rankings – insbesondere bei nicht zulassungsbeschränkten Studiengängen – also bereits auf die Finanzierung der Hochschulen. Mit dem Anteil derjenigen Studierenden, die sich an Rankings orientieren, dürfte sich diese Wirkung verstärken (vgl. hierzu auch Orr 2001: 161).

<sup>67</sup> In der Wirtschaft wurde von Unternehmen mit BSC diese bislang zu 70% zur Leistungs- und Erfolgsbeurteilung verwendet. Zwei Drittel von diesen verknüpfen dies direkt mit (materiellen) Anreizen. Dagegen wird in der Literatur das Potential und der innovative Charakter der BSC v.a. in der Strategiekommunikation gesehen (vgl. Speckbacher/ Bischof 2000: 808, Grüning 2002: 88).

### 3.3.1 DIN ISO 9000ff

Die *DIN ISO 9000ff.* (Normenreihe des Deutschen Instituts für Normung und der International Standard Organisation) können eingeordnet werden als ein Modell zur Zertifizierung der Qualitätsfähigkeit einer Institution (vgl. Pasternack 2004: 38). Die DIN ISO 9000 („Leitfaden zur Auswahl und Anwendung der Normen über Qualitätsnachweisführung“) erläutert v.a. Schlüsselbegriffe des Qualitätsmanagements. Die Normen 9001 bis 9003 beinhalten Modelle zur externen Darlegung des QM-Systems eines Unternehmens, d.h. zur Nachweisführung über die geplanten, realisierten, dokumentierten und kontrollierten QM-Maßnahmen, die Norm 9004 stellt einen Leitfaden zum Aufbau eines internen QM für alle Ebenen der Aufbauorganisation und alle Phasen der Ablauforganisation dar. Zur Normenfamilie werden außerdem die Teile 8402 (Definitionen – vgl. Abschnitt zur Klärung des Begriffes Qualität), 10011 (Qualitätsaudit)<sup>68</sup> sowie 10013 (QM-Handbuch) gezählt (vgl. Pasternack 2004: 37).

Die Stärken der Normenfamilie DIN ISO 9000ff. liegen nach Zink darin, daß sie v.a. dort, wo es noch kein System der Qualitätssicherung gibt, die internen Strukturen einer Organisation in den Blick nimmt, um technische, zeitliche und personelle Ablaufprozesse zu optimieren, dabei Standards zu erreichen hilft und dies nach außen kommunizierbar macht. Der Vorteil ihres Einsatzes im Hochschulbereich liegt nach Stawicki (vgl. HRK 1998a: 69ff.) darin, daß man „systematisch und kontrolliert aus einer Bestandsaufnahme zu Veränderungen gelangen“ könne, was im Lehrberichtsverfahren sowie in der internen und externen Evaluation nicht automatisch entstehe, da diese primär eine Bestandsaufnahme lieferten. In diesem Sinne sei Qualitätsmanagement kein weiterer, konkurrierender Ansatz, sondern ein umfassendes Konzept, in dem Evaluation u.ä. ihren Platz haben.

Kritik wird v.a. daran geäußert, daß zwar das korrekte Arbeiten nach den selbstgesetzten Qualitätsstandards bestätigt werde, dies aber genau genommen nichts über die Inhalte bzw. über die Qualität des „Produktes“ oder der „Dienstleistung“ selbst aussage und daß der Aufwand zur Dokumentation im Verhältnis zum erwartbaren Nutzen zu hoch sei (vgl. HRK 2003b: 14) bzw. „Bürokratisierungsprozesse“ entstehen können (vgl. Stockmann 2002b: 217). Die Gefahr dieses Ansatzes bestehe grundsätzlich darin, daß formale Gesichtspunkte Übergewicht vor inhaltlichen finden, schreibt auch Bartel-Lingg (1996: 35). Insbesondere der Nutzen für die Qualitätsverbesserung wird bezweifelt. So führt Zink (1999: 29) eine Studie an, wonach der Nutzen häufiger darin gesehen wird, die Zertifizierung als Marketinginstrument zu gebrauchen, als zur Qualitätsverbesserung zu nutzen. Dies geht jedoch am Anspruch des Modells vorbei. Im Jahre

2000 wurde die Normenreihe jedoch revidiert (und wird nun auch als ISO 9000: 2000 bezeichnet). In diesem Zusammenhang nimmt deren Prozeßorientierung zu und bis Ende 2003 mußten alle bereits zertifizierten Qualitätsmanagementsysteme auf die neue Normversion umgestellt werden (vgl. Wuppertaler Kreis/ Certqua 2002: 25ff.). Schwierigkeiten bereitet jedoch weiterhin der für die DIN ISO 9000ff. zentrale Begriff des „Kunden“ bei der Anwendung im Hochschulbereich (vgl. HRK 1998a: 74), da insbesondere Studierende nicht nur als „Abnehmer“, sondern auch als Mitwirkende oder auch „Koproduzent“ im Hochschulbereich gesehen werden sollten (vgl. auch Abschnitt zur Definition von Lehrqualität auf Lehrveranstaltungsebene).<sup>69</sup>

Und schließlich bestehe – so eine weitere Kritik – Qualität an der Hochschule nicht allein in der (möglichst) mustergetreuen Fertigung, sondern anders als in der Wirtschaft wesentlicher noch in der Normabweichung bzw. Erzeugung des Nichtnormierbaren (vgl. bereits Pasternack 2001: 47). Daher bleibe die Adaptionfähigkeit von QM nach DIN ISO 9000ff. auf standardisierbare Prozesse an den Hochschulen beschränkt. „Dort indes – etwa bei Verwaltungsabläufen, aber auch in Bereichen wie der Studienfachberatung – wird es sicher auf große Optimierungsreserven stoßen.“ (vgl. Pasternack 2004: 39).

Trotz aller Kritik an der Angemessenheit der DIN ISO 9000ff. haben einige Lehrstühle (z.B. Knoll (Universität Leipzig) ein Qualitätsmanagement nach dieser Norm entwickelt und sich zertifizieren lassen (vgl. HRK 2000b: 9ff.). Als Hauptertrag für den Lehrstuhl wurde von Knoll die feste Verankerung von regelmäßigen Arbeitsvorgängen im Zusammenhang mit der Einbeziehung von Studenten und Planungen genannt, wie z.B. die Ideenbörse zur Gestaltung des Lehrangebotes. Die wichtigste Auswirkung für die Hochschule sei die Weiterentwicklung des Lehrberichtverfahrens zu einem Regelkreis gewesen (vgl. auch HRK 2000d: 50). Ebenfalls eine Rolle gespielt haben dürfte – auch wenn dies nicht erwähnt wird – der „Marketingeffekt“ einer solchen Zertifizierung. So wies z.B. Knoll in verschiedenen Zusammenhängen mehrfach darauf hin, daß sein Lehrstuhl der erste in diesem Bereich sei, der nach DIN ISO 9000ff. zertifiziert wurde. Zwar ist die Zahl von Einrichtungen, die Erfahrungen mit diesem Modell haben, inzwischen leicht gestiegen, insgesamt haben QM-Verfahren nach DIN ISO 9000ff. aber nach wie vor einen vergleichsweise geringen Stellenwert an Hochschulen (vgl. HRK 2003b: 14). Allerdings haben wesentliche Grundgedanken, wie die Implementierung von Zielvereinbarungen

---

<sup>68</sup> Pasternack (2004: 85) erläutert Qualitätsaudit in seinen weiteren Ausführungen als separates Instrument der Qualitätsbewertung. Dies soll hier aufgrund der bereits weiter o.g. von Pasternacks Analyse abweichenden Zielstellungen nicht erfolgen.

<sup>69</sup> Buschor (2002: 72) berichtet, daß aus den genannten Gründen in neuerer Zeit in Deutschland der Begriff des Kunden eher vermieden wird.

im Sinne eines Regelkreises mittlerweile Eingang in die Evaluationsverfahren gefunden (vgl. Stockmann 2002b: 209ff.).

### **3.3.2 Total Quality Management**

Das *Total Quality Management (TQM)* wird als Modell *umfassenden* Qualitätsmanagements gesehen, das noch „einen Schritt weiter“ geht als Zertifizierungen (vgl. Zink 1999:27). Definiert ist es als „auf der Mitwirkung aller ihrer Mitglieder beruhende Führungsmethode einer Organisation, die Qualität in den Mittelpunkt stellt und durch Zufriedenstellung der Kunden auf langfristigen Geschäftserfolg sowie auf Nutzen für die Mitglieder der Organisation und die Gesellschaft zielt“ (vgl. DIN ISO 8402).

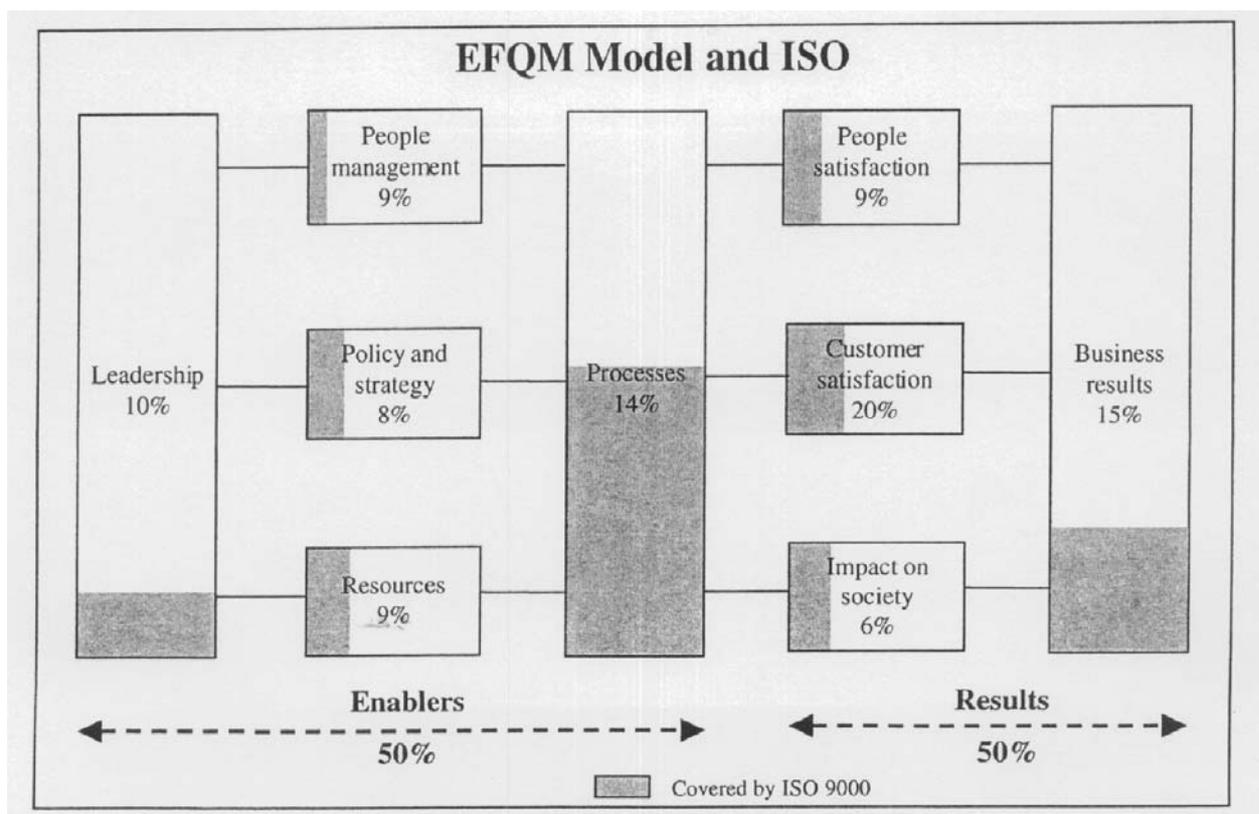
Im Unterschied zur Zertifizierung geht es beim TQM stärker um eine Steigerung der Leistungsfähigkeit und weniger um die Einführung von Standards. Durch Selbstbewertungen (und nicht wie bei der Zertifizierung durch Auditierung) soll anhand eines transparenten Modells ein kontinuierlicher Verbesserungsprozeß (KVP) in Gang gesetzt werden. Anknüpfungspunkte für die Hochschulen werden insbesondere darin gesehen, daß die beim TQM wichtige Mitarbeiterorientierung „dazu beitragen könnte, Mitarbeiterinnen eine positive Arbeitsmotivation zu verschaffen und damit etwa auch das Arbeitsklima zu verbessern: insofern würde manche Hochschulleitung (erst) dadurch sensibilisiert etwa für die demotivierenden Effekte von Hierarchien, erzwungener Unselbständigkeit und befristeten Arbeitverträgen auf halben bzw. Drittelstellen“ (vgl. Pasternack 2004: 40).

In Ländern wie Großbritannien, die TQM bereits länger im Hochschulbereich anwenden, wird dieses als durch fünf Hauptelemente charakterisiert beschrieben (vgl. Sims & Sims 1995: 9ff., eigene Übersetzung):

- 1.) ein Fokus auf einem gekennzeichneten Prozeß (oder System) der mit den Zielen der Institution verbunden werden kann
- 2.) ein Fokus auf die Identifikation, das Verstehen und das Erreichen der Bedürfnisse eines jeden individuellen Nutzers
- 3.) ein Vertrauen auf solide Daten (und wissenschaftliche Analyse), um die Notwendigkeiten der Institution zu definieren, Probleme entdecken zu helfen und Lösungen zu finden
- 4.) die Miteinbeziehung des Personals in allen Funktionen, Situationen und Ebenen in der Beschlußfassung über Verbesserungen
- 5.) die Achtung der Meinungen aller beteiligten Menschen innerhalb des Prozesses unabhängig davon, ob sie Nutzer, Mitarbeiter oder Leitungsmitglieder sind

Sims & Sims (1995: 12ff.) nennen noch eine Reihe weiterer Aspekte, die für die effektive Anwendung von TQM an Hochschulen nützlich seien, sowie Fehler, die es zu vermeiden gelte. Deren Aufzählung würde jedoch den Rahmen dieser einordnenden Ausführungen sprengen.

Ein spezielles Modell des TQM und zugleich das, mit dem derzeit in Europa die TQM-Inhalte bearbeitet werden, stellt das Modell für Business Excellence der European Foundation for Quality Management<sup>70</sup> (EFQM) dar. EFQM greift in gewisser Weise die Kritik an der DIN ISO 9000ff. auf, in dem es sich (nur) zu 50% auf Elemente bezieht, die die *Befähigung* zur Qualitätserzielung erfassen (Mittel und Wege) und zu weiteren 50% auf die Prozeßergebnisse. Zink (1999: 33) spricht als wesentlichen Aspekt auch an, daß mit den Ergebnisgrößen die Dinge eher vergangenheitsorientiert betrachtet werden, während mit den Befähigergrößen die Potentiale angeschaut werden, die für eine Zukunftssicherung vorhanden sind. Die beiden Bereiche werden jeweils weiter untergliedert in neun verschiedene Kriterien und diese mit Gewichtungsfaktoren versehen (für die Gesamtbewertung).

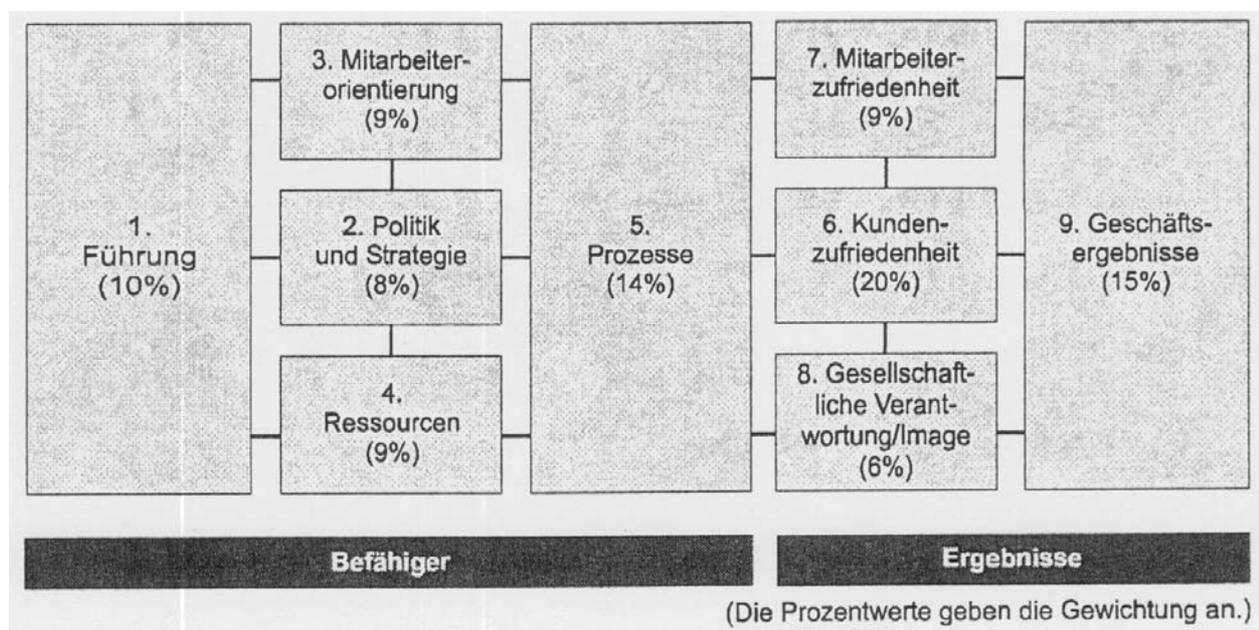


Vergleich von EFQM und DIN ISO 9000ff. nach Zink (vgl. HRK 1999:35)

Die neun Kriterien sind weiter untergliedert in insgesamt 32 Einzelaspekte. Für diese sind Bewertungen auf einer fünfstufigen Skala vorzunehmen, inwieweit das jeweilige Kriterium

erfüllt ist. Insgesamt ist es das Ziel des EFQM, eine Organisation abzubilden, die auch über die nächste Zukunft hinaus erfolgreich sein will und kann (vgl. Zink 1999: 32). Dies schließt ggf. auch den Vergleich mit anderen Organisationen als einen Teil der Bewertung und damit die Idee des Benchmarkings<sup>71</sup> ein, so Zink weiter. Organisationen, die nach EFQM arbeiten, können sich außerdem für den sogenannten „European Quality Award“ bewerben. In Japan,<sup>72</sup> später in den USA und auch im übrigen Europa ist TQM häufig mit solchen Qualitätspreisen verbunden (vgl. Stockmann 2002b: 214ff.). In Deutschland ist dies jedoch bisher eher selten der Fall.<sup>73</sup>

Zink (1999: 144ff.) hat einen Vorschlag erarbeitet, wie das EFQM speziell für die Qualität der Hochschullehre angepasst werden kann.



Vorschlag zur Anpassung des EFQM für die Hochschullehre nach Zink (1999: 144)

Das TQM und als seine spezielle Ausformung das EFQM werden als relativ offene Modelle eingeschätzt, die auf „jede“ (vgl. Zink 1999: 34) oder zumindest viele Arten von Organisationen

<sup>70</sup> Die European Foundation for Quality Management wurde Ende der 1980er Jahre gegründet, um Europa in der internationalen Qualitätsdiskussion gegenüber den USA und Asien voranzubringen (vgl. Zink 1999: 32).

<sup>71</sup> Von der HRK (2003b: 15) wie auch von Pasternack (2004: 87) werden Benchmarkings als mehr oder weniger eigenständige Instrumente bzw. Verfahren zur Qualitätsbewertung genannt. Sie werden jedoch nur relativ selten explizit von den Hochschulen als genutztes Verfahren angegeben. Deren Grundgedanke – Nutzung vorhandener Problemlösungen (best practice) – sei allerdings häufig als Element im Kontext etablierter Evaluationsverfahren anzutreffen (vgl. HRK 2003b: 15).

<sup>72</sup> Von Japan wird zugleich die stärkste Mitarbeiterorientierung des TQM-Ansatzes (im Vergleich zu europäischen und amerikanischen Ansätzen) berichtet (vgl. Bartel-Lingg 1996: 37). Mehrere Autoren vermuten, daß gerade diese stärkere Mitarbeiterorientierung – und nicht etwa kulturelle oder mentale Unterschiede – der ausschlaggebende Erfolgsfaktor des japanischen Qualitätsmanagements war.

<sup>73</sup> Pasternack (2004: 88) nennt Qualitätspreise als eigenes Instrument der Qualitätsbewertung, wobei er einschätzt, daß „in Deutschland vor allem eine größere Wirkung entfaltet: Das Total-E-Quality-Prädikat“. In dieser Arbeit werden Preise im Abschnitt zu Leistungsanreizmechanismen diskutiert.

anwendbar sind. So kämen im Gegensatz zur DIN ISO 9000 ff.<sup>74</sup> sehr wohl Elemente wie Politik und Strategie, Auswirkungen auf die Gesellschaft vor und kontinuierliche Verbesserung und Innovation sind wesentliche Bewertungskriterien (vgl. auch HRK 2001a: 31). In dieser Offenheit liegen, die Vorteile von TQM bzw. EFQM – neben der bereits erwähnten stärkeren Orientierung auf die Qualität des betrachteten Gegenstandes (und nicht nur seiner ihn herstellenden Institution) in Bezug auf die Kundenbedürfnisse und neben seiner stärkeren Mitarbeiterorientierung. In einem Handbuch für Personalräte zum Thema „Flexibilisierung der Hochschulhaushalte“ wird sogar explizit darauf hingewiesen, daß bei konsequenter Umsetzung des TQM-Ansatzes „korrespondierende positive Effekte der Leistungsverdichtung erzeugt werden, welche die Leistungsverdichtung selbst aus der Perspektive der MitarbeiterInnen tragbar werden lassen. Ein solcher Effekt dürfte v.a. erhöhte Arbeitszufriedenheit sein.“ (vgl. Pasternack 2001: 19).

Als Problem des TQM wird die Tendenz zur Verselbständigung der Qualitätssicherung gesehen. Als Beispiel hierfür wird die „Verwechslung einer Hochschule mit einer Konservenfabrik durch die Anhänger des New Public Management“ erwähnt (vgl. HRK 2001a: 32). Außerdem wäre aber auch hier der nicht ohne weiteres übertragbare Kundenbegriff (vgl. HRK 1999a: 41ff.) zu nennen. Speziell am EFQM wird kritisch bewertet, daß die vorgegebenen Parameter und Gewichtungsfaktoren nicht theoretisch hergeleitet oder begründet seien<sup>75</sup> und der zeitaufwändige Implementierungsprozeß bei der Selbstbewertung beanstandet (vgl. Stockmann 2002b: 218). Diese Kritik entschärft sich allerdings etwas, da es durchaus Praxisbeispiele gibt, die mit einem für Gutachter in Peer-review-Verfahren vergleichbaren Zeitaufwand auskommen (vgl. Olbertz/Otto 2001: 13ff.). Letztlich muß der Aufwand ohnehin v.a. im Verhältnis zu seinem Nutzen gesehen werden. Nach einer Umfrage der HRK (2003b: 27) zeigen Universitäten gegenüber dem TQM- bzw. EFQM-Verfahren eine größere Aufgeschlossenheit als Fachhochschulen. Insgesamt wird auch dieses QM-Verfahren im Vergleich zu bereits etablierten Evaluationsverfahren bisher nur von relativ wenigen Einrichtungen genutzt. Gleichwohl überwiegen bei Universitäten wie auch bei Fachhochschulen die positiven Bewertungen des Stellenwertes der QM-Verfahren für die Qualitätsverbesserung.

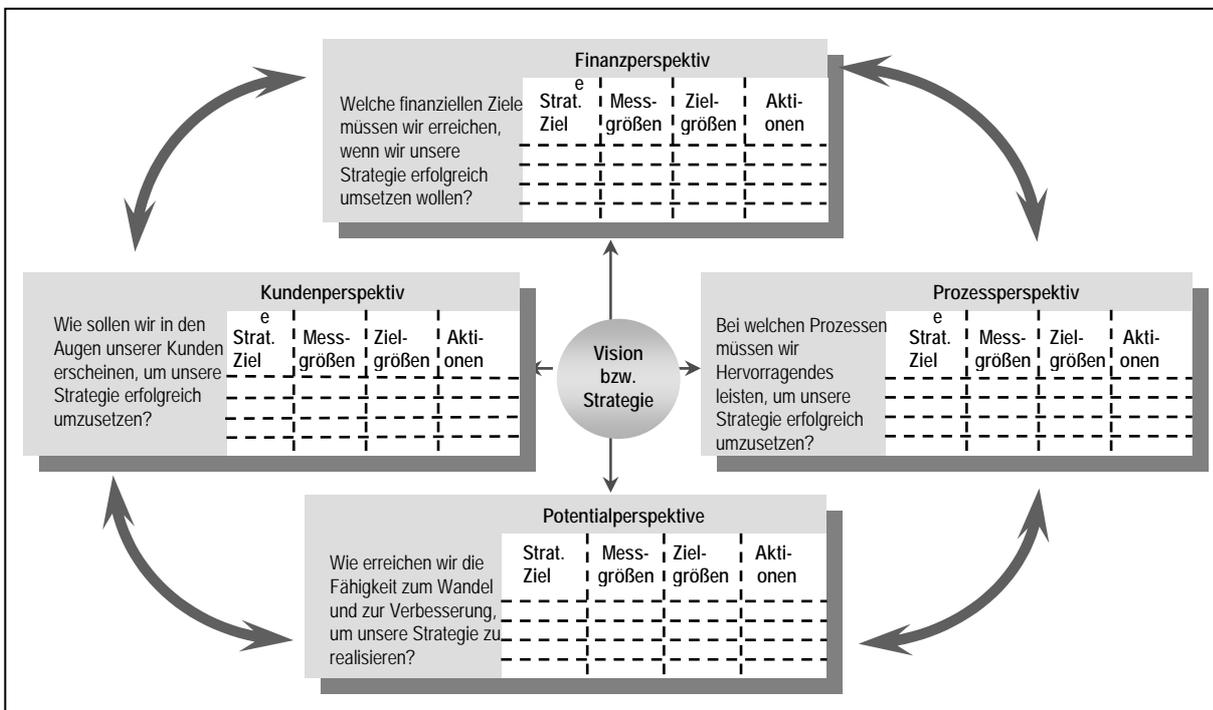
---

<sup>74</sup> Mehrere Autoren weisen darauf hin, daß DIN ISO 9000ff. und TQM nicht als konkurrierende Ansätze oder als sich ausschließende Alternativen betrachtet werden müssen, da die Anforderungen in wesentlichen Punkten deckungsgleich seien. TQM könne auch als Weiterführung von DIN ISO 9000ff. betrachtet werden, wobei die Modelle lediglich verschiedene Schwerpunkte setzen, aber ineinander überführbar sind (vgl. Zink 1999: 27ff., Stockmann 2002b: 217).

<sup>75</sup> Diese Kritik wurde z.T. aufgenommen, indem die Parameter nicht als absolut fix angesehen, sondern im Sinne des Modells angepasst und weiterentwickelt wurden (vgl. Zink 1999: 147ff.), was bei Stockmann noch nicht in die Diskussion einfloß. Auch die Kritik an den Gewichtungsfaktoren könnte ggf. aufgenommen werden, indem diese

### 3.3.3 Balanced Scorecard

Die *Balanced Scorecard (BSC)* ist wie DIN ISO 9000 und TQM ein Begriff aus der Betriebswirtschaft, der ursprünglich von Kaplan/ Norton (1997) als strategisches Managementkonzept für die Unternehmensführung entwickelt wurde, inzwischen aber auch in die Leistungs- und Qualitätsdebatte an den Hochschulen Eingang fand. „Balanced“ meint hier eine ausgewogene Berücksichtigung der wesentlichen Perspektiven, die einzubeziehen zur Umsetzung einer Organisationsstrategie notwendig ist. „Scorecard“ beschreibt ein übersichtliches Erfassungs- und Darstellungsschema der einzelnen strategischen Ziele, der dazugehörigen Meßgrößen, Zielwerte und Aktivitäten (vgl. Pasternack 2004: 64ff.). Dem entsprechend besteht die BSC konzeptionell aus zwei Elementen: a) Verknüpfung von 4 Betrachtungsperspektiven und b) klare und systematische Ableitung von Zielen und Maßstäben einschließlich der Integration eines transparenten Zielsystems und Leistungsindikatorensystems. Die vier Perspektiven sind die finanzwirtschaftliche Perspektive, die Kundenperspektive, die Prozeßperspektive und die Entwicklungs- oder auch Potentialperspektive, die jeweils zu operationalisieren sind (vgl. Horvath u.a. 2000: 10ff.).



Grundmodell der Balanced Scorecard nach Horvath u.a. (2000: 11)

nicht als fix angesehen, sondern als gemäß den zu verfolgenden Zielen auszubalancieren angesehen würden. Dieser Grundgedanke soll im nachfolgenden Abschnitt mit der Balanced Scorecard aufgegriffen werden.

Als Vorzug der BSC schätzt Pasternack (2004: 64) ein, daß sie wie auch das EFQM-Modell auch Auskunft über vorhandene Potentiale gibt und damit versucht, auf Probleme zu reagieren, welche die meisten Steuerungssysteme haben. So benennt er die vergangenheitsorientierte Ausrichtung auf Zahlen, wie sie das Rechnungswesen kennzeichnet, als problematisch. Auch Bieri (vgl. HRK 2001: 27) von der ETH Zürich nennt die „einseitig zahlenmäßige Fixierung“ von Kennzahlensystemen als Schwierigkeit, die durch die BSC überwunden werden kann.<sup>76</sup> Speckbacher/ Bischof (2000: 797) bezeichnen die Erkenntnis, daß eine Steuerung nicht allein anhand finanzieller Kennzahlen erfolgen kann, allerdings als keineswegs neu. Zwar sehen auch sie Vorteile in der Nutzung nichtfinanzieller Kennzahlen (v. a. in der Möglichkeit der besseren Früherkennung von Fehlentwicklungen und deren besserer Lokalisierbarkeit), der innovative Gehalt der BSC liege jedoch nicht darin, sondern in der Frage, welcher Art zusätzliche, nicht finanzielle Kennzahlen sein sollten.

Dies führt zu einer weiteren Besonderheit der BSC, die häufig als Vorzug genannt wird: Sie berücksichtigt neben den traditionellen quantitativen insbesondere auch qualitative Größen (vgl. Horvath u.a. 2000: 9, HRK 2001: 27). Neben objektiven Kennzahlen wird auch ermöglicht, subjektive Faktoren, wie z.B. Kundenzufriedenheit oder auch Mitarbeiterzufriedenheit<sup>77</sup>, über Meßgrößen zu erfassen (vgl. Ederer 1999: 2). Seeger u.a. (1998: 208) werten dies für den Bereich Controlling wie folgt: „Einer Ergänzung des Controlling durch die Messung nicht finanzieller, sondern wirkungsorientierter Ziele steht damit nichts mehr im Wege.“ Mittels BSC wäre es also möglich, sehr viel stärker den Aspekt der Qualität einzubeziehen als bisher<sup>78</sup> (vgl. auch Abschnitte zum Qualitätsbegriff in dieser Arbeit). Hierin kann die eigentliche Stärke der BSC für die Förderung der Qualitätsorientierung gesehen werden.

Des weiteren gilt die BSC – stärker als z.B. EFQM – als flexibel einsetzbar, da sie in einem kontinuierlichen Kommunikationsprozeß zu erarbeiten und anzupassen bzw. auf die jeweiligen

---

<sup>76</sup> Diese Schwierigkeit entsteht dadurch, daß nichtfinanzielle Kennzahlen schwerer zu erfassen sind. Außerdem ist der Nutzen von Investitionen in immaterielle Vermögenswerte (wie Mitarbeiterqualifikation, Qualität von Leistungen und Kundenzufriedenheit) im Gegensatz zu den hierfür notwendigen Ausgaben in der Regel nicht klar auf einzelne Bereiche zuordenbar sind. Erzeugt beispielsweise ein Bereich durch Forschungs- und Entwicklungsaktivitäten Know-How, so kann (und soll) nicht verhindert werden, daß auch andere Bereiche davon profitieren. Ähnliche Effekte entstehen in hohem Maße auch bei Weiterbildungsprogrammen. Bei der bereichsbezogenen Gegenüberstellung von Kosten und finanziellem Nutzen werden daher zwar die gesamten Kosten, nicht aber der gesamte Nutzen berücksichtigt. Werden Bereiche anhand rein finanzieller Erfolgsgrößen beurteilt und gesteuert, bestehe die Tendenz zur Unterinvestition in immaterielle Vermögenswerte (vgl. Speckbacher/ Bischof 2000: 797).

<sup>77</sup> In der Praxis hat die stärkere Berücksichtigung der Mitarbeiterperspektive jedoch häufig (noch) einen untergeordneten Stellenwert (vgl. Speckbacher/ Bischof 2000: 806). Andererseits nannten von Unternehmen, auf die Frage, welche Voraussetzungen sie bei der Umsetzung eines Performance Measurement-Systems (als das die BSC dort eingeordnet wurde – Anm. d. Verf.) als erfolgskritisch ansehen, mit 77 von 108 die große Mehrheit die Mitarbeiterbeteiligung und Kommunikation an erster Stelle (vgl. Grüning 2002: 206).

Bedürfnisse hin auszurichten ist (vgl. Horvath 1999: 6). Zurwehme (2000: 36) führt hierzu aus, daß der Grundaufbau der BSC als „flexibles Raster aufzufassen ist, daß zahlreiche Spielräume zur individuellen Gestaltung unter Berücksichtigung der Institutionsspezifika zulässt“.<sup>79</sup> Dies waren Gründe, warum die BSC zunehmend im Bereich der öffentlichen Verwaltung Anwendung fand (wobei in diesem Sektor oftmals eine Erweiterung des BSC-Grundmodells um eine Politikperspektive vorgenommen wurde) und bereits vor einigen Jahren auf ihre Eignung auch für den Bildungssektor hin untersucht wurde (vgl. Zurwehme 2000: 38ff.). Exemplarisch für den Einsatz an deutschen Fachhochschulen hat Lerchenmüller (2002: 6ff.) eine BSC entwickelt und an spezielle Erfordernisse angepasst. Längere Erfahrungen mit dem Einsatz einer BSC liegen an der Universität Reykjavik vor, die in Anlehnung an Pasternack (2004: 68) wie folgt zusammengefasst werden können: Dort erwies sich als hinderlich, daß keine klaren Organisationsziele bzw. Strategien existierten, die die BSC jedoch voraussetzt. Als förderlich erwies sich die bereits recht stark ausgeprägte Dokumentationskultur mit dem Vorhandensein z.B. von Forschungsdatenbanken und das hohe Ausbildungsniveau des Personals, welches erlaubte, die BSC in einer sehr dezentralen Weise umzusetzen. Als wesentlicher funktionaler Vorzug der BSC wird außerdem beschrieben, daß Entwicklungsszenarien auf ihre Erfolgsaussichten hin überprüft werden können.

Im Vergleich zu DIN ISO 9000ff. und TQM/ EFQM ist zur BSC zusammenfassend einzuschätzen: Ebenso wie EFQM enthält die BSC auch eine Potentialperspektive. Sie bezieht jedoch (mit der Kundenperspektive) noch stärker explizit die „Kundenbedürfnisse“ ein. Über die Potentialperspektive kann auch die Mitarbeiterorientierung stärker einbezogen werden.<sup>80</sup> Außerdem wird die Möglichkeit zur Einbeziehung qualitativer Größen (i.S. Qualität erfassender Größen)<sup>81</sup> bei der BSC im Vergleich zum EFQM als stärker und diese als flexibler einsetzbar eingeschätzt. Nachteilig wirkt sich aus, daß die BSC in der Praxis als relativ komplex erscheint (vgl. Zurwehme 2000: 82, Mende/ Stier 2002: 96) Zudem liegen bisher erst relativ wenige Erfahrungen mit dem Einsatz der BSC im Bildungs- bzw. Hochschulbereich vor. Und schließlich setzt die BSC im Unterschied zur DIN ISO 9000ff. voraus, daß bereits eine Art Leitbild bzw.

---

<sup>78</sup> Laut Speckbacher/ Bischof (2000: 798) gingen auch Kaplan und Norton offensichtlich davon aus, daß es sinnvoll sei, nicht nur finanzielle Performance, sondern auch das Erreichen nichtfinanzieller Ziele (etwa Qualitätsziele) zu belohnen. Dies sei auch in vielen Teilen der Unternehmenspraxis üblich.

<sup>79</sup> Lichka u.a. (2002: 915) gehen davon aus, daß die BSC neben der finanziellen lediglich „mindestens drei weitere Perspektiven“ umfasst. „Weitere Perspektiven sind je nach Bedarf möglich.“

<sup>80</sup> Die Mitarbeiter gelten bekanntlich als das wichtigste Potential, insbesondere im wissensproduzierenden Bereich. Aus diesem Grund schlägt Zurwehme (2000: 54) vor, die Potentialperspektive in Bildungseinrichtungen eher als Mitarbeiterperspektive zu bezeichnen.

<sup>81</sup> Auch in gewerkschaftsnahen Publikationen wird betont, daß es gilt, „die Balance zwischen Modernisierungszielen wie Steigerung von Qualität, Quantität, Wirtschaftlichkeit und Mitarbeiterorientierung zu halten“ (vgl. Tondorf/ Jochmann-Döll 2004: 433). Dies könnte die BSC vom Ansatz her am ehesten leisten.

Strategie vorhanden ist, was sich ebenfalls – jedenfalls beim derzeitigen Entwicklungsstand an den meisten Hochschulen – als problematisch erweisen dürfte.

### 3.4 Leistungsanreizmodelle

Nachdem in den beiden vorhergehenden Abschnitten 3.2 und 3.3 ausgewählte Konzepte bzw. Instrumente der Leistungs- und Qualitätsbewertung sowie Modelle des Qualitätsmanagements diskutiert wurden, sollen nun Leistungsanreizmodelle diskutiert werden. Durch die Politik wurden diese in den letzten Jahren an den Hochschulen stärker als bisher in den Fokus der Aufmerksamkeit gerückt. Innerhalb der Qualitätsdebatte existieren aber bereits seit längerer Zeit Diskussionen in dieser Richtung.

Bereits 1995 wurde in der Dokumentation des 1. Bildungspolitischen Forums St. Marienthal „Qualitätsmanagement im Bildungswesen“ darauf hingewiesen, daß sich bei aller Unterschiedlichkeit der Qualitätsdiskussion in den verschiedenen Institutionen des Bildungswesens auch Gemeinsamkeiten finden. Eine dieser Gemeinsamkeiten ist die Frage: „In welcher Weise können die Ergebnisse von Qualitätsprüfungen mit Handlungskonsequenzen (unter Einschluß von Anreizen oder Sanktionen) so rückgekoppelt werden, daß sich daraus auch eine tatsächliche Qualitätsverbesserung ergibt?“ (vgl. Gnahs u.a. 1995: 7). Etwa 10 Jahre später kommt Pasternack (2004: 56) in seiner Bestandsaufnahme zur Qualitätsorientierung an Hochschulen zu dem Schluß: „Leistungsanreizmodellen können die Hochschulen künftig nicht mehr ausweichen“ und ordnet diese als Instrument der Qualitätsentwicklung ein.

Auch andere Autoren forderten (meist im Zusammenhang mit bestimmten Problemen im Hochschulbereich, die als verbesserungsbedürftig eingeschätzt werden), dies durch Anreize zu stimulieren (vgl. z.B. Klockner 1997: 79, Meyer 1999: 636, Lemke 2000: 172). Häufig wird bei der Forderung nach Anreizen nicht genauer ausgeführt, welche Art von Anreizen gemeint ist. Grundsätzlich kann eine Unterscheidung in materielle und nichtmaterielle Anreize getroffen werden. Teilweise ist auch von monetären und nichtmonetären Anreizsystemen die Rede (vgl. Ziegele/ Handel 2004: 6). Bei letzterer bleibt aber unklar, wo z.B. die Ausstattung eines Lehrstuhles (z.B. mit vorhandenen Möbeln, PC usw.) zuzuordnen ist. Zu nichtmateriellen Anreizsystemen zählen Ziegele/ Handel (2004) neben der Reputation auch Freiheit (im Sinne

---

<sup>83</sup> Auch in der Wirtschaft gibt es Beispiele, in denen Unternehmen es vorziehen, Qualitätsarbeit mit persönlicher Anerkennung statt mit Geld zu honorieren; „denn persönliche Anerkennung hat sich als äußerst wirksames Motivationsinstrument erwiesen“ (vgl. Hochschild 1998: 33).

von Autonomie), Zeitallokation (z.B. Forschungsfreisemester) und Transparenz (für alle Beteiligten, um kostenbewußtes Handeln auszulösen).

Darüber hinaus können materielle Anreize (nach der Ebene, auf der sie wirken sollen) unterschieden werden in das Einkommen bzw. leistungsorientierte Besoldungszulagen (individuelle Ebene) und in leistungsorientierte Verteilung finanzieller Mittel bzw. Ressourcen an Hochschulen, Fakultäten/ Fachbereiche/ Institute usw. (institutionelle Ebene). Ziegele/ Handel (2004: 6) sprechen darüber hinaus von neben solchen auch als „simulierte Märkte“ bezeichneten Modellen auch von „echten“ Märkten als Leistungsanreizsystem und nennen hierzu für den Bereich der Lehre Studiengebühren als Beispiel. Deren (potentieller) Effekt als Anreizsystem ist jedoch von verschiedenen Bedingungen abhängig (z.B. Verbleib der Einnahmen), die noch ungeklärt sind. Zudem waren Studiengebühren für Erststudien bis Anfang 2005 gesetzlich untersagt. Daher werden in dieser Arbeit nur die ersten beiden materiellen Leistungsanreize vorgestellt und (soweit anhand der verfügbaren Literatur möglich) einer Einordnung und ersten Bewertung unterzogen.

### **3.4.1 Nichtmaterielle Anreize**

*Nichtmaterielle Anreize* wirken über die Anerkennung, z.B. Auszeichnungen bzw. Preise. So hat der Deutsche Hochschulverband (2000) formuliert: „Stärker als durch finanzielle Honorierung sind Hochschullehrer durch immaterielle Anreize zu motivieren.“<sup>83</sup> Daher hat er gefordert, besondere Leistungen durch die Gewährung zusätzlicher Forschungsfreisemester oder durch eine zeitweise Absenkung des Lehrdeputats zu honorieren. Als nichtmonetäre Anreize werden auch von Witte u.a. (2002: 16) Lehrdeputatsreduktionen und sogenannte „Sabbaticals“ gefordert, um zu vermeiden, „daß monetäre Anreize die vorhandene intrinsische Motivation ersetzen oder gar zerstören“. Bei Anreizen im Zusammenhang mit dem Bereich der Hochschullehre ist jedoch zu hinterfragen, ob Lehrdeputatsreduktionen und wie auch immer genannte Freistellungen vom Lehrbetrieb hier ein richtiges Signal setzen können, um wie angestrebt den Stellenwert der Lehre gegenüber der Forschung zu stärken. Von Witte u.a. (2002: 16) wird außerdem die Ausstattung als nichtmonetärer Anreiz genannt. Da eine Verbesserung der Ausstattung aber materielle Aufwendungen bedeutet und zudem die laufenden Zuweisungen für die Erhaltung und ggf. Verbesserung der Ausstattung inzwischen häufiger auch im Zusammenhang mit Zielvereinbarungen diskutiert werden, wird diese hier als materieller Anreiz aufgefasst.

Außerdem wird als wesentlicher Anreiz das Renommee innerhalb der scientific community gesehen, was durch die öffentliche Anerkennung wissenschaftlicher Leistung gefördert werden könne. Der Hochschulverband (2000) schlug daher die Einrichtung eines jährlich zu

vergebenden und hochdotierten Deutschen Wissenschaftspreis (National Scientific Award) vor, der von Bund und Ländern zu finanzieren und direkt vom Bundespräsidenten in den einzelnen Fächern zu verleihen sei. Weiterhin sollten Sonderpreise für exzellente fächerübergreifende Forschung, herausragende Lehre, Förderung des wissenschaftlichen Nachwuchses oder verständliche Vermittlung von Ergebnissen der Wissenschaft in der Öffentlichkeit geschaffen werden. Dies wurde jedoch bislang nicht umgesetzt. Pasternack (2004: 88) nennt nur einen Preis, der bisher in Deutschland größere Wirkung entfalten würde: das „Total-E-Quality-Prädikat“ als Beispiel für einen Qualitätspreis im Zusammenhang mit TQM.

Überlegungen zum Thema Preise für herausragende Lehre gibt es durchaus schon seit längerer Zeit (vgl. bmb+f 1991: 12ff.). An mehreren deutschen Hochschulen wurden auch bereits solche Lehrpreise vergeben (z.B. Universität Münster, TU Dresden, TU Bergakademie Freiberg).<sup>84</sup> Deren Vergabe erfolgte jedoch über Kommissionen und durch Gremienbeschlüsse und orientierte sich nicht an Bewertungen wie beim TQM üblich. Bei Preisvergaben im Zusammenhang mit TQM wird darauf hingewiesen, daß die Vergabe der Preise äußerst transparent nach nachvollziehbaren Kriterien erfolgen sollte, um anerkannt zu werden. Die Vergabe durch Verhandlung in universitären Gremien ist als relativ wenig transparent einzuschätzen. Dies mag einer der Gründe sein, warum solche Preise in Deutschland bisher keine größere Rolle spielen. Es wird aber auch eingewandt, daß Lehrpreise nur die bereits besten Dozenten erreichen würden und sich daher als Anreiz für schwächere Dozenten nicht eignen würden. Ein Anreiz zur Verbesserung der Lehre wird daher eher in studentischen Lehrveranstaltungsbewertungen in Verbindung mit Beratung/ Fortbildungsmaßnahmen gesehen (vgl. HRK 2004a: 96).

Ein weiterer Aspekt ist die bisherige Dotierung von Lehrpreisen: Zwar geht es bei solchen Preisen v.a. darum, die symbolische Anerkennung zu befördern. Dennoch ist vermutlich eine nennenswerte Dotierung von Lehrpreisen in ähnlicher Höhe wie bei renommierten Forschungspreisen notwendig, um mit solchen Preisen tatsächlich eine gewisse Anerkennung zu erzielen (und das grundsätzlich damit angestrebte Ziel der Erhöhung des Stellenwertes der Lehre gegenüber der Forschung zu erreichen).<sup>85</sup>

Von den nichtmateriellen Anreizen lassen sich die materiellen Anreize unterscheiden. Materielle Anreize sind v.a. leistungsabhängige Besoldungszulagen (vgl. Arnhold/ Schreiterer 2002: 3ff.) sowie die leistungsorientierte Mittelverteilung oder, wie Pasternack (2004: 56) dies formuliert, „leistungsabhängig differenzierter Ressourcenzugriff“. Diese werden nachfolgend vorgestellt.

---

<sup>84</sup> Solche Lehrpreise gibt es seit längerem auch an britischen Hochschulen (vgl. z.B. Orr 2001: 289).

<sup>85</sup> Die „New Australian Awards for University Teaching“ sind mit \$ 50.000 (Prime Ministers Award „Teacher of the year“), \$ 25.000 (40 Preise, entspricht etwa Anzahl der Universitäten) bis \$ 10.000 (210 Preise, entspricht etwa Anzahl der Fakultäten) dotiert (vgl. [www.backingaustraliasfuture.gov.au/archive/8.htm](http://www.backingaustraliasfuture.gov.au/archive/8.htm)).

Auch hierbei gilt es – wie bereits zuvor erwähnt – ein grundsätzliches Problem zu vermeiden, „daß monetäre Anreize die vorhandene intrinsische Motivation ersetzen oder gar zerstören“ (vgl. Witte u.a. 2002: 16). Hierauf weist auch Grüning (2002: 185, 192) im Zusammenhang mit Performance Measurement Systemen hin. Am pointierteren hat die Kernthese (allerdings für den Bereich der privaten Wirtschaft) hierzu aber wohl Sprenger (2000: 8) in seinem zwar keinen wissenschaftlichen Anspruch erhebenden, aber häufig zitierten Buch „Mythos Motivation“ formuliert: „Je mehr versucht wird, Motivation zu kaufen, desto schwächer die Führung, desto bedrohter das Unternehmen.“ Jedoch räumt an anderer Stelle selbst Sprenger (2000: 118) ein, daß unter bestimmten Bedingungen (z.B. von gegenseitiger Achtung geprägter Führung) nichts gegen eine längerfristig gedachte leistungsorientierte Einkommensdifferenzierung auf der Basis klarer Leistungsabsprachen einzuwenden sei. Werden zudem subjektive Beurteilungen mit eingesetzt, so haben diese den Vorteil, daß sie im Gegensatz zu sogenannten quantitativen oder objektiven Produktivitätskennzahlen nicht beliebig gesteigert werden können und somit nicht zu einer schleichenden ständigen Normerhöhung beitragen könnten (vgl. Sprenger 2000: 118ff.). Verbunden mit dem Argument der Leistungsgerechtigkeit und unter Nennung weiterer Randbedingungen sehen auch gewerkschaftsnahe Vertreter der Arbeitsforschung bei aller häufig geäußerten Kritik durchaus Vorteile in materiellen Anreizmodellen. Darauf wird insbesondere im Zusammenhang mit den nachfolgend vorzustellenden leistungsorientierten Besoldungen noch genauer einzugehen sein. Ziegele/ Handel (2004: 20) weisen jedoch darauf hin, daß diese im Hochschulbereich immer mit nichtmateriellen Anreizen ergänzt werden sollten, um nicht-intendierte Wirkungen zu vermeiden oder wenigstens zu vermindern.

### **3.4.2 Leistungsabhängige Besoldungszulagen**

*Leistungsabhängige Besoldungszulagen* sind eine Form leistungsorientierter Entgeltbestandteile und seit Februar 2002 mit dem in Kraft getreten sogenannten Professorenbesoldungsreformgesetz in der Bundesrepublik beschlossen. Sie müssen im Einzelnen noch von den Ländern umgesetzt werden. Da zur konkreten Ausgestaltung noch keine endgültigen Entscheidungen getroffen wurden, beschränken sich die nachfolgenden Ausführungen auf eine grundsätzliche Diskussion der zugrunde liegenden Ansätze. Bevor jedoch versucht wird, die Ansätze hierzu zu systematisieren und anhand der bisher vorliegenden Literatur zu einer ersten einordnenden Bewertung zu gelangen, soll zunächst kurz deren jüngere historische Entwicklung nachgezeichnet werden:

Bereits einige Jahre vor dem Gesetz gab es eine Art „white paper“ hierzu mit Vorschlägen einer Bund-Länder-Arbeitsgruppe (1998). In diesem Bericht wurden bereits Eckpunkte eines

künftigen leistungsorientierten Besoldungssystem festgehalten. Danach sollten Hochschullehrer der Besoldungsgruppen C 2 bis C 4 bis zu 3600 DM monatlich als Leistungszulage erhalten können. Im Gegenzug sollten die bisher mit dem Alter ansteigenden Besoldungszulagen bis auf eine Grundvergütung wegfallen. Hierzu wurden insgesamt fünf Modelle diskutiert, von denen drei in die engere Wahl kamen. Als Leistungskriterien wurden insbesondere Leistungen in den Bereichen Lehre, Forschung, Hochschulmanagement und Hochschulsebstverwaltung genannt. „Dabei können sowohl quantitative Leistungsparameter, als auch qualitative Ergebnisse der Evaluation von Forschung und Lehre sowie von Berufungsverfahren berücksichtigt werden.“ - hieß es wörtlich in dem Papier.

Die genauere Entscheidung über die Kriterien sollte beim jeweiligen Wissenschaftsminister liegen, aber auch auf die Hochschulleitungen übertragen werden können. Als eine „denkbare“ Spezifizierung der zu erbringenden Leistungen wurde genannt, daß diese sich für die Lehre aus dem Lehraufwand nach Stunden, der Hörerzahl in den Lehrveranstaltungen (bezogen auf die Gesamtzahl der Studierenden im jeweiligen Studiengang und Fachsemester), der Anzahl der abgenommenen Prüfungen, der betreuten Diplomarbeiten, der Mentorentätigkeit und dem Engagement bei der Studienreform ergeben können, da dies auch die am häufigsten genannten Kriterien bei der Vergabe von Lehrpreisen seien. Für die Forschung wurde die Anzahl der betreuten Promotionen, Habilitationen, eingeworbene Drittmittel, Publikationen und Patente als Kriterien vorgeschlagen. Der Streit entzündete sich zuvor häufig an den Kriterien und an der Verknüpfung von Evaluationen mit der Vergabe finanzieller Mittel. Inzwischen gehört die Frage, ob nicht die Messung von Hochschulleistungen durch Evaluation wegen des grundgesetzlichen Schutzes der Wissenschaftsfreiheit unzulässig sei, der Vergangenheit an (vgl. Behrens 2001: 73). Dieser Streit entschärfte sich vermutlich aber auch deshalb etwas, da die Funktion von Evaluation von einigen politischen Funktionsträgern, die Mitglied in dieser Bund-Länder-Arbeitsgruppe waren und den Prozeß entscheidend mitgestalteten<sup>86</sup>, stärker bzw. primär in der Selbstreflektion der Hochschullehrer gesehen wurde (vgl. Meyer 1997).<sup>87</sup> Damit wurde indirekt eine Art Grenzziehung zwischen Evaluationen (v.a. zur Selbstreflektion bzw.

---

<sup>86</sup> In diesem Sinne äußerte sich der Sächsische Wissenschaftsminister Meyer auf dem Symposium zum Thema „Qualitätsentwicklung in einem differenzierten Hochschulsystem“ des Stifterverbandes für die Deutsche Wissenschaft am 9. Januar 1997 im Wissenschaftszentrum Bonn. Der Sächsische Wissenschaftsminister Meyer war Vorsitzender der Expertenkommission „Reform des Hochschuldienstrechts“ (2000).

<sup>87</sup> In diesem Sinne äußerte sich der Sächsische Wissenschaftsminister Meyer auf dem Symposium zum Thema „Qualitätsentwicklung in einem differenzierten Hochschulsystem“ des Stifterverbandes für die Deutsche Wissenschaft am 9. Januar 1997 im Wissenschaftszentrum Bonn.

Verbesserung) und Berichtssystemen (v.a. zur Rechenschaftslegung) gefördert,<sup>88</sup> wie sie im Abschnitt zu Instrumenten der Leistungs- und Qualitätsbewertung formuliert wurde.

Mit dem 2002 in Kraft getretenen Professorenbesoldungsreformgesetz wurden nun im einzelnen drei prinzipielle Anlässe für Leistungszulagen definiert und deren Arten wie folgt in Form eines kurzen Begriffes<sup>89</sup> charakterisiert:

1. Übernahme besonderer Funktionen in der Selbstverwaltung (Funktionszulagen)
2. Berufungs- und Bleibeverhandlungen (Berufungszulagen)
3. besondere Leistungen in Forschung, Lehre, Kunst, Weiterbildung und Nachwuchsförderung (Leistungszulagen)

Nachfolgend sind für das mit dieser Arbeit zu verfolgende Thema v.a. die Leistungszulagen interessant. Zudem werden diese als „der Dreh- und Angelpunkt der Besoldungsreform gesehen“ (vgl. Arnhold/ Kandel 2004: 17). Die bisher ausführlichste Dokumentation anhand von praktischen Erfahrungen bei der Einführung solcher Zulagen im Hochschulbereich liegt in Form der Publikationen des CHE vor, die allerdings in der Tendenz eher optimistisch ausfällt. Dem gegenübergestellt werden können Dokumentationen der Hans-Böckler-Stiftung zu Erfahrungen mit betrieblichen Leistungslohnsystemen, die in der Tendenz eher skeptisch ausfallen. Die jeweils begrenzte Reichweite dieser sehr unterschiedlich gearteten Publikationen muß natürlich berücksichtigt werden. Die Dokumentation des CHE eignet sich die aufgrund ihrer relativ systematischen Aufarbeitung zur Systematisierung möglicher Vorgehensweisen bei der Vergabe von Leistungszulagen. Die Dokumentationen der Hans-Böckler-Stiftung ermöglichen eine kritische Reflektion anhand von Erfahrungen aus der betrieblichen Praxis. Allerdings ist hierbei zu beachten, daß diese Erfahrungen aufgrund ihres anderen Kontextes nur teilweise auf den Hochschulbereich übertragen werden können. Sie könnten jedoch den Blick schärfen für möglicherweise künftig auch im Hochschulbereich auftretende Probleme.

---

<sup>88</sup> Zwei Jahre zuvor vertrat Meyer noch die Auffassung, daß Evaluation für die Zuweisung finanzieller Mittel an die Hochschulen und Fakultäten herangezogen werden sollte. Zwar war dies nach Paragraph 118 des damals gültigen Sächsischen Hochschulgesetzes nicht ausgeschlossen, jedoch sah Meyer (1997) dies nicht mehr als Normalfall, sondern nur noch als eine Art „Notstandsrecht“.

<sup>89</sup> Die juristisch korrekten Begriffe für die Zulagenarten werden hier aus Gründen der Verständlichkeit und Einfachheit nicht verwendet (vgl. auch Witte u.a. 2002: 3).

### **Möglichkeiten der Vergabe von Leistungszulagen**

Grundsätzlich sind zwei Möglichkeiten der Vergabe von Leistungszulagen denkbar, die nachfolgend genauer erläutert werden (vgl. Witte u.a. 2002: 8, Kundelbusch 2001: 23, Klein-Schneider 1999: 93ff.):

- eine formelorientierte Zuweisung oder
- eine diskretionäre Zuweisung.

Die formelorientierte Verteilung (zunächst vom Land an die Hochschulen) könne demnach aus einer Formel errechnet werden, in die neben dem landesweiten und dem Hochschuldurchschnitt der Besoldung je C-Stelle auch Leistungsindikatoren eingehen, deren Gewichtung nach und nach erhöht würde. Die diskretionäre Zuweisung könne nach Witte u.a. (2002: 8) z.B. über Zielvereinbarungen erfolgen.

Für die einzelnen Lehrenden wird die Frage nach angemessenen, d.h. wissenschaftsadäquaten Verfahren und Kriterien zum zentralen und gleichzeitig neuralgischen Punkt. Hierbei müssen die Zulagen „dem Grunde und der Höhe nach willkürfrei bestimmt werden und nachvollziehbar sein“ (vgl. Witte u.a. 2002: 19, ähnlich Tondorf/ Jochmann-Döll 2003: 430).

Hierfür könne bei der formelgebundenen Vergabe ein verbindlicher, abgeschlossener Katalog messbarer Leistungskriterien durch das Land oder die Hochschule mit definierter Wertigkeit jedes einzelnen Merkmals festgelegt werden. Die Höhe der Zulagen ergäbe sich demnach „automatisch“ als Algorithmus aus dem Grad, in dem die einzelnen Kriterien erfüllt wären (vgl. Witte u.a. 2002: 19ff.).

Bei einer diskretionären Vergabe bzw. Vergabe nach Abwägung für die einzelnen Lehrenden müssen vor allem die Verfahrensregeln so gesetzt werden, daß sie eine nachvollziehbare, wissenschaftsadäquate Entscheidungsfindung gewährleisten könne. Die konkrete Vergabe der Zulagen erfolgt dann „im Zuge konkurrierender Abwägungsentscheidungen, die auch Ergebnisse einer regelmäßigen, stichtagsbezogenen Beurteilung aller Professoren berücksichtigen“ kann (vgl. Witte u.a. 2002: 20).

Bei der Vergabe nach Abwägung könne dies wiederum unterteilt werden in eine komplett individualisierte Bemessung ohne jede Vorgabe bezüglich Höhe (und Dauer) oder in Zulagenstufen. Während sich erstere für Zielvereinbarungen anböten, gelte dies für letztere vor allem (für i.d.R. unbefristete) Zulagenstufen im Rahmen von regelmäßigen Leistungs-Reviews. Der Vorteil individueller Bemessung sei, daß grundsätzlich die konkrete Situation eines einzelnen Professors flexibler berücksichtigt werden könne. Dies würde allerdings mit einem hohen Verfahrensaufwand und dem Risiko erkaufte, daß das System als intransparent und die

Einzelentscheidungen als nicht ausreichend nachvollziehbar und legitimiert erscheinen könnten. Da mangelnde Transparenz auch ein Grund sein kann, Systeme der Leistungsvergütung gerichtlich anzufechten, sollte dies nicht unterschätzt werden (vgl. Tondorf/ Jochmann-Döll 2004: 430, Jackson 1997: 105). Und schließlich erlauben Zulagenstufen, die Zulagen hinsichtlich ihrer Streuung, Spreizung und Höhe besser zu planen (vgl. Arnhold/ Schreiterer 2002: 18). Bisher haben alle Hochschulen Stufenmodelle favorisiert (vgl. Arnhold/ Kandel 2004: 17).

Allerdings ist hier aus den Ausführungen der Autoren nicht ersichtlich, warum beides, also eine individualisierte und eine abgestufte Bemessung, nicht auch bei formelgebundener Vergabe grundsätzlich möglich sein soll. So könnten auch bei formelgebundener Vergabe entweder vorab bestimmte zu überspringende Schwellen (Mindestanforderungen bezüglich einzelner Kriterien oder Leistungspunktintervalle für eine bestimmte Leistungsstufe) festgelegt werden oder (nach einschlägigen mathematischen Verfahren bestimmbar) signifikant über dem Durchschnitt liegende Leistungen zur Entscheidung über eine gestufte Vergabe herangezogen werden. Die Autoren sprechen selbst weiter davon (vgl. Witte u.a. 2002: 19), daß die Zulagen die Erbringung besonderer, „*erheblich über dem Durchschnitt liegender* individueller – und individualisierbarer – Leistungen“ voraussetze.<sup>90</sup> Auch die HRK (1998c: 5) sprach in ihren Empfehlungen von „überdurchschnittlichen Lehrleistungen“ als Grundlage für Leistungszulagen. Mit einer Entscheidung über am Durchschnitt orientierte Leistungsstufen wäre gleichzeitig die über eine gewisse Kontingentierung verbunden, da hiermit die Anzahl derjenigen begrenzt würde, die diese überdurchschnittliche Leistungsstufe erreichen können. Dies wäre hierbei abzuwägen. So weist Kundelbusch (2001: 24) darauf hin, daß bei solchen Vorgaben Personen auch bei gleich bleibender Leistung heruntergestuft werden können, wenn das Leistungsniveau der Mehrheit der Mitarbeiter steigt. Dadurch könne ein solches Zulagensystem ggf. keine Motivation entfalten und sogar zur Demotivation führen. Allerdings wird dies je nach Kontext und Ziel (wenn z.B. kontinuierliche Verbesserung angestrebt wird) selbst von Betriebsräten bzw. Gewerkschaftsvertretern sehr unterschiedlich bewertet.<sup>91</sup>

---

<sup>90</sup> Nach dem Verständnis empirischer Sozialwissenschaften könnte der Begriff signifikant (z.B. nach einschlägigen Konventionen auf dem 5%-Alpha-Fehler-Niveau) nachprüfbarer bestimmt werden als der Begriff „erheblich“.

<sup>91</sup> Beispiele für Unternehmen, in denen nicht nur diese Begrenzung der Leistungszulagenbudgets nach oben praktiziert werde, sondern darüber hinaus jedes Jahr die fünf bis 10 Prozent der sogenannten „Minderleister“ ausgetauscht werden, nennt Conrady (2003: 13). Von der IG Metall Bayern wurde dies als Kultur kritisiert, in der „Angst und Streß als Motivatoren“ dienen. Dagegen wird das Entgeltrahmenabkommen (ERA), das die IG-Metall-Bezirke Küste und Baden-Württemberg abschlossen (welches solche Listen von „Minderleistern“ auch nicht verhindern könne, sich aber auf messbare und damit überprüfbare Kennziffern stütze) als faktische Erweiterung der Mitbestimmung bewertet, „da die Tarifparteien nun gemeinsam Leistungskriterien festlegen“ und die Überprüfbarkeit der Leistungskriterien erlaube, „Vorgesetzten mit sachbezogenen Argumenten Paroli zu bieten“ (vgl. Conrady 2003: 15).

Wissenschaftliche Untersuchungen, die die Zielerreichung (z.B. Motivationssteigerung) nach der Einführung von Leistungsentlohnung in Deutschland untersuchen, gibt es nach Bahn Müller (2001: 428) kaum. Auch ein Blick in Länder mit längeren Erfahrungen hiermit wie Großbritannien hilft nicht unbedingt weiter: So zitiert Jackson (1997: 101) in einem Überblicksartikel zu variablen Vergütungen an britischen Universitäten mehrere Studien, die (je nach Auftraggeber) zu unterschiedlichen Ergebnissen kamen. Während die eine Studie eine Motivationssteigerung, Reduktion der Personalfuktuation, Veränderung der Organisationskultur sowie einen höheren Grad an Fairneß konstatierte, kam eine andere Studie zu dem Schluß, daß leistungsorientierte Bezahlung auch die Motivation verringern kann (vgl. Jackson 1997: 101).<sup>92</sup>

Bahn Müller interpretiert das in Deutschland „offensichtliche Desinteresse an einer Zielerreichungskontrolle“ allerdings nicht als Rationalitätsmythos, sondern als „durchaus rationale Strategie der Legitimations- und damit Bestandssicherung“ der jeweiligen Institution zur „Herstellung von Isomorphie mit den Erwartungen relevanter Akteure der Umwelt“ (vgl. Bahn Müller 2001: 429). Daher stelle sich die Frage nach dem „ob“ solcher Leistungsanreizsysteme gar nicht erst und es gehe nur noch um die Frage, „wie“ dies umzusetzen ist.

Da sich die Frage des „ob“ nach der gesetzlichen Festlegung für den Hochschulbereich ebenfalls zunächst nicht mehr stellt, soll nach diesem kleinen Exkurs weiter der Systematisierung nach Arnhold/ Schreiterer (2002) gefolgt werden: Wie hoch oder niedrig<sup>93</sup> man die Schwellen im einzelnen legt oder wie viele Stufen festgelegt werden, hängt demnach von der „Wettbewerbskultur“ der jeweiligen Hochschule ab (z.B. drei Stufen: für Spitze/ überdurchschnittlich, im Schnitt und unterdurchschnittlich – vgl. auch Teichler 2003: 5).<sup>94</sup> Zwar sah seit 1997 das Bundesbesoldungsgesetz bereits die Vergabe von Leistungszulagen vor, jedoch für maximal 10% der grundsätzlich Empfangsberechtigten.<sup>95</sup> Im Gegenzug ist „bei unterdurchschnittlicher Leistung“ eine zeitliche Hemmung und Streckung des Aufstiegs möglich, so daß sich auch hier grundsätzlich die drei Gruppen für im Schnitt, über und unter dem Schnitt liegende Empfangsberechtigte ergeben. Allerdings wurde dies bisher nur auf Bundesebene rasch umgesetzt, im Bereich der Länder wurden leistungsorientierte Besoldungselemente nur zurückhaltend und selektiv genutzt, da es u.a. aufgrund der Finanzlage nicht praktikabel erschien (vgl. Tondorf/

---

<sup>92</sup> Hinweise dazu, unter welchen Bedingungen eine Verringerung der Motivation durch leistungsorientierte Bezahlung vermieden werden kann, geben Ziegele/ Handel (2004: 20 – vgl. Ausführungen am Ende dieses Abschnittes).

<sup>93</sup> Zur Diskussion über die Höhe von Leistungszulagen im öffentlichen Sektor vgl. Tondorf (2003: 10ff.).

<sup>94</sup> Denkbar wären natürlich auch andere Modelle mit z.B. 5, 6 oder mehr Leistungsstufen wie z.B. in England, die sich an bestimmten Prozentvorgaben orientieren, wie viele Professoren (maximal) eine bestimmte Leistungsstufe erhalten können (vgl. Expertenkommission „Reform des Hochschuldienstrechts“ 2000: 50, Arnhold/ Schreiterer 2002: 25ff., Arnhold/ Kandel 2004: 17ff.).

<sup>95</sup> 2002 wurde die Quote für Leistungszulagen auf 15% erhöht. (vgl. Tondorf/ Jochmann-Döll 2004: 428).

Jochmann-Döll 2004: 428). Generell befinde sich der öffentliche Sektor im Hinblick auf solche Leistungsanreize noch im „Experimentierstadium“ (vgl. Jochmann-Döll/ Tondorf 2004: 159).

Was die eingangs aufgeworfene Frage nach formelorientierter oder diskretionärer bzw. abwägender Vergabe von Zulagen betrifft, beschreiben Arnhold/ Schreiterer (2002: 17) die Entscheidung dieser Frage als ein „Dilemma“: Eine formelorientierte Vergabe sei „ein immer auch ‚mechanisch‘ wirkender Formelautomatismus“, der u. U. weder den Unterschieden zwischen den einzelnen Fächern angemessen gerecht werde (wenn man das Verfahren nicht separat für die einzelnen Fächer durchführt) noch der unterschiedlichen Leistungsfähigkeit und individuellen Leistungsentwicklung einzelner Professoren“. Andererseits könnte allein der Verfahrensaufwand eine abwägende Vergabe bei größeren Hochschulen (insbesondere bei individueller Bemessung) unpraktikabel werden lassen, von den zu erwartenden Akzeptanzproblemen nicht ausreichend nachvollziehbar erscheinender Vergabe einmal abgesehen.

Ein gangbarer Ausweg wird in einer gesicherten Regelmäßigkeit und Transparenz des Verfahrens gesehen, wie sie ihr Vorbild in Prüfungsentscheidungen und Berufungsbeschlüssen hätten. Diese unterliegen im Kern nicht der juristischen Anfechtbarkeit, sondern nur im Hinblick auf Verfahrensfragen. Hierbei müsse sich das Verfahren auf Informationsgrundlagen stützen und daher auch mit Indikatoren begründen lassen, so die Wertung von Arnhold/ Schreiterer (2002: 17). Die Expertenkommission „Reform des Hochschuldienstrechts“ (2000: 5) hatte zuvor in diesem Sinne formuliert: „Im Bereich der Lehre werden die Ergebnisse studentischer Veranstaltungsevaluation *eine Basis* der Beurteilung durch Fachgutachterinnen/ Fachgutachter bilden.“ Solche Befragungsergebnisse sollen demnach nicht direkt einfließen, aber Grundlage sein. Auch das bmb+f (2000b: 19) und die HRK (1998: 5) hatten bereits zuvor studentische Lehrveranstaltungsevaluationen als Basis der Beurteilung durch die Fachgutachter vorgeschlagen.

Abschließend formulieren Arnhold/ Schreiterer (2002: 17) hierzu, daß eine klare Umschreibung der Leistungsdimensionen und eine hochschulinterne Einigung über entsprechende Leistungsindikatoren zwar unverzichtbare Voraussetzungen für ein stimmiges, rationales und akzeptanzfähiges Verfahren wären. Sie entbinde aber nicht von Ermessensentscheidungen, sondern strukturiere diese vielmehr.

Jackson (1997: 103) formulierte bei der Einführung von leistungsabhängigen Vergütungen an britischen Hochschulen wesentliche Aspekte in ähnlicher Weise: „Universities and local university managements must ensure they have necessary resources and expertise to define appropriate and objective criteria against which performance can be assessed, as well as carrying out assessment in a fair and defensible manner.“ An anderer Stelle fügt er hinzu: “The absence

of clear and visible systems for measuring performance has been recognized by many within universities as a major problem facing the introduction of schemes for distributing discretionary pay” (vgl. Jackson (1997: 107).

### **Erfahrungen mit leistungsorientierter Vergütung in der Wirtschaft**

Ähnliche Einschätzungen wie von Arnhold/ Schreiterer (2002) und Jackson (1997) finden sich auch zur leistungsorientierten Entlohnung in der Wirtschaft. In Veröffentlichungen aus der deutschen Arbeitsforschung wird ebenfalls darauf hingewiesen, daß Transparenz und Klarheit von Leistungsindikatoren erfolgskritische Aspekte von monetären Leistungsanreizen sind, dies allein aber nicht ausreicht (vgl. Tondorf/ Jochmann-Döll 2004: 429, Bahn Müller 2001: 426ff.). Insbesondere wird betont, „daß die Bedeutung der Prozeßgestaltung für die Akzeptanz und Anwendung des Leistungsvergütungssystems vielerorts unterschätzt wird“ (vgl. Tondorf/ Jochmann-Döll 2004: 432).

Dennoch ist leistungsorientierte Entlohnung (häufig in Zusammenhang mit Zielvereinbarungen) in der betrieblichen Praxis bereits weit verbreitet. Nach den Ergebnissen von Betriebsbefragungen gibt es in der Hälfte bis zu zwei Dritteln der Angestellten „eine mehr oder weniger direkte Verknüpfung“ zwischen der Bewertung der Leistungen und der Entlohnung (vgl. Bahn Müller 2001: 431). Allerdings ist hier – das zeigt die Spannweite der Einschätzungen – zwischen der Perspektive des Managements und der Arbeitnehmer bzw. ihrer Vertretungen zu unterscheiden. So stammt die Einschätzung, daß diese Verknüpfung bei zwei Dritteln der Angestellten existiere, aus direkten Betriebsbefragungen und zeigen in diesem Fall die Managementperspektive. Die Einschätzung, daß diese Verknüpfung bei der Hälfte der Angestellten existiere, stammt aus Betriebsrätebefragungen. Der Unterschied ergibt sich u. a. daraus, daß ein Teil der Beschäftigten das, was das Management als Zielvereinbarung einstuft, als Zielvorgabe ansieht. Wenn aber ein gesetztes Ziel als *Zielvorgabe* (und evtl. dazu noch als unrealistisch hohes Ziel)<sup>96</sup> angesehen wird, kann es nicht die gewünschte Anreizwirkung entfalten. Als wesentliche Grundlage wird daher betont, daß bei Zielvereinbarungen die Ziele beeinflussbar und erreichbar sein müssen. Diesen Aspekt gilt es auch bei Überlegungen zu Leistungsanreizen an Hochschulen insbesondere bei Zielvereinbarungen zu beachten.

In diesem Zusammenhang ist auf eine britische Erfahrung der Verknüpfung der Evaluation von Forschern und Mittelzuweisung hinzuweisen, nach der die schlechter bewerteten Forscher durch die finanziellen Einbußen ihrer Einrichtung infolge der schlechteren Bewertung kaum eine

---

<sup>96</sup> Auch für Zielvereinbarungen gelte, daß die Ziele vom Mitarbeiter beeinflussbar sowie für ihn erreichbar und zumutbar sein müssen (vgl. Kundelbusch 2001: 24).

Chance haben, wieder „aufzusteigen“. Ihre Chancen zum Wechsel an eine besser ausgestattete Hochschule bei wiederholter schlechterer Bewertung würden ebenfalls gering sein. Ähnlich verhalte es sich bei der Neugründung von Forschungseinrichtungen. Ahrens (2000: 184) schätzt daher ein: „Nach der alten Regel `Wer hat, dem wird gegeben` etablieren sich andauernde Ranglisten, die der Dynamik und Innovationskraft von neuen Forschungseinrichtungen und einzelnen Forschern die Schubkraft nehmen und sich kontraproduktiv auswirken“. Hier könnten – richtig angewendet – Zielvereinbarungen entgegenwirken, da sie im Gegensatz zu der stärker auf die Vergangenheit orientierten formelgebundenen Vergabe eher in die Zukunft gerichtet sind und damit in stärkerem Maße Chancen für angestrebte Verbesserungen eröffnen.

### **Mögliche nicht-intendierte Wirkungen bzw. „Fehlanreize“**

Während die bisher genannten Autoren auf Probleme von leistungsorientierter Besoldung (bzw. leistungsorientierter Entlohnung in der betrieblichen Praxis) hinweisen, deren Angemessenheit sie jedoch nicht grundsätzlich bestreiten, stellen einige Autoren die Realisierbarkeit der angestrebten Ziele im Hochschulbereich zum derzeitigen Zeitpunkt grundsätzlich in Frage. So formuliert Stadler (2003: 334ff.) daß die aufgrund der derzeitigen Rahmenbedingungen entstehenden „Opportunitätskosten“ eines Hochschullehrers „Fehlanreize für Professoren verursachen kann, die die Realisierung der Zielvorgaben in Frage stellen“.

Stadler beschreibt anhand einer vertragstheoretischen Analyse mit Hilfe des Prinzipal-Agenten-Ansatzes (den auch Ziegel/ Handel 2004: 4 zugrundelegen) unter Einbeziehung der beiden Haupttätigkeitsfelder eines Professors (Forschung und Lehre) die Erfolgsaussichten der 2002 in Kraft getretenen Reform der Professorenbesoldung. Stadler diskutiert hierbei insbesondere die Auswirkungen des Problems, daß sich bislang „Forschungserfolge vergleichsweise gut objektiv messen lassen, nicht aber die Leistungen in der Lehre“ (vgl. Stadler 2003: 336). Als Indikatoren stünden bislang lediglich Studierendenzahlen (bzw. Auslastungsquoten) und Lehrevaluationen auf einer zunehmend breiteren Basis zur Verfügung. Laut Stadler stellen die so gewonnenen Informationen „eine wertvolle Hilfe (...) für die verbesserte Gestaltung der Lehre dar. Über die Qualität der wissenschaftlichen Ausbildung, die die Studierenden eigentlich erst in ihrem späteren Beruf kritisch einschätzen können, sagen diese Daten allerdings wenig aus. Eine (...) erforderliche Rückrechnung der wissenschaftlichen bzw. beruflichen Erfolge (einschließlich der Vergütungen) der Absolventen scheitert an fehlenden Daten.“<sup>97</sup> Daher ging er in seinem Modell

---

<sup>97</sup> Diese Einschätzung wird vom Verfasser dieser Arbeit nicht geteilt. Es gab Ende der 1990er Jahre bereits eine Reihe von Untersuchungen zum beruflichen Erfolg von Hochschulabsolventen, wenngleich diese bei weitem noch nicht flächendeckend vorliegen. Einen Überblick hierzu geben Burkhardt / Schomburg / Teichler (2000). Inzwischen dürfte deren Zahl noch deutlich gestiegen sein. Allerdings werden diese Untersuchungen bisher kaum

von der ausschließlichen Messbarkeit des Forschungserfolges aus. Demzufolge ergibt sich in keiner seiner Modellrechnungen (nicht einmal in einem Modell unter Berücksichtigung diverser „spill-over“-Effekte zwischen Forschung und Lehre) ein klarer positiver Anreiz für einen Hochschullehrer, mehr Zeit und Engagement in die Lehre zu investieren. Im realistischeren Szenario „substitutiver Anstrengungen in immer spezialisierterer Forschung einerseits und möglichst breit und praxisnah angelegter Lehre“ kommt es gar zu einem als negativ einzuschätzenden Anreiz, die Anstrengungen für die Lehre zu minimieren. Stadler (2003: 336) spricht daher von „erheblichen Risiken in Form von Fehlanreizen zu Lasten der Lehre“.

Stadler (2003: 338) schließt daher mit einem Plädoyer für eine Entlohnung „auf der Basis subjektiver Leistungsbeurteilungen“, wobei er darauf verweist, daß solche subjektiven Beurteilungen im Rahmen von Berufungs- und Bleibeverhandlungen längst existieren. Neben den dabei natürlich wichtigen Publikationen und eingeworbenen Drittmitteln nennt er auch Lehrveranstaltungen, die dazu zählen sollten. Ziegele/ Handel (2004: 20) formulieren, daß neben Angaben zu entscheidungsrelevanten Bereichen, die Professoren selbst zusammenstellen, auch unabhängige Evaluationen und studentische Veranstaltungskritik in Beurteilungen eingehen sollen. Andere Autoren, die die Realisierbarkeit der angestrebten Ziele mittels leistungsorientierter Besoldungszulagen aufgrund von Fehlanreizwirkungen grundsätzlich in Frage stellen, argumentieren häufig damit, daß viele Hochschullehrer v.a. intrinsisch motiviert wären und derart motivierte Professoren durch materielle Anreize ihre bisherige Motivation eher abbauen würden (vgl. z.B. Minssen/ Wilkesmann 2003: 117, Ronge 2000: 189). Allerdings liegen für den Hochschulbereich bislang kaum empirische Untersuchungen vor und die wenigen haben z.T. methodische Mängel, die deren Aussagekraft in Frage stellen (vgl. ausführlichere Diskussion der Studie von Minssen/ Wilkesmann (2003) im Abschnitt zu leistungsorientierter Mittelvergabe).

Zusammenfassend kann zu leistungsabhängigen Besoldungszulagen festgehalten werden: Neben der grundsätzlichen Eignung der Leistungskriterien, nebst grundsätzlich beeinflussbaren und erreichbaren Leistungszielen und einem dem Hochschulbereich angemessenen und transparenten Verfahren der Vergabe von Leistungszulagen ist es notwendig, den Gegenstand der Leistungsbeurteilung umfassend in allen seinen wesentlichen Dimensionen in die Leistungsbewertung einzubeziehen (vgl. Ziegele/ Handel 2004: 20). Um die vorgegebenen Ziele zu erreichen und Fehlanreize zu vermeiden, wären also nicht aufgrund der leichteren Verfügbarkeit nur quantita-

---

als eine Grundlage für die Bemessung von leistungsabhängigen Besoldungszulagen diskutiert (z.T. aus guten Gründen, da längst nicht alle diese Studien methodischen Standards entsprechen). Unabhängig davon wird ist aber in Frage zu stellen, ob solche – meist für Aussagen auf Fächerebene konzipierten Studien – Aussagen über den Beitrag einzelner Hochschullehrer zum beruflichen Erfolg ermöglichen.

tive Kriterien einzusetzen, oder solche, die die Forschungsdimension erfassen. Vielmehr müssen auch Kriterien gefunden oder geschaffen werden, die die Qualität bzw. die Lehre erfassen.

### **3.4.3 Leistungsorientierte Verteilung finanzieller Mittel bzw. Ressourcen**

Grundsätzlich ist zu dieser Problematik zu erwähnen, daß in Deutschland anders als in vielen entwickelten Industrieländern die Hochschulfinanzierung nicht über einen „Zentralstaat“, sondern durch die einzelnen Bundesländer erfolgt. Zudem befindet sich dieser Bereich derzeit in Deutschland in einer äußerst dynamischen Entwicklung (vgl. Leszcensky u.a. 2004: 1). Es ist daher von vornherein davon auszugehen, daß Informationen „nur lückenhaft vorhanden und vielfach veraltet“ sind (vgl. Hartwig/ Küpper 2002: 5). Zwar geben Hartwig/ Küpper (2002) einen Überblick über die Entwicklung in einigen Bundesländern. Und z.B. mit den Veröffentlichungen von Ziegele (2003b), Minssen/ Wilkesmann (2003) sowie Leszcensky/ Orr (2004) gibt es auch detaillierte Ausführungen zu einzelnen Bundesländern. Andere Autoren betonen allerdings auch: „Leistungs- und belastungsbezogene Mittelverteilung bezieht sich nicht allein auf das Verhältnis zwischen Staat und Hochschulen, sondern ist auch innerhalb der Hochschulen relevant“ (vgl. Behrens 2001: 68). Da in dieser Arbeit der Schwerpunkt auf den Qualitätsaspekt von Leistung gelegt werden soll und Möglichkeiten zur Förderung der Qualitätsorientierung – soweit es denn mit solchen Anreizmodellen möglich ist – stärker auf der Ebene der Hochschulen und Fakultäten bzw. Fachbereiche gesehen werden (vgl. Abschnitt zur Eingrenzung des Qualitätsbegriffs anhand von Zielklärungen), erfolgt hier eine Schwerpunktsetzung auf die Ebene innerhalb der Hochschulen.

Einen relativ aktuellen Überblick über die Entwicklung „leistungsorientierter Ressourcensteuerungsverfahren an deutschen Hochschulen“ gibt Schröder (2004: 28-59). Allerdings schränkt Schröder ein, daß es aufgrund des kontinuierlichen Veränderungsprozesses zwischenzeitlich bereits wieder zu Änderungen gekommen sein kann, womit er Ungenauigkeiten einräumt.<sup>98</sup>

Da im Rahmen dieser Arbeit nicht einmal annähernd die zeitlichen und personellen Ressourcen<sup>99</sup> für ein solches Unterfangen vorhanden sind, muß nachfolgend jeder Versuch einer flächendeckenden aktuellen Bestandsaufnahme unterbleiben und eine Beschränkung auf die Vorstellung wesentlicher Grundprinzipien anhand einiger ausgewählter Beispiele erfolgen.

Unter „leistungsorientierter Ressourcensteuerung“ versteht Schröder eine outputorientierte Steuerung. Mit Steuerung meint Schröder allerdings keine strategische Steuerung (wie z.B. im

---

<sup>98</sup> So gab es z.B. an der TU Dresden Weiterentwicklungen, auf die zwecks Erhaltung einer einheitlichen Datengrundlage jedoch nicht weiter eingegangen wird. Zudem sind die Informationen nicht öffentlich zugänglich.

Controlling), sondern lediglich die Steuerung von Finanzströmen (vgl. Schröder 2004: 53). Um Mißverständnissen vorzubeugen, ist nachfolgend nur noch von der Verteilung finanzieller Mittel bzw. kurz von Mittelvergabe die Rede. Die Verteilung der finanziellen Mittel erfolgt dabei „in Abhängigkeit eines vereinbarten oder bereits erzielten Ergebnisses“ (vgl. Schröder 2004: 30).

### **Arten leistungsorientierter Mittelvergabe**

Bei der leistungsorientierten Mittelvergabe ist grundsätzlich zu unterscheiden in Formelmodelle und diskretionäre, also einzelfallbezogene Verteilung (vgl. Ziegele 2003b: 3, Behrens 2001: 76 sowie Abschnitt in dieser Arbeit zu leistungsorientierten Besoldungszulagen). Auch Schröder nennt (wie bereits andere Autoren bei leistungsorientierten Besoldungszulagen) an erster Stelle der Varianten diskretionärer Verteilung Zielvereinbarungen. Eine andere Variante ist die über Finanz- und Stellenpools (vgl. Schröder 2004: 30) oder zweckgebundene Projektmittel (vgl. Leszcensky u.a. 2004: 5). Hierbei handelt es sich aber lediglich um spezielle Formen der diskretionären Verteilung mittels abwägender Entscheidung. Die Vor- und Nachteile der beiden grundsätzlichen Unterscheidungen wurden bereits relativ ausführlich im Zusammenhang mit leistungsorientierten Besoldungszulagen ausgeführt. Da die Problematik ähnlich ist, soll an dieser Stelle nicht noch einmal auf die bereits bekannten Aspekte eingegangen werden, sondern v.a. die Besonderheiten der Bewertung leistungsorientierter Mittelvergabe angesprochen werden. Eine Besonderheit ist, dass hier grundsätzlich auch nach Sach- und Personalmitteln zu unterscheiden ist. Bisher wurden jedoch im wesentlichen Sachmittel in die leistungsorientierte Mittelvergabe einbezogen, so dass sich die Erfahrungen zunächst hierauf beschränken. Für eine Einordnung der von Schröder vorgebrachten Pro- und Contra-Argumente erscheint es wichtig, noch einmal kurz auf die Datengrundlage seiner empirischen Untersuchung einzugehen: Schröder bezog sich hierbei auf fünf Hochschulen in verschiedenen Bundesländern, für die er im einzelnen Analysen von hochschuleigenen Veröffentlichungen und Statistiken sowie Interviews mit Mitgliedern der Hochschulleitungen, der Verwaltungen und mit Wissenschaftlern anhand „teilstandardisierter Fragebögen“ durchführte (vgl. Schröder 2004: 33). Schröder weist darauf hin, daß es sich aufgrund der Grundgesamtheit der von ihm interviewten Personen nach statistischen Maßstäben nicht um eine repräsentative Größe handelt. Er sieht die Stärken seiner Vorgehensweise v.a. darin, daß sie hilfreiche Aufschlüsse über die Wirkungsweise, Probleme und Grenzen der Anwendung leistungsbezogener Ressourcenverteilung gebe. Diese Interviews sind neben zur Verfügung gestelltem schriftlichem Material auch die Datengrundlage für die

---

<sup>99</sup> Schröder arbeitete in einem von der DFG geförderten Projekt (Laufzeit: 5 Jahre) mit, welches sich auf die beispielhafte Untersuchung von 5 technisch dominierten Hochschulen beschränkte. Auswahlkriterium war aber u.a.,

nachfolgende tabellarische Übersicht zur Gliederung der leistungsorientierten Mittelvergabe nach funktionaler Zuordnung in Anlehnung an Schröder (2004: 35).

Die Übersicht kann – obwohl sie zur Herstellung einer ausgewogeneren Fächerstruktur nach eigenen Recherchen um zwei nicht technisch dominierte Hochschulen erweitert wurde<sup>100</sup> – keine Übertragbarkeit auf den gesamten Hochschulbereich beanspruchen. Wie bereits erwähnt, existiert für Deutschland bisher keine vollständige Übersicht. Die Gliederung sollte aber dennoch für eine Systematisierung hilfreich sein. Sie erfolgte bei Schröder einerseits nach leistungs- und belastungsorientierten Anteilen der Verteilungsmodelle, andererseits nach forschungs- und lehrbezogenen Anteilen. Da die Zuordnung zu leistungs- und belastungsorientierten Anteilen sehr unterschiedlich gesehen wird,<sup>101</sup> und es in der hier vorliegenden Arbeit v.a. um die lehrbezogenen Anteile (im Verhältnis zu forschungsbezogenen Anteilen) geht, wird nachfolgend lediglich dieser Teil diskutiert.<sup>102</sup> Der Tabelle voranzustellen ist außerdem der Hinweis, daß die leistungsorientierte Mittelverteilung in Deutschland zunächst keine Personalmittel betraf und deren Anteil am staatlichen Zuschuss insgesamt lediglich 4 bis 9 Prozent ausmachte, so daß sie bisher nach Einschätzung der von Schröder interviewten Gesprächspartner keine wesentliche Steuerungsfunktion (bzw. Anreizfunktion) übernahm. Dies könnte sich jedoch mit der geplanten Erhöhung der Anteile am staatlichen Zuschuss verändern. Im Vergleich zu Vorjahren ist festzustellen, daß es nur wenige Jahre vor der Erhebung Schröders nach einer Länderumfrage der KMK (1996: 426) jedenfalls auf Länderebene mit 1 bis 5,8% noch deutlich geringere Anteile waren. Die in der Erhebung Schröders höheren Anteile können aber auch damit zusammenhängen, daß sich die von ihm untersuchten Hochschulen ausdrücklich als Vorreiter sehen und daher die Anteile höher setzten.

---

daß sie in verschiedenen Bundesländern liegen sollten, um ein breites Spektrum möglicher Varianten zu erfassen.

<sup>100</sup> Zur Wahrung der Vergleichbarkeit in der Darstellung wurden Dokumente aus denselben Erhebungsjahren wie bei Schröder (2000/ 2001) herangezogen (Universität Oldenburg 2000, Universität Dortmund 2000). Aktuellere Dokumente waren teilweise auch nicht zugänglich, so daß eine vollständige Aktualisierung nicht möglich wäre.

<sup>101</sup> So wird das Vorhandensein (und die damit nötige Betreuung) z.B. einer großen Anzahl von Studierenden an den Hochschulen teilweise als Leistung, teilweise aber auch als Belastung eingestuft und entsprechend auch die hiermit im Zusammenhang stehenden Finanzierungsanteile eingeordnet.

<sup>102</sup> Diese Einteilung der Aufgabenbereiche in Lehre und Forschung (und z.T. in einen dritten Aufgabenbereich Gleichstellung) wird auch in mehreren Bundesländern für die Mittelverteilung auf der Ebene Land-Hochschulen vorgenommen, so neben Berlin in Baden-Württemberg, Bayern und Niedersachsen (vgl. HIS 2004: 14).

### Gliederung der leistungsorientierten Mittelverteilung ausgewählter Hochschulen (in %)

Funktion	Universität Hannover	TU Dresden	TU Berlin	TU München	RWTH Aachen	Universität Oldenburg	Universität Dortmund
forschungsbezogen	20	20	13,0	25	27,2	20	20
lehrbezogen	20	60	8,4	45	72,8	20	20
Sockelbetrag	40	20	78,6	20	-	55	40
Ziel- und Leistungsvereinbarungen	20	-	-	-	-	5	20
Summe	100	100	100	90 <sup>103</sup>	100	100	100

Quellen: Schröder (2004: 35) sowie Universitäten Oldenburg und Dortmund: eigene Darstellung

Die vorgestellten Modelle ähneln sich relativ stark in ihrem Aufbau. Ein wesentlicher Unterschied besteht in ihren funktionalen Anteilen. So ist an der TU Berlin der Sockelbetrag besonders hoch. Der Sockelbetrag übernimmt die Funktion der Grundausstattung und wird häufig an der bereits vorhandenen Personalausstattung bemessen. Auffällig ist außerdem, daß an der TU Berlin<sup>104</sup> ein hoher Anteil forschungsbezogen vergeben wird, während an der RWTH Aachen und der TU Dresden relativ hohe Anteile lehrbezogen vergeben werden.

International sind höhere Anteile für die Lehre jedoch nicht ungewöhnlich. Bei vielen OECD-Ländern ist der Anteil für die Lehre höher als für die Forschung (vgl. Leszcensky u.a. 2004: 189). Bei einer solchen Einordnung der Gliederung leistungsorientierter Mittelvergabe deutscher Hochschulen in Erfahrungen anderer Länder muß natürlich beachtet werden, daß die Angaben sich auf unterschiedliche Ebenen beziehen. Häufig findet sich diese Gliederung jedoch mehr oder weniger (z.B. fachspezifisch) angepasst in mehreren Ebenen des Hochschulsystems, um die gewünschten Anreize weiterzugeben (vgl. Orr 2001: 144). Ein Bericht der OECD (1998: 34) hält diese Folge sogar für unvermeidlich. Daher erscheint dies für eine grobe Einordnung deutscher Modelle in internationale Erfahrungen durchaus geeignet. Für dieses Ziel wird deshalb nachfolgende Übersicht vorgestellt, die die Bandbreite der Anteile leistungsorientierter Mittelvergabe für die Aufgabenbereiche Forschung und Lehre verschaulichen soll.<sup>105</sup>

<sup>103</sup> Die übrigen 10%, die weder lehr- noch forschungsbezogen verteilt werden, wurden vorab für andere Aufgaben abgezogen (vgl. Schröder 2004: 35).

<sup>104</sup> Inzwischen wurde für das Land Berlin auf der Ebene Land-Hochschulen ein weitgehend einheitliches Mittelverteilungsmodell eingeführt, wobei Lehre und Forschung mit gleicher Gewichtung eingehen (vgl. HIS 2004: 14).

<sup>105</sup> Grundlage sind einzelne ausführlichere Länderstudien, die in Kooperation mit Projektpartnern vor Ort durchgeführt wurden. Die Datenquellen befinden sich im Anhang des Projektberichtes von Leszcensky u.a. (2004: 207ff.).

Diese lehnt sich an die Systematik Schröders an und gibt einen Überblick zur Gliederung der leistungsorientierten Mittelverteilung ausgewählter OECD-Länder.<sup>106</sup> Hierbei werden als Abkürzungen die internationalen Länderkennzeichen (nach ISO 3166) verwendet.

Zur Wahrung der Übersichtlichkeit werden dort, wo Angaben nach Hochschularten differieren, nur Universitäten einbezogen. Dadurch werden z.T. (Belgien - frz. Teil, Finnland, Niederlande) die Anteile indikatorgestützter Finanzierung in Forschung (F.) und Lehre (L.) unterschätzt, da der Anteil an den Hautes Ecoles, Yrke-högskola bzw. Hogescholen höher ist.<sup>107</sup>

### Gliederung der leistungsorientierten Mittelverteilung ausgewählter OECD-Länder (in %)

Land/ Funktion	AU	BE frz.	BE fla.	CZ	DK	ES	FI	FR	GB	HU	IE	JP	NL	NO	SE
forschungs bezogen	17	-	-	7	-	-	-	-	19	1	-	30 (F.)	10	8	-
lehr- bezogen	77	100	-	51	26	84	68	60	64	49	30 (L.)	+	25	25	62
Sockel- betrag	-	-	-	12	57	4	12	30	-	47	63	70	63	60	
diskre- tionär: Zielver- einbarg./ Proj.-mittel	6	-	100 (L.) )	30	17	12	20	10	17	4 (F.)	7	-	2 (F.)	7 (F.)	38 (F.)

Quelle: Leszcensky u.a. 2004: 188ff., Tabelle: eigene Darstellung

Betrachtet man nur die OECD-Länder mit hohen indikatorgestützten Anteilen an der Gesamtfinanzierung,<sup>108</sup> liegt der Anteil für die Lehre an der indikatorgestützten Finanzierung zwischen 60 und 88%. Angesichts dessen scheint eher der höhere oder gleich hohe Anteil für Forschung gegenüber der Lehre an deutschen Hochschulen international ungewöhnlich zu sein. Hier bleibt es sicherlich abzuwarten, ob sich dies bewährt, da sich die Modelle in Deutschland im Vergleich zu anderen Ländern noch in der Erprobungsphase befinden. Beispielsweise in Großbritannien

<sup>106</sup> Die bei Leszcensky u.a. (2004: 189) als diskretionär-inkrementalistisch bezeichnete Komponente wurde hier wie bei Schröder (2004: 35) als Sockelbetrag eingeordnet, da dieser Bestandteil sich oft an Vorjahren orientiert bzw. die im vorhergehenden Zeitraum gezahlten Summen Ausgangsbasis für Verhandlungen sind.

<sup>107</sup> Deutschland konnte nicht einbezogen werden, da nicht genügend Informationen vorlagen. Grundsätzlich wäre es zwar möglich, die Spannweite der verschiedenen Anteile an einzelnen Hochschulen anzugeben. Die lehrbezogenen Anteile differieren aber so stark (8,4% bis 72,8%), daß dies nicht genügend Aussagekraft hat.

<sup>108</sup> Dies sind die Länder Australien, Frankreich, Spanien (Valencia) und England. Hier betragen die Anteile indikatorgestützter Finanzierung (Lehre + Forschung zusammen) in allen Hochschularten insgesamt über 80%.

wurden bereits 1986 Leistungsanreizmodelle eingeführt, in den folgenden Jahren schrittweise der Leistungsanteil erhöht und die Modelle modifiziert (vgl. Jackson 1997: 102ff., Orr 2001: 52). Weitere wesentliche Unterschiede zwischen Mittelverteilungsmodellen an deutschen Hochschulen zeigt nachfolgend dargestellte Auswahl der Parameter [(X) bedeutet geplant]:

### Gegenüberstellung der in den Mittelverteilungsmodellen verwendeten Parameter

Parameter	Funkt. Zuordnung	Univ. Hannover	TU Dresden <sup>109</sup>	TU Berlin	TU München	RWTH Aachen	Univ. Oldenburg	Univ. Dortmund
Auslastung	Le	X				X		
Absolventen	Le	X			X		X	(X)
Studierende im 1.-12. Fachsemester/ RSZ	Le			X			X	(X)
Studienanfänger	Le				X			(X)
Studienplätze	Le	X						
Bestandene Hauptprüfungen	Le			X				
Bedarf an Absolventen	Le				X			
Lehrangebot	Le		X			X		
Lehrnachfrage	Le					X		
Promotionen	Fo	X			X	X	X	(X)
Habilitationen	Fo	X			X		X	
Drittmittel	Fo	X	X		X	X	X	
Publikationen							X	X
Patente								X
Wiss. MA aus Drittmitteln	Fo			X				
Wissenschaftler-Arbeitsplätze	Fo	X					X	
Professuren	Fo		X		X		X	
Hochschullehrer	Fo			X				
Akademische Mitarbeiter	Fo			X	X		X	

Quellen: Schröder (2004: 37) sowie Universitäten Oldenburg und Dortmund: eigene Darstellung

<sup>109</sup> Für die TU Dresden liegen dem Verfasser Angaben zu Parametern vor, die sich z.T. deutlich unterscheiden. Es kann jedoch nicht ausgeschlossen werden, dass die Angaben zum Erhebungszeitpunkt bei Schröder korrekt waren.

Bei der Gegenüberstellung der Parameter ist auffällig, daß kein Parameter an allen Hochschulen zugleich Anwendung findet. Bestimmte Parameter, wie z.B. Drittmittel und Promotionen für die Forschung oder Absolventenzahlen für die Lehre, werden an den betrachteten Hochschulen jedoch häufiger als andere angewendet. Diese Parameter werden von der Systematik her den „Ergebnissen“ des „Prozesses“ Hochschullehre zugeordnet. Wenn sich bei weiteren Untersuchungen zu solchen Parametern die sich hier anhand einiger weniger Hochschulen abzeichnende Nutzungshäufigkeit der Parameter bestätigen sollte, könnte man daraus schließen, daß über die Ergebnisse des Qualitätsprozesses an Hochschulen eine stärkere Einigkeit besteht als über deren Ausgangsbedingungen (denen z.B. Studienanfänger, Studienplätze oder Stellen zugeordnet werden können).<sup>110</sup> Auf diese Systematik wird nachfolgend noch einzugehen sein.

Zunächst soll aber versucht werden, auch zu den Parametern die verfügbaren internationalen Erfahrungen in die Diskussion einzubeziehen. Leider ist dies nicht in ähnlich übersichtlicher Form wie zuvor in einer Tabelle möglich. Da die Parameter nicht vollständig für alle bereits zuvor bei der Gliederung der leistungsorientierten Mittelverteilung einbezogenen Länder verfügbar sind bzw. aufbereitet wurden und in einigen Fällen keine Angaben gemacht wurden, erscheint eine „zählbare“ Darstellung nicht angemessen. Es gibt allerdings für die Länder, zu denen Informationen verfügbar sind, eine bei den am häufigsten verwendeten Parametern relativ große Übereinstimmung. Im Wesentlichen wurden nach Leszcensky u.a. (2004: 189) für den Bereich der Lehre folgende Parameter verwendet:

#### **In ausgewählten OECD-Ländern verwendete Parameter für die Lehre:**

- Anzahl der Absolventen bzw. Abschlussprüfungen
- Anzahl der Studierenden
- Anzahl der Studienanfänger bzw. „Erstjahrestudierenden“
- Schwund
- Anzahl der erworbenen Kreditpunkte
- Anzahl internationaler Studierendenaustausche
- qualitative Indikatoren

---

<sup>110</sup> Dies würde der angestrebten Umstellung von der sogenannten ex-ante-Steuerung (mit klassischen kameralistischen Haushalten auf ex-post-Steuerung (meist mit Globalhaushalten) entsprechen (vgl. Behrens 2001: 70). Ziel- und Leistungsvereinbarungen werden als Mischform zwischen ex-ante- und ex-post-Steuerung betrachtet (vgl. Leszcensky/ Orr 2003: 7). Die Übersicht zeigt aber auch, daß einige Hochschulen bei der leistungsorientierten Mittelvergabe sehr viel stärker nach Parametern steuern, die eher Ausgangsbedingungen bzw. „Input“ zuzuordnen sind, als den Ergebnissen bzw. „Output“. Darauf wird im Abschnitt zu Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen zurückzukommen sein.

Die Parameter Absolventenzahl und Studierendenzahl werden in den meisten untersuchten Ländern verwendet; hier stimmt das Bild der untersuchten deutschen Hochschulen mit den internationalen Erfahrungen in etwa überein. Die übrigen Parameter werden in deutlich weniger Ländern verwendet: Anzahl der Studienanfänger bzw. „Erstjahrestudierenden“ lediglich in Finnland und den Niederlanden, der Schwund in den Niederlanden, die Anzahl der erworbenen Kreditpunkte in Schweden und Norwegen (dort auch die Anzahl internationaler Studierendenaustausche) sowie qualitative im Sinne von Qualität erfassenden Indikatoren (bislang nur in Ungarn, geplant jedoch auch in Großbritannien). Außerdem existieren umfassende Erhebungen zur Studienqualität und zum Absolventenverbleib in Australien,<sup>111</sup> die jedoch „nur“ mittelbar über die Studentenzahlen Einfluß auf die Mittelverteilung haben.<sup>112</sup>

Der Vollständigkeit halber sollen hier auch wie in der Übersicht der untersuchten deutschen Hochschulen die für die Forschung verwendeten Parameter leistungsorientierter Mittelvergabe aufgeführt werden: Am häufigsten werden hierfür die Drittmittel und die Anzahl der abgeschlossenen Promotionen verwendet; auch hier stimmt das Bild der untersuchten deutschen Hochschulen mit den internationalen Erfahrungen in etwa überein. In einem Land werden auch die Anzahl der Dissertationen und die als „Top-Forschungsschulen“ bezeichneten Nachwuchsforschungseinrichtungen (Niederlande) genutzt. In Australien geht in geringem Umfang auch die Anzahl der Veröffentlichungen in die Forschungsfinanzierung ein, in Großbritannien v.a. die qualitative Bewertung der Forschung (vgl. Leszcensky u.a. 2004: 189). Als ergänzende „Hintergrundinformationen“, die nicht direkt in die Bewertung einfließen, aber zur Unterstützung der Entscheidungsfindung dienen, werden in mehreren Ländern weitere Datenquellen hinzugezogen, so in Irland die Selbstevaluationsberichte, in den Niederlanden Zitationsindizes für Publikationen und in Großbritannien Angaben sieben spezieller Formulare zur Forschungsevaluation (RAE).

An der Universität Dortmund wurden als einziger der untersuchten deutschen Hochschulen zusätzlich zu den in Anlehnung an die Systematik Schröders aufgeführten Parametern eine Reihe von eher qualitativen Kriterien bzw. Indikatoren aufgeführt. Diese waren mit der weiter oben

---

<sup>111</sup> Beschreibungen hierzu finden sich in: [www.backingaustraliasfuture.gov.au/fact\\_sheets/14.htm](http://www.backingaustraliasfuture.gov.au/fact_sheets/14.htm) (Stand März 2004).

<sup>112</sup> Die Erhebungen in Australien werden ausdrücklich auch mit dem Ziel der Unterstützung der Studienanfänger bei der Hochschulwahl durchgeführt und entsprechend aufbereitet im Internet präsentiert. Es ist davon auszugehen, daß sie weit größeren Einfluß auf die Hochschulwahl haben, als bisher dies bei den bisher vorliegenden Rankings in Deutschland der Fall sein dürfte (vgl. ähnlich für UK: Orr 2001: 161). An dieser Stelle soll deshalb daran erinnert werden, daß 77% der Mittelzuweisungen an die australischen Hochschulen von der Anzahl der Studierenden abhängig sind. Außerdem wurden ab 1998 Fragen der internen Qualitätssicherung auch ein Element der jährlichen Einzelverhandlungen über Finanzen und Studierendenzahl (vgl. Leszcensky u.a. 2004: 24).

angewandten Systematik in einer Tabelle jedoch nicht übersichtlich darstellbar. Daher wird nachfolgend eine weitere Übersicht entwickelt, die versucht, auch diese stärker qualitativen Kriterien einzubeziehen. Zunächst soll aber die Bewertung der von Schröder vorgestellten Modelle dargestellt werden.

### **Mögliche Wirkungen und Nebenwirkungen der leistungsorientierten Mittelvergabe**

Im Anschluß an die Gliederung der Modelle leistungsorientierter Mittelverteilung nimmt Schröder unter Rückgriff auf die Ergebnisse seiner Interviews eine Bewertung dieser Modelle bezüglich angestrebter bzw. erwarteter Wirkungen, Probleme und Grenzen vor. Wesentliches Ergebnis dieser Bewertungen ist, daß die leistungsorientierte Mittelverteilung von den Interviewten insgesamt vorrangig als für die Steigerung der Leistungsgerechtigkeit und der Transparenz über die Verwendung der finanziellen Mittel geeignet gesehen wird, jedoch kaum für die Erhöhung der Leistungsbereitschaft von Professoren und Mitarbeitern.<sup>113</sup>

Als erwartete Wirkungen werden v.a. eine Ausweitung der Aktivitäten zur Drittmittelwerbung und der Kontakte zur Wirtschaft genannt. Für die anderen angesprochenen Aspekte (u.a. Bedeutung der Lehre, Abwanderung qualifizierter Wissenschaftler) wird mehrheitlich keine Veränderung erwartet. Da ursprünglich als Ziele der leistungsorientierten Mittelvergabe aber sowohl die Erhöhung der Leistungsbereitschaft als auch eine Anreiz- bzw. Steuerungswirkung formuliert wurden, liegen hier offenbar Probleme der untersuchten Modelle. Schröder führt dies jedoch nicht auf eine generelle Nicht-Steuerbarkeit im Hochschulbereich zurück, sondern sieht dies vielmehr als Indiz, daß die notwendigen Rahmenbedingungen bislang nicht hinreichend erfüllt sind und verweist auf ausländische Erfahrungen, die eine Steuerbarkeit belegen würden.<sup>114</sup>

Als eine problematische Rahmenbedingung für eine Anreizwirkung wird die geringe absolute Höhe der leistungsorientierten Mittelvergabe genannt. Ein weiteres Problem ist, daß häufig die Steuerungsintention unklar ist, die mit dem angewandten Verfahren verknüpft wird. Hinzu kommt mit hoher Wahrscheinlichkeit auch der Umstand, daß zum Erhebungszeitraum (2000/2001) der Schröderschen Untersuchung die Diskussionen um Zielvereinbarungen noch in den Anfängen steckten. Häufig wird hierzu auf folgendes hingewiesen: Um auch auf zukünftige

---

<sup>113</sup> Schröder (2004: 43ff.) analysiert die Bewertungen seiner Interviewten auch noch detailliert nach Fächergruppen und prozentuiert deren Antworten. Dies ist allerdings weder seinem Erhebungsverfahren (mangelnde Repräsentativität) noch der Fallzahl innerhalb der einzelnen Fächergruppen angemessen (vgl. hierzu Bortz/ Döring 2002: 613)

<sup>114</sup> In der Tat gibt es Hinweise dafür, daß bestimmte Anreizmechanismen sehr effektiv wirken. So hat England, das seine Finanzierungsmethodik mit dem Ziel, die Studierquote zu erhöhen, relativ stark auf Studierendenzahlen ausrichtete, seine Studierendenzahl trotz negativer demographischer Entwicklung zwischen 1995 und 2001 um 14% gesteigert (vgl. Lesczensky u.a. 2004: 57).

Entwicklungen Einfluß nehmen zu können, sollte die Mittelverteilung nicht allein vergangenheitsbezogen nach Formelmodellen erfolgen, sondern mit diskretionären Elementen wie z.B. Zielvereinbarungen ergänzt werden (vgl. Schröder 2004: 41).

Als positive Aspekte der untersuchten leistungsorientierten Mittelverteilungsmodelle wurden erwähnt, daß sie Verteilungskämpfe verhindern, wobei als Voraussetzung wiederum deren Transparenz genannt wird. Ein weiterer Vorzug sei, daß durch die leistungsorientierte Mittelverteilung historisch bedingte, unangemessene Ausstattungen und Budget von Fachbereichen und Instituten zumindest teilweise korrigiere und damit zu höherer Fairneß führe. Diese Aspekte wurden auch in den vorliegenden Dokumenten der Universitäten Oldenburg und Dortmund sowie von Minssen/ Wilkesmann (2003: 124) als Vorzüge genannt. Auch in der britischen Diskussion wurde dies als wichtiger Aspekt hervorgehoben (vgl. z.B. Jackson 1997: 103ff.). Außerdem nennt Schröder (2004: 50) als zumindest potentielle Vorzüge (die künftig an Bedeutung gewinnen könnten), die Steigerung des Rechenschaftsbewußtseins über die Verwendung finanzieller Mittel, die Transparenz über den Leistungsstand (die allerdings eine hochschulweite Veröffentlichung der Kennzahlen voraussetze) und die Förderung des Wettbewerbs durch Messung der Aufgabenerfüllung.

Bei letzteren als potentielle Vorzüge erwähnten Aspekten erscheint jedoch Skepsis angebracht, da diese entweder von vielen Beteiligten nicht als Vorzüge gesehen werden oder durch andere Instrumente bzw. Modelle besser erreicht werden könnten.<sup>115</sup>

Wie bei den leistungsorientierten Besoldungszulagen wird zudem argumentiert, daß viele Hochschullehrer v.a. intrinsisch motiviert wären und derart motivierte Professoren durch materielle Anreize ihre bisherige Motivation eher abbauen würden (vgl. z.B. Minssen/ Wilkesmann 2003: 117). Gerade bei dieser Studie, die die Diskussion mit empirischen Beiträgen erhellen könnte, zeigten sich jedoch leider einige methodische Probleme, die die Aussagekraft der Ergebnisse in Frage stellen.<sup>116</sup> Dennoch hat die Studie interessante Ansatzpunkte, die ggf. aufgegriffen werden könnten. Minssen/ Wilkesmann führten eine schriftliche, standardisierte Professorenbefragung<sup>117</sup> mit dem Ziel durch, die Akzeptanz und die Wirkungen der indikatori-sierten Mittelverteilung an Nordrhein-westfälischen Hochschulen zu untersuchen. Hierzu legten

---

<sup>115</sup> So wird die Schaffung von Transparenz über den Leistungsstand oder die Messung der Aufgabenerfüllung häufig als geeignete Aufgabe von Instrumenten der Leistungsbewertung wie z.B. der Evaluation angesehen, die dies differenzierter und wirkungsvoller leisten könnten (vgl. Abschnitt zur Evaluation).

<sup>116</sup> Es wurden zwar 732 Befragte aus ganz Nordrhein-Westfalen in die als Vollerhebung geplante Untersuchung einbezogen. Jedoch war die Rücklaufquote mit 12 Prozent bei den Professoren sehr gering und nach Befragten-gruppen sehr unterschiedlich (vgl. Minssen/ Wilkesmann 2003: 111). Zudem erfolgte kein Hinweis, ob die Repräsentativität der Befragung bezüglich der Grundgesamtheit anhand wichtiger, potentiell die Ergebnisse beeinflussender Merkmale der Befragten überprüft wurde (z.B. Fächerzusammensetzung). Daher ist davon auszugehen, daß dies nicht geschah und die Ergebnisse bei diesem geringen Rücklauf stark verzerrt sein können.

sie den Professoren Itemlisten vor, anhand derer die Professoren ihre Einschätzung der Mittelverteilung insgesamt sowie bezüglich einzelner positiver und negativer Wirkungen beurteilen sollten. Insgesamt beurteilte mit 49% rund die Hälfte der Professoren die indikatorisierte Mittelverteilung positiv.<sup>118</sup> Die unabhängig davon befragten Dekane und Rektoratskollegien sahen dies noch positiver. Hierbei bewerteten die befragten Professoren dies „um so positiver, je informativer und partizipativer der Einführungsprozeß verlaufen ist“ (vgl. Minssen/ Wilkesmann 2003: 113). Insgesamt fühlten sich zwei Drittel von ihnen ausführlich informiert, drei Viertel gaben an, sich an fachbereichsinternen Diskussionen beteiligt zu haben. Allerdings erfuhr auch ein Fünftel erst im Zuge ihrer Berufung von der Neuordnung der Mittelverteilung und bewerteten diese dann deutlich negativer. Aus diesen Ergebnissen könnte – Repräsentativität<sup>119</sup> oder zumindest Tendenzaussagen ermöglichende Ergebnisse vorausgesetzt – zunächst geschlossen werden, daß die Akzeptanz für eine leistungsorientierte Mittelverteilung jedenfalls in Nordrhein-Westfalen weit größer ist als bisher angenommen und zudem noch ein deutliches Potential bei nicht ausreichend informierten Professoren existiert.

Minssen/ Wilkesmann kamen jedoch zu einer anderen Interpretation, indem sie die Professoren unter Zuhilfenahme einer Faktorenanalyse einer ihrer Itemlisten bezüglich einzelner positiver und negativer Wirkungen in „Ökonomen“ und „Traditionalisten“ einteilten. Demzufolge würden sich 33% der Professoren in der Gruppe der „Ökonomen“ und 67% in der Gruppe der „Traditionalisten“ versammeln.<sup>120</sup> Gestützt auf diese Einteilung wird argumentiert, daß ein großer Teil der Professoren der Anreizwirkung sehr skeptisch gegenüberstehe und durch die leistungsorientierte Mittelvergabe kaum zu einer Verhaltensänderung zu bewegen sei. Die wenigen Aufgeschlossenen, die erreicht werden könnten, so wird weiter argumentiert, würden ohnehin schon unabhängig von leistungsorientierter Mittelvergabe nach deren Prinzipien handeln und daher solche Anreize nicht benötigen (vgl. Minssen/ Wilkesmann 2003: 119). Für die Mehrheit der Professoren seien es ganz andere Faktoren, die motivierende Wirkung erzielen: Attraktive Arbeitsbedingungen, interessante Forschungsprojekte und insbesondere die Gestaltung des Denkens junger Menschen (vgl. Minssen/ Wilkesmann 2003: 122). Der stärkste Anreiz für Professoren sei schließlich ihre Reputation bei Fachkollegen.

---

<sup>117</sup> Außerdem führten sie 12 Experteninterviews durch. Deren Ergebnisse werden aber nicht weiter ausgeführt.

<sup>118</sup> Es wird zudem auf eine Studie von Hübner/ Rau (2001) verwiesen, die fast zu denselben Ergebnissen kam.

<sup>119</sup> Der Begriff Repräsentativität wird häufig kritisiert, da er aufgrund seiner allgemeinsprachlichen Verwendung Anlaß zu Mißverständnissen geben kann (vgl. z.B. Schnell u.a. 1997: 12). Hier wird er in Anlehnung an Kromrey (1998: 259) verwendet.

<sup>120</sup> Wie genau die Autoren zu diesen Anteilen kommen, wird nicht weiter erklärt. Ebenso wenig wird erklärt, warum nicht z.B. eine Clusteranalyse zur Gruppenbildung verfolgt wurde, da diese als für das Ziel der Bildung von *Personengruppen* als wesentlich geeigneter eingeschätzt wird als die Faktorenanalyse, die das Verfahren der Wahl für die Bildung von *Variablengruppen* sei (vgl. z.B. Backhaus u.a. 2000: 329ff., Brosius 2002: 661ff.).

Diese gedankliche „Dichotomisierung“, die Minssen/ Wilkesmann vornehmen, verkennt jedoch einen wesentlichen Zusammenhang, der regelmäßig auch in der Argumentation vieler Professoren (z.B. bei Berufungs- und Bleibeverhandlungen) eine große Rolle spielt: Die Arbeitsbedingungen und Forschungsmöglichkeiten eines Professors sind keineswegs unabhängig von der finanziellen Ausstattung einer Professur bzw. eines Institutes. Und da Reputation v.a. über (innovative) Forschungsprojekte und entsprechende Publikationen erlangt wird, sind die Möglichkeiten zum Erwerb von Reputation schließlich *nicht* unabhängig von der finanziellen Ausstattung einer Professur bzw. eines Institutes. Zwar gilt dies stärker für die Ingenieur- und Naturwissenschaften und die empirischen Sozialwissenschaften als für die sogenannten „reinen Buchwissenschaften“. Doch selbst bei letzteren bedarf es einer guten Ausstattung an Büchern und Fachzeitschriften, die keineswegs grundsätzlich vorhanden ist und außerdem auf einem aktuellen Stand gehalten werden muß.

Einmal von dieser Problematik der Argumentation von Minssen/ Wilkesmann abgesehen, ist die gedankliche Übertragung der Ergebnisse von den Befragten auf die Grundgesamtheit angesichts der erwähnten methodischen Probleme bei der Erhebung und Auswertung der Daten fragwürdig. Außerdem erscheint auch die Einteilung in die beiden genannten Gruppen und das Schließen auf deren künftiges Verhalten aufgrund der wenigen vorhandenen Informationen hierzu als zumindest schwer nachvollziehbar.

Als tragfähiger erscheint jedoch (unter der Voraussetzung einer weitgehend repräsentativen Stichprobe) der bei Minssen/ Wilkesmann nur kurz gestreifte Ansatz, Professoren zur Akzeptanz sowie (wenn bereits genügend Erfahrungen hiermit vorliegen) bezüglich positiver und negativer Wirkungen leistungsorientierter Mittelvergabe zu befragen und anhand von Zusammenhangsanalysen nach Einflußfaktoren auf die Akzeptanz und die Einschätzungen (wie z.B. Grad der Informiertheit und Partizipation) zu suchen. Dies könnte ein Desiderat für weitere noch vorzunehmende empirische Untersuchungen darstellen.

Zusammenfassend kann zu Modellen leistungsorientierter Mittelverteilung formuliert werden, daß sie als geeignetes Mittel gesehen werden, Verteilungskämpfe zu verhindern und historisch bedingte, unangemessene Ausstattungen von Fachbereichen und Instituten zumindest teilweise zu korrigieren. Diese Sicht wird auch in einigen gewerkschaftlichen Publikationen zum Thema geteilt (vgl. Behrens 2001: 68). Ähnlich wie bei leistungsabhängigen Besoldungszulagen ist neben der grundsätzlichen Eignung der Parameter bzw. Leistungskriterien hierfür ein dem Hochschulbereich angemessenes und transparentes Verfahren notwendig, über deren Entwicklung die Betroffenen zudem ausreichend informiert und daran soweit wie möglich beteiligt werden sollten. Außerdem muß auch hier der Gegenstand der Leistungsbeurteilung umfassend in allen

seinen wesentlichen Dimensionen in die Leistungsbewertung einbezogen werden, um die angestrebten Ziele zu erreichen und Fehlanreize zu vermeiden. Besonders betont wird zur leistungsorientierten Mittelverteilung, daß - um auch auf zukünftige Entwicklungen Einfluß nehmen zu können – diese nicht allein vergangenheitsbezogen nach Formelmodellen erfolgen, sondern mit diskretionären Elementen wie z.B. Zielvereinbarungen ergänzt werden sollte. Insgesamt werden die Möglichkeiten, mit leistungsorientierter Mittelverteilung Qualitätsorientierung zu befördern, als derzeitig (noch) stark eingeschränkt beschrieben, aber zugleich unausgeschöpftes Potential gesehen.

### **3.5 Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen**

#### **3.5.1 Schwächen von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen als Ausgangspunkt**

Wie bereits bei der Einbeziehung des Modells der Universität Dortmund in die Übersicht zu Parametern der leistungsorientierten Mittelvergabe kurz angesprochen, wurden zusätzlich zu den in Anlehnung an die Systematik Schröders aufgeführten Parametern eine Reihe von eher qualitativen (i.S.v. stärker die Qualität erfassenden) Kriterien bzw. Indikatoren aufgeführt. Auch in Zusammenhang mit leistungsorientierten Besoldungszulagen wurde von mehreren Autoren darauf hingewiesen, daß stärker qualitative Indikatoren z.B. in Form von subjektiven Bewertungen einfließen sollten. Dies wurde v.a. vor dem Hintergrund diskutiert, daß die bisher vorhandenen und eingesetzten Anreizmodelle zu stark quantitative, z.T. auch als „objektiv“ bezeichnete Indikatoren betonen und zu wenig Aspekte der Qualität erfassen und damit zu einer Fehlsteuerung bzw. Fehlanreizen führen können, wie dies bereits von anderen Ländern berichtet wurde (vgl. z.B. Leszcensky u.a. 2004: 190ff., Orr 2001: 128). Häufig wird versucht, diese Probleme lediglich nachträglich zu kontrollieren oder transparent zu machen. Bislang eher selten (v.a. in Ländern mit längerer Erfahrung wie Großbritannien und Australien, aber auch in Ungarn) wird versucht, sie durch entsprechende Gegensteuerung bereits im (oder parallel zum) Anreizsystem möglichst gar nicht erst entstehen zu lassen. Eine Möglichkeit der Weiterentwicklung wird also in der Ergänzung von Leistungsanreizmodellen durch (qualitative oder auch als subjektiv bezeichnete) Indikatoren gesehen, die stärker Qualität und damit nach dem zugrundegelegten Begriffsverständnis immer auch subjektive Aspekte erfassen sollten.

Auf der anderen Seite werden bei Leistungs- und Qualitätsbewertungen, insbesondere bei Evaluationen, als sogenanntes „Follow-up“ zunehmend Zielvereinbarungen üblich (vgl. HRK

2001b: 10, 2002: 83ff., 2004b: 97ff.). Als Schwäche solcher für die Bewertung (und Verbesserung) der Qualität entwickelten Instrumente wird jedoch eingeschätzt, daß die Umsetzung der Zielvereinbarungen bzw. der Follow-up-Maßnahmen aufgrund der geringen Verbindlichkeit in sehr unterschiedlichem Ausmaß stattfindet, wie eine im Auftrag der HRK durchgeführte Metaevaluation ergab: „Ein großer Teil der Interviewpartner bezeichnet die Umsetzung der Ergebnisse als Schwachstelle bzw. als Problem. Es fehle eine Systematik und auf die Umsetzung werde zu wenig geachtet“ (vgl. HRK 2004b: 105).<sup>121</sup> Daraufhin befragt, ob eine Verknüpfung von Evaluationsergebnissen mit leistungsorientierter Mittelvergabe vorgenommen werden sollte, antwortete ein (angesichts der bislang großen Widerstände gegen eine solche Verknüpfung) nur geringer Anteil der Befragten, daß sie dagegen wären.<sup>122</sup> Es ergab sich sogar eine klare Mehrheit für den Fall, daß es bei Erreichung der Ziele ggf. eine Erhöhung der Mittel gäbe (und bei Nichterreichung zumindest der Status quo erhalten bliebe).<sup>123</sup>

Da diese Konstellation bei Zielvereinbarungen gewährleistet werden könnte und solche Zielvereinbarungen sowohl Bestandteil von Anreizmodellen als auch von Follow-Up-Maßnahmen zu Evaluationen darstellen, könnten sie ein konzeptionelles Bindeglied sein, daß Leistungsanreizmodelle und bereits etablierte Leistungs- und Qualitätsbewertungen an den Hochschulen zu ergänzen in der Lage ist. So subsumiert Teichler (2003: 6) „indikatorenbezogene Mittelzuweisung“ und „Individual-Bewertungen bei leistungsbezogener Besoldung“ u.a. gemeinsam mit Evaluation und Rankings in einer Liste von Verfahren multipler Qualitätsbewertung an deutschen Hochschulen. Aber auch andere (z.B. Pasternack 2004: 73) gehen davon aus, daß Evaluationen „im Zuge von formelgebundenen Mittelzuweisungen, Zielvereinbarungen, leistungsorientierten Besoldungen (...) künftig sehr viel stärker in den Hochschulen präsent sein“ werden als in der Vergangenheit. Wieder andere forderten dies sogar, wie z.B. die damalige hessische Wissenschaftsministerin Wagner (2000: 188). Betrachtet man die internationalen Erfahrungen, so zeigt sich in vielen OECD-Ländern eine Wechselwirkung zwischen leistungsbezogenen Finanzierungsmodellen und Qualitätsmaßnahmen. Daher gehen Leszczeny u.a. (2004: 1) davon

---

<sup>121</sup> In anderen Ländern wie z.B. England ist bei Feststellung von Mängeln beim Quality Audit innerhalb von drei Monaten ein Aktionsplan zu erstellen und nach spätestens 18 Monaten müssen die Probleme beseitigt sein. Anderenfalls kann dies Folgen für die Finanzierung haben (vgl. Leszczeny u.a. 2004: 65). Ähnlich sieht dies in Australien aus, wo die Fristen nur wenig länger bemessen sind (vgl. Leszczeny u.a. 2004: 26).

<sup>122</sup> Nur 40% der über 600 Befragten meinten, dies sollte in keinem Fall verknüpft werden, die Mehrheit sprach sich also dafür aus, dies unter bestimmten Umständen zu tun: 39% gaben an, dies sollte sowohl im Falle einer Erhöhung als auch im Falle einer Verringerung der Mittel geschehen (Bonus-Malus-System). 20% sprachen sich dafür aus, dies nur im Falle einer Erhöhung zu tun (Bonus-System). Lediglich 1% befürwortete dies nur im Falle einer Verringerung der Mittel (Malus-System).

<sup>123</sup> In dieser Art fand bisher eine Verknüpfung von Quality Audit und finanziellen Anreizen an den Hochschulen in Australien statt. Außerdem wurden ab 1998 Fragen der internen Qualitätssicherung Element der jährlichen Einzelverhandlungen über Finanzen und Studierendenzahl (vgl. Leszczeny u.a. 2004: 24). Das Bonus-System wird auch bei der Vergabe zusätzlicher Forschungsmittel in Irland angewendet (vgl. Leszczeny/Orr 2003: 68).

aus, daß deren internationale Entwicklung „einen deutlichen Vorsprung gegenüber dem deutschen Status quo aufweist“. Die Art und Weise sowie der Grad der Verknüpfung ist jedoch sehr unterschiedlich gestaltet, beispielsweise ob Anreizsysteme und Qualitätssicherungsmaßnahmen eher nebeneinander existieren oder ob sie sich sinnvoll ergänzen (vgl. Leszcensky u.a. 2004: 187ff.).

Eine solche – mehr oder weniger enge – Verknüpfung ist also eigentlich nichts grundsätzlich Neues. Orr (2001: 92) sieht in der Verknüpfung der leistungsorientierten Finanzierung und der Bewertung von Qualität sogar ein Hauptmerkmal des „New Public Management“ an den Hochschulen in Großbritannien. Bereits weiter vorn erwähnte Qualitätsmanagementmodelle praktizieren dies in anderen Bereichen auch in Deutschland mit mehr oder weniger großen Erfolgen. Im deutschen Hochschulbereich gibt es allerdings bislang nur sehr wenige Erfahrungen mit solchen Modellen.<sup>124</sup> Ohnehin ist – wie in diesem Zusammenhang bereits angesprochen – darauf hinzuweisen, daß diese Erfahrungen nicht ohne weiteres übertragbar sind, sondern nach angemessenen und ggf. für den Hochschulbereich spezifisch zu entwickelnden Lösungen gesucht werden muß.

Daher soll, wenn es um die angestrebte stärkere Einbeziehung qualitativer Kriterien und Indikatoren geht, nachfolgend zunächst auf theoretische Überlegungen aus dem deutschsprachigen Hochschulbereich zurückgegriffen werden. Überlegungen, wie man umfassend alle wesentlichen quantitativen und qualitativen Aspekte der Lehre erfassen könnte, existieren bereits seit längerem, z.B. im Zusammenhang mit Verfahren zur Erstellung von Lehrberichten (vgl. Webler u.a. 1993) bzw. im Zusammenhang mit der Evaluation von Hochschulleistungen im Rahmen eines Performance Measurements (vgl. Schenker-Wicki 1996). Bisher wurden diese Anregungen – u.a. weil die Erhebung qualitativer Indikatoren als schwierig gilt – nur z.T. aufgegriffen (vgl. z.B. Leszcensky u.a. 2004: 12). International gibt es jedoch bereits Erfahrungen auch mit Absolventen- und Studentenbefragungen z.B. in Australien (vgl. Leszcensky u.a. 2004: 26). In England gibt es ebenfalls Absolventenstudien sowie Erhebungen von Abbrecherquoten, die als Leistungsindikator verwendet werden und an einer Studentenbefragung als weiteres Element des nationalen Systems der Qualitätssicherung wird gearbeitet (vgl. Leszcensky u.a. 2004: 66).

---

<sup>124</sup> Leszcensky/ Orr (2003: 35) berichten von ersten Erfahrungen mit der Transformation qualitativer Bewertungen zu Notenskalen in 5 von 12 veröffentlichten Endberichten zur Fächerevaluation an niedersächsischen Hochschulen. Nach Leszcensky/ Orr (2003: 8) müsse zunächst die Frage einer solchen Transformation in numerische Skalen gelöst werden, wenn Qualität Bestandteil von Zielvereinbarungen sein soll und die Erreichung von vereinbarten Qualitätszielen überprüft werden soll. Dies könne neben der Transformation von Expertenurteilen (vgl. zu Erfahrungen für den Bereich der Lehre Orr 2001: 183) aber auch durch studentische Urteile im Rahmen der Lehrberichterstattung erfolgen (vgl. Leszcensky/ Orr 2003: 8), die bereits in Form von Notenskalen vorliegen.

Um zu verdeutlichen, welche Indikatoren im deutschsprachigen Raum bereits seit längerem zur Erfassung auch qualitativer Aspekte der Lehre diskutiert werden, wird nachfolgend eine Gegenüberstellung von theoretischen Überlegungen zu Indikatoren und von Modellen vorgenommen, die teilweise bereits Indikatoren einbezogen oder diskutierten, die Qualitätsaspekte erfassen sollen. Dies geschieht am Beispiel dreier Hochschulen: Universität Oldenburg, Universität Dortmund und TU Dresden.<sup>125</sup> Die Gliederung der Indikatoren in der Übersicht erfolgt angelehnt an Qualitätsmodelle (vgl. Dubs 1998: 34, Zurwehme 2000: 40ff., Teichler 2003: 5), bzw. Wirkungsmodelle (vgl. Orr 2001: 97, Bayer 2004: 171ff.) im Bildungswesen in folgender Systematik: Ausgangsbedingungen („Input“), Prozeß (Prozeßqualität) und Ergebnisse (Output) bzw. Wirkungen (Outcome).

---

<sup>125</sup> Zu den von Schröder einbezogenen vier weiteren Hochschulen liegen keine Informationen hierzu vor, daher können sie hier nicht aufgenommen werden.

### 3.5.2 Gegenüberstellung theoretischer Überlegungen und eingesetzter bzw. diskutierter Indikatoren der leistungsorientierten

#### Mittelverteilung an ausgewählten Hochschulen

	<i>Webler 1993</i>	<i>Schenker- Wicki 1996</i>	<i>Olden- burg 2000</i>	<i>Dort- mund 2000</i>	<i>TU Dresden 2001</i>	<i>Erläuterungen, Beispiele in bisherigen Überlegungen/ Modellen:</i>
<b>Ausgangsbedingungen/ "Input":</b>						
Studentenzahl	x		x	(x)	x	Anzahl der immatrikulierten Studenten (z.B. im grundständigen Präsenzstudium)
Anzahl Studenten in der Regelstudienzeit			x		x	o.g. Studentenzahl in der Regelstudienzeit (in der Prüfungsordnung des Studienganges festgelegte Regelstudienzeit)
Anfängerzahl	x			(x)	x	Anzahl der Studienanfänger (im grundständigen Präsenzstudium)
Attraktivität	x	x				Bewerber pro Studienplatz (im grundständigen Präsenzstudium)
Personalstärke und -Struktur	x	x				Anzahl der Professoren (evtl. wiss. Mitarbeiter/ stud. Tutoren separat aufführen)
Anfängerqualität	(x)	x				Durchschnitt der Abiturnoten (evtl. berufliche Vorerfahrungen, Praktika usw.)
Anfängermotivation	(x)					Stärke des fachlichen Interesses bei der Studienfachwahl
Internationalität	x		x			Anteil ausländischer Studenten
Betreuung(spotential)	x	x			x	Betreuungsrelation (Anzahl der Studenten pro Professor/ u. evtl. wiss. Mitarbeiter)
<b>Prozeß:</b>						
Qualität der Studienberatung	x			x		Freundlichkeit, inhaltlich zutreffende Information zu Studienangebot/ Berufsfeldern usw.
Qualität der Studienziele	x			x		klare Formulierung der Studienziele, mögl. Berufsfelder, Wahlmögl./ Studienpläne
Lehr- und Studienqualität (1)	x	x				"Studierzufriedenheit" bei Studenten, z.B. stud. Lehrbewertung/ Studienbedingungen
Lehr- und Studienqualität (2)	x					Absolventenbefragung zur rückblickenden Bewertung von Lehre und Studium
Lehr- und Studienqualität (3)						Einschätzungen durch Mitarbeiterbefragung/ Lehrendenbefragung
speziell: Beratungsqualität im Studium	x			x		Ansprechbarkeit der Lehrenden,
Prüfungen	x			x		Wiederholungsmöglichkeit nach angemessener Zeit, Transparenz über Anforderungen
Internationalität		x		x		fremdsprachige Lehrveranst., Anteil ausländischer Studenten/Dozenten, Austauschrate
Interdisziplinarität		x		x		Dienstleistungsverflechtungen, Lehrexporte, gemeinsame Studien-/ Forschungsprojekte
Gleichstellung			x	x		Frauenanteil ggf. relativiert mit Bundesschnitt, Anteil an Studienanfängern o. ä.)
Studierbarkeit	x	(x)		x		Anteil der Studenten in der Regelstudienzeit (in %)
interne Qualitätssicherung	x	x		x	(x)	Vorhandensein stud. Lehrbewertung und/ oder anderer Qualitätssicherungsprojekte

(Fortsetzung der Tabelle auf der nächsten Seite)

(Fortsetzung der Tabelle von der vorhergehenden Seite)

	<i>Webler 1993</i>	<i>Schenker-Wicki 1996</i>	<i>Oldenburg 2000</i>	<i>Dortmund 2000</i>	<i>TU Dresden 2001</i>	<i>Erläuterungen, Beispiele in bisherigen Überlegungen/ Modellen:</i>
<b>Ergebnisse/ "Output":</b>						
Anzahl der Zwischen-/ Hauptprüfungen			(x)		x	Fallzahlen (tats. Anzahl abgenommener Prüfungen mit Teilprüfungen, Wiederholungen usw.)
Absolventenzahl	x	x	x	(x)	x	Anzahl der Absolventen/ Promovierten
Abschlußnoten der Absolventen		x				
Absolventenquote	x	x				Absolventenquote (Verhältnis Absolventenzahl zu Anfängerzahl vor 5 Jahren, Regelstudienzeit o.ä.)
Studiendauer	x	x	x		x	Mittel/ Median Fachsemester oder Anteil derer, die Studium in Regelstudienzeit abschlossen
Effizienz und Effektivität	(x)				x	Kosten pro ausgebildeten Absolventen, Zielerreichungsgrad
<b>Wirkungen/ "Outcome":</b>						
Erwerbsquote/ Beschäftigungsgrad	x	x	x		x	Anteil Absolventen in regulärer Erwerbstätigkeit (z.B. 1 Jahr nach Studienabschluß)
Ausbildungsadäquanz	x				x	z.B. berufliche Positionen der Absolventen entsprechend Studienabschluß
Einkommen	x				x	Einkommen der Absolventen (z.B. 1 Jahr nach Studienabschluß)
berufliche Zufriedenheit	(x)					berufliche Zufriedenheit

(x) = angesprochen, aber nicht weiter konkretisiert/ noch nicht einbezogen

Quelle: eigene Darstellung

Nachdem Möglichkeiten der Ergänzung von Modellen leistungsorientierter Mittelvergabe mit Überlegungen zu Indikatoren aus Berichtssystemen wie z.B. dem Lehrberichtsverfahren und der Evaluation von Hochschulleistungen betrachtet wurden, soll nun auch noch betrachtet werden, welche Möglichkeiten sich ggf. aus Überlegungen zu Indikatoren aus sogenannten Rankings bzw. vergleichenden Studienführern ergeben. Begründet werden kann die Einbeziehung dieser Überlegungen zu Indikatoren einerseits mit den relativ hohen methodischen Ansprüchen, den einige Verfasser solcher vergleichenden Bewertungen daran stellen (vgl. Berghoff u.a. 2004: 4) und deren weitgehende Erfüllung inzwischen vielerorts zugestanden wird, andererseits aber auch damit, daß sie von einigen Hochschulen als Möglichkeit der Indikatorengewinnung diskutiert wurden, hier also eine gewisse Akzeptanz vermutet werden kann (vgl. Abschnitt zu Rankings in dieser Arbeit). Bedingung für die Einbeziehung von Rankings in nachfolgende Übersicht war, daß diese sowohl die objektive als auch die subjektive Dimension erfaßten, um den multidimensionalen Charakter von Qualität abzubilden (vgl. Abschnitt zur Entwicklung des Qualitätsbegriffes in dieser Arbeit). Internationale Rankings, z.B. Times Higher Education Supplement (November 2004) oder U.S. News (Januar 2005), wurden hier aufgrund der anderen Ausgangslage nicht einbezogen.

### 3.5.3 Indikatoren zur Leistungs- bzw. Qualitätsbewertung in mehrdimensionalen Rankings/ vergleichenden Studienführern

	<i>Hornbostel/ Daniel 1996</i>	<i>Eberle u.a. 1997</i>	<i>Daniel 1999</i>	<i>Kremp- kow u.a. 2000</i>	<i>Berghoff u. a. 2004</i>	<i>Erläuterungen, Beispiele:</i>
<b>Ausgangsbedingungen/ Input":</b>						
Studentenzahl		x		x	x	Anzahl der immatrikulierten Studenten (im grundständigen Präsenzstudium, evtl. andere Studienformen separat))
Anfängerzahl		x		x	x	Anzahl der Studienanfänger (im grundständigen Präsenzstudium, evtl. andere Studienformen separat))
Attraktivität		x		x		Bewerber pro Studienplatz (im grundständigen Präsenzstudium, evtl. andere Studienformen separat))
Anzahl PC-Arbeitsplätze					x	Anzahl der PC-Arbeitsplätze für den Fachbereich (anteilig je Student)
Personalstärke und -Struktur		x			x	Anzahl der Professoren (evtl. wiss. / nichtwiss. Mitarbeiter und stud. Tutoren)
Anfängerqualität						Durchschnitt der Abiturnoten, evtl. nur in bestimmten Fächern, evtl. berufliche Vorerfahrungen (Berufsausbildung), Praktika und/ oder soziales Engagement)
Anfängermotivation						Stärke des fachlichen Interesses bei der Studienfachwahl
Internationalität		x			x	Anteil ausländischer Studenten
Betreuung		x	x	x	(x)	Betreuungsrelation (Anzahl Studenten pro Prof./ u. evtl. wiss. Mitarbeiter)

<b>Prozeß:</b>						
Überfüllung	x		x			Überfüllung von Lehrveranstaltungen, Skala von 1 = "nur selten" bis 6 = "sehr häufig"
Ausfall von Lehrveranstaltg.	x					Skala von 1 = "nur sehr selten" bis 6 = "sehr oft"
Platz in Pflichtveranstaltungen	x				(x)	Skala von 1 = "sehr häufig" bis 6 = "nur selten"
Qualität der Studienberatung	x				x	Öffnungszeiten, Freundlichkeit, inhaltlich zutreffende Information zu Studienangebot/Berufsfeldern, Verhindern falscher Erwartungshaltungen, usw.)
Qualität der Studienziele				x		klare Formulierung der Studienziele, mögl. Berufsfelder, Wahlmögl. Im Studium bis hin zu Studienplänen o. ä.
Einschätzung Lehrpersonal	x	x	x	x		Vorbereitung der Dozenten auf Lehrveranstaltungen, Didaktik (Verständlichkeit) Skala von 1 = "sehr viele" bis 6 = "nur wenige"
Lehrinhalte, Lehrangebot	x	x	x	x	x,	Breite des Lehrangebotes, Forschungsbezug
Räume	x	x	x	x	x	Gestaltung Unterrichts- und Aufenthaltsbereich
Bibliothek	x	x	x	x	x	Einschätzung Ausleihmöglichkeiten, Arbeitsplätze
Computer	x	x	x		x	Einschätzung Arbeitsplätze
Examensvorbereitung	x	x		x	x	allgemeine Bewertung bzw. Besprechung der Ergebnisse von Tests, Klausuren und Hausarbeiten so daß die Studierenden genau wissen, warum sie mehr oder weniger gut abschneiden
Beratungsqualität im Studium	x				x	Ansprechbarkeit der Lehrenden, individuelle Beratung und Betreuung
Praktika					x	Anzahl, Wochen oder Stunden
Prüfungen				(x)		Prüfungs-Wiederholungsmöglichkeiten nach angemessener Zeit, Transparenz über Prüfungsanforderungen usw.
Zu-/ Abwanderung/ Schwund			x			Schwundquote bzw. Zuwachs an Studenten in Prozent
Internationalität					x	fremdsprachige LV, Austauschrate = Anteil ausländischer Studenten/Dozenten und Anteil eigener Stud./Dozenten im Ausland, Doppelabschlüsse, Credit-Point-System
Interdisziplinarität					(x)	Dienstleistungsverflechtungen, Lehrexporte, gemeinsame Studien-/Forschungsprojekte o. ä.)
Gleichstellung				x	x	Frauenanteil bzw. Männeranteil, ggf. relativiert mit Bundesschnitt, Anteil an Studienanfängern o. ä.)
Studierbarkeit					(x)	Anteil der Studenten in der Regelstudienzeit
rollstuhlgerecht		x		x		jeweils bezogen auf den Fachbereich/ Studiengang: ja /nein
interne Qualitätssicherung					x	Vorhandensein stud. Lehrbewertung oder Qualitätssicherungsprojekte
externe Evaluationen				(x)	x	(peer Reviews)
Absolventenbefragungen					x	Befragungen der Absolventen v. a. zum beruflichen Verbleib und zur rückblickenden Bewertung des Studiums
Arbeitsmarktbezug	x				x	versch. Aspekte, s. Frage 24 im CHE-Fachbereichs-Fragebogen (2001)

<b>Ergebnisse/ „Output“</b>						
Absolventenzahl		x		x	x	Anzahl der Absolventen
Notenschnitt				(x)	x	Notenschnitt der Abschlußprüfungen
Absolventen/ Erfolgsquote				x		Absolventenquote (Verhältnis Absolventenzahl zu Anfängerzahl vor Regelstudienzeit oder mittlerer Fachstudiendauer)
Studiendauer	x	x	x		x	Median Fachsemester oder Anteil derer, die Studium in Regelstudienzeit abschlossen
Anzahl Promotionen		x			x	Anzahl der Promotionen absolut am Fachbereich oder pro Professor
Effektivität bzw. Effizienz						Kosten pro ausgebildeten Absolventen bzw. Zielerreichungsgrad
<b>Wirkungen/ „Outcome“<sup>126</sup></b>						
Erwerbsquote						Anteil erwerbstätiger Absolventen/ reguläre Erwerbstätigkeiten (z.B. 1 Jahr nach Studienabschluß, aber nur diejenigen, die Stelle gesucht haben)
Ausbildungsadäquanz						berufliche Positionen der Absolventen
beruflicher Erfolg, Einkommen						Einkommen der Absolventen (z.B. 1 Jahr nach Studienabschluß, aber besser langfristig untersuchen)
berufliche Zufriedenheit						berufliche Zufriedenheit

<sup>126</sup> Zu (längerfristigen) Wirkungen des Hochschulstudiums wurden bisher in Deutschland in komplexen Rankings keine entsprechenden Indikatoren veröffentlicht. Das CHE plant jedoch für die Zukunft die Einbeziehung der Ergebnisse von Absolventenstudien. In einigen OECD-Ländern wie Großbritannien und Australien wurden diese bereits einbezogen.

Ähnlich verhält es sich mit der Anfängerqualität, die in Deutschland zwar häufig (meist in Form der Abiturnote) als Ausgangsbedingung angesprochen, jedoch bislang nicht einbezogen wurde. Beispielsweise in Großbritannien wird dies in Rankings regelmäßig mit erhoben und veröffentlicht (vgl. Orr 2001:227).

### 3.5.4 Fazit

Zusammenfassend kann zu Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen zunächst festgestellt werden, daß relativ häufig der Bedarf einer Weiterentwicklung von Anreizmodellen in Richtung qualitativer Kriterien bzw. subjektiver Indikatoren formuliert wird. Es gibt hierzu jedoch bislang lediglich einzelne konzeptionelle (vgl. Webler u.a. 1993, Schenker-Wicki 1996) und empirische Vorarbeiten (wobei diese – wie die in die Diskussion einbezogenen Rankings – z.T. aus etwas anderen Zusammenhängen stammen und aus verschiedenen Gründen bisher eher selten in diesem Zusammenhang diskutiert wurden).

Eine andere Möglichkeit der Weiterentwicklung und Ergänzung wird in der Schaffung einer stärkeren Verbindlichkeit von Maßnahmen als „Follow-up“ bzw. zur Umsetzung der Ergebnisse von Evaluationen und Qualitätsbewertungen gesehen, die z.B. über Zielvereinbarungen erfolgen könnte (vgl. Abschnitt 3.2.1 in dieser Arbeit).

Nach Einschätzung des Verfassers liegt in der Verbindung dieser beiden bislang meist unverbunden betrachteten Weiterentwicklungsmöglichkeiten ein Potential, die jeweils als Schwächen angesehenen Aspekte von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen zu mindern. So könnten die Anreizmodelle durch die Nutzung von Erfahrungen aus Qualitätsbewertungen durch Evaluationen u.U. stärker auch Qualität erfassen, die Umsetzung der Ergebnisse von Evaluationen könnte umgekehrt durch Verknüpfung mit Anreizen verbindlicher gestaltet werden. Pellert (1999: 292ff.) hat in einer umfassenden Organisationsanalyse der (österreichischen) Universitäten hierzu bereits formuliert: „Es mangelt sowohl auf der Ebene des Hochschulsystems als auch auf der Ebene der einzelnen Universität an entscheidungsrelevantem Wissen. Hierzu kann und soll Evaluation einen Beitrag leisten. (...) Daher sollte es sich die Wissenschaft zur Aufgabe machen, sich in Evaluationsverfahren zu engagieren und die Qualitätskriterien selbst weiterzuentwickeln“.

Dabei können die vorhandenen Vorarbeiten sicherlich nicht ohne weiteres übernommen werden und so wäre zunächst auszuloten, wie realistisch die genannte Einschätzung ist bzw. welche Bedingungen ggf. erfüllt sein müssten, um eine solche Verbindung bislang unverbundener Ansätze möglich und sinnvoll werden zu lassen. Da ein solches Wissen nicht bei allen Mitgliedern der Hochschule, sondern vor allem bei einigen „Experten“ zu finden ist, bietet sich hier die Durchführung von Experteninterviews an (vgl. auch Schröder 2004,

Minssen/ Wilkesmann 2003, Orr 2001). Hierbei wäre der Schwerpunkt auf Kriterien bzw. Indikatoren und Anreizmodelle zu legen, da zu dieser Thematik bisher in Deutschland die wenigsten Erfahrungen vorhanden sind.

Bevor jedoch im folgenden Abschnitt näher auf Anlage, Durchführung und schließlich Ergebnisse der Experteninterviews eingegangen wird, soll noch einmal zusammengefasst werden, welche Aspekte in der Literatur zur Leistungs- und Qualitätsbewertung, Qualitätsmanagement- und Leistungsanreizmodellen angesprochen wurden, die für die Experteninterviews relevant erscheinen. Hierbei ist zu beachten, daß die Experteninterviews bereits im Verlauf der Jahre 2002 und 2003 konzipiert und durchgeführt wurden und daher hierbei die in diesen (mit dem Fazit abzuschließenden) Abschnitt eingearbeitete aktuellste Literatur in die Experteninterviews nicht mehr einfließen konnte.

Da 2002 die Reform des Hochschuldienstrechts und dabei insbesondere die Professorenbesoldungsreform aktuell war, wurde es als sinnvoll angesehen, die Interviews mit der Bitte um eine allgemeine Einschätzung der bislang diskutierten Reformvorhaben einzuleiten, wobei neben der Professorenbesoldung auch Leistungsbewertungs- und anreizverfahren auf Studiengangs-, Fakultäts- und Hochschulebene angesprochen werden sollten.

Bevor in den Interviews auf einzelne Modelle eingegangen werden könnte, sollte – da in der Literatur vielfach deren Bedeutung betont wurde – zunächst erfragt werden, welche Kriterien und Verfahren als angemessen bzw. geeignet gesehen werden. Im Sinne der in der Literatur geforderten umfassenden Einbeziehung aller wesentlichen Dimensionen des zu bewertenden Gegenstandes sollte anschließend – sofern von den Gesprächspartnern noch nicht selbst angesprochen – konkret nach Möglichkeiten und Bedingungen der Einbeziehung subjektiver Bewertungen wie Studenten- und Absolventenbefragungen als Kriterien für die Erfassung von Qualität gefragt werden.

Ein weiterer Aspekt, der Leistungsanreizmodelle allgemein betrifft und in den vorliegenden Publikationen mehrfach erwähnt wurde, ist die angemessene Höhe des Leistungsanteils sowie die Gewichtung der Aufgabenbereiche (innerhalb des Leistungsanteils).

Dann sollte auf die unterschiedlichen Arten von Leistungsanreizmodellen eingegangen werden, wobei die Frageformulierung bzw. der „Stimulus“ zunächst möglichst allgemein zu halten wäre, um das Gespräch nicht von vornherein auf spezielle Arten von Leistungsanreizmodellen zu verengen. Letztlich sollte, wenn dies von den Gesprächspartnern nicht von sich heraus angesprochen wird, eine Einschätzung zur Problematik direkter Kopplung von Leistungsbewertung und -anreizen (z.B. mittels formel-

gebundener Vergabe) vs. diskretionärer Vergabe (z.B. mittels Zielvereinbarungen) bzw. zu Alternativen erfragt werden. Und schließlich soll – möglichst erst zum Schluß, da dies wohl der sensibelste Aspekt ist – eine Einschätzung zur Größe und Art der „Spreizung“<sup>127</sup> von Zulagen erfolgen, wobei die Einteilung in Leistungsgruppen bzw. Leistungsstufen oder alternative Formen der Differenzierung thematisiert werden können.

Durch die Experteninterviews zu den genannten Themenbereichen – so die Hoffnung des Verfassers – könnten über die zur Verfügung stehende Literatur hinausgehende Erkenntnisse zu Möglichkeiten und Grenzen der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen gewonnen werden.

---

<sup>127</sup> Der Begriff der „Spreizung“ soll hier nicht nur für (Besoldungs-)Zulagen im Sinne einer Lohnspreizung verwendet werden, sondern auch die Spreizung von Leistungsstufen wie in der britischen leistungsorientierten Mittelverteilung für die Forschung erfassen. Dort erhält die höchste Stufe der Exzellenz (5\*) etwa neunmal so viele Forschungsmittel wie die Hochschulen mit der letzten noch „finanzierungsfähigen“ Leistungsstufe (3a) von insgesamt 7 Leistungsstufen (vgl. Lesczensky u.a. 2004: 60). Dies wird jedoch kritisch gesehen und eine Reform des Verfahrens diskutiert, da damit die Chancen für einen „Wiederaufstieg“ oder überhaupt auf einen „Aufstieg“ sehr begrenzt sind (vgl. Orr 2001, Ahrens 2000: 182ff.).

## 4. Experteninterviews zu Leistungsbewertung und Leistungsanreizen im Hochschulbereich

*Leistungsbewertungen und Leistungsanreize werden in der deutschen Hochschulpolitik bereits seit einiger Zeit diskutiert. Wie dies jedoch von denjenigen eingeschätzt wird, die für deren konkrete Einführung an den Hochschulen (mit)verantwortlich sind und was die Gründe für deren Einschätzungen sind, wurde bislang eher selten untersucht. Hierzu mittels Experteninterviews Erkenntnisse zu sammeln und einen Diskussionsbeitrag zu leisten, ist das Anliegen dieses Abschnittes. Die folgenden Ausführungen sind in fünf Unterabschnitte gegliedert: Zunächst wird vorgestellt, welche Ziele den Experteninterviews zugrunde liegen. Danach wird das methodische Vorgehen und die Durchführung der Interviews erläutert, um das Zustandekommen der Ergebnisse transparent zu machen. Dies erscheint notwendig, da Experteninterviews zwar häufig durchgeführt, aber nach wie vor nicht sehr häufig methodisch reflektiert werden. Den vierten Schwerpunkt der Ausführungen bildet die Auswertung der Experteninterviews bezüglich der von den Gesprächspartnern dargelegten Einschätzungen von Leistungsbewertung und -anreizen. Abschließend erfolgt eine zusammenfassende Darstellung ausgewählter Ergebnisse, in der diese zu zentralen Thesen verdichtet werden.*

Gliederung:

- 4. Experteninterviews zu Leistungsbewertung und Leistungsanreizen im Hochschulbereich
  - 4.1 Ziele der Experteninterviews im Forschungsvorhaben
  - 4.2 Methodisches Vorgehen
  - 4.3 Durchführung der Experteninterviews
  - 4.4 Auswertung der Experteninterviews
    - 4.4.1 Arbeitsschritte der Auswertung
    - 4.4.2 Thematischer Vergleich
    - 4.4.3 Soziologische Konzeptualisierung
  - 4.5 Fazit

## 4.1 Ziele der Experteninterviews im Forschungsvorhaben

Ziel des Forschungsvorhabens insgesamt ist es, bei der Umgestaltung des Hochschulwesens gesammelte und v.a. bisher wenig beachtete Erfahrungen und Ergebnisse aufzugreifen. Um die Experteninterviews besser in den Gesamtzusammenhang des Forschungsvorhabens einordnen zu können, werden dessen Schwerpunkte noch einmal kurz genannt:

- Erstens sollte ein Überblick über die Konzepte und Modelle geschaffen werden, auf deren Grundlage die (das Thema dieser Arbeit betreffenden) Veränderungsprozesse in den verschiedenen Ebenen des Hochschulsystems stattfanden bzw. stattfinden (siehe vorhergehende Kapitel).
- Zweitens sollen Einschätzungen derjenigen einbezogen werden, die direkt vor Ort an den Hochschulen in die Einführung von Leistungsbewertungen und Anreizen involviert sind (Aufgabe der Experteninterviews in diesem Kapitel).
- Drittens sollen für Anreizsysteme verwendete bzw. diskutierte Leistungskriterien bzw. Indikatoren untersucht und anhand empirischer Analysen deren Eignung diskutiert werden (Aufgabe des nachfolgenden fünften Kapitels).
- Viertens soll eine Diskussion der Bedingungen erfolgreicher Institutionalisierung von Leistungsanreizen erfolgen (Aufgabe des sechsten Kapitels, wo die Diskussion anhand von Empfehlungen aus der Literatur, eigener empirischer Analysen und Erfahrungen anderer Länder zusammengeführt werden soll).

Während der Überblick über und die Einordnung der Konzepte und Modelle vor allem anhand wissenschaftlicher Literatur und schriftlicher Unterlagen aus der Hochschulpolitik (z.B. sogenannter "white papers") erfolgte, sollte die Diskussion der bei der Einführung solcher Konzepte und Modelle an den Hochschulen relevanten Aspekte anhand von Experteninterviews vertieft werden. Außerdem wird bei diesem zweiten Schwerpunkt mit Hilfe der Experteninterviews die (Einschätzung der) Akzeptanz der Kriterien bei den Experten eruiert. Für den dritten Schwerpunkt sind die Experteninterviews ebenfalls von Bedeutung, da hier z.B. bestimmte Voraussetzungen oder Bedenken, die von Experten geäußert werden, in die Diskussion der Kriterien einbezogen werden könnten. Dies soll jedoch hier nicht dargestellt werden, sondern im entspr. Teil der Forschungsarbeit.

Neben der Einordnung und vergleichenden Diskussion der Konzepte war die Einschätzung der Akzeptanz von Bedeutung, weil sich gezeigt hat, daß neben der Zuverlässigkeit und Validität von Leistungskriterien für die Einführung von Leistungsbewertungen (und deren

Kopplung mit Anreizsystemen) vor allem die *Akzeptanz bei den zu Bewertenden* (bzw. deren Einschätzung durch Experten) von großer Bedeutung ist. In den Interviews sollten deshalb insbesondere Personen zu Wort kommen, die entweder diese Systeme (mit) einführen oder die als Vertreter der (von der Einführung) Betroffenen agieren oder die beratend bzw. durch ihre wissenschaftliche Expertise bei der Einführung eine Rolle spielen. Diese Personen wirken ebenfalls direkt auf den Prozeß ein und verfügen über das zur Bearbeitung der Forschungsfrage relevante Wissen. Durch die relativ breite Auswahl soll versucht werden, bei den im Rahmen des Forschungsvorhabens maximal zu leistenden etwa zehn Interviews die notwendige Vielfalt der Perspektiven abzusichern. Um die Akzeptanz der Kriterien bei den zu Bewertenden und die Auswirkungen auf das Arbeitsklima einschätzen zu können und evt. auch für Experten unvermutete neue Informationen zu erlangen<sup>128</sup>, wurden aber nicht nur die Vertreter der Betroffenen an den ausgewählten Hochschulen, sondern zusätzlich auch diese selbst direkt befragt.<sup>129</sup> Hiermit wurden jedoch weitere Ziele verfolgt, daher wurden sie mit einer anderen Methode befragt, welche in einem gesonderten Abschnitt hierzu ausführlicher zu erläutern ist.

## 4.2 Methodisches Vorgehen

Zum methodischen Vorgehen soll entsprechend den Zielen der Untersuchung zunächst die Auswahl der einzubeziehenden *Personengruppen* (Wer kommt als „Experte“ in Frage?), dann die methodische Einordnung der Experteninterviews in das Forschungsvorhaben und schließlich die Auswahl geeigneter *Institutionen* (z.B. Hochschulen) und konkreter *Personen* erfolgen.

### Wer kommt als „Experte“ in Frage? – Auswahl der Personengruppen

Bei der Auswahl der einzubeziehenden Personengruppen für die Experteninterviews war Grundprinzip, daß insbesondere Vertreter der Hochschulleitungen bzw. mit der Einführung von Leistungsbewertungssystemen befaßte Personen der Hochschulen, Ministerien, hochschulübergreifender (wissenschaftlicher) Institutionen und (gewerkschaftliche) Vertreter der

---

<sup>128</sup> Hochschullehrerbefragungen (Vgl. Kopp/Weiß 1995) ergaben auch für Hochschulforscher z. T. unvermutete Ergebnisse.

<sup>129</sup> Eine direkte Befragung wird auch bei Mitarbeiterbefragungen in Unternehmen als unumgänglich angesehen. Die tatsächliche Akzeptanz von Bewertungskriterien wäre nicht allein durch Befragung von Arbeitgeber- und Gewerkschaftsvertretern als „Experten“ feststellbar. (Vgl. z.B. Bartel-Lingg 1996)

Personalräte zu Wort kommen sollten.<sup>130</sup> Letztere wurden, obwohl ihr Haupttätigkeitsfeld nicht unbedingt im Bereich von Leistungsbewertungssystemen liegen mußte, wegen ihrer Mitsprache- und Multiplikatorfunktion für die Betroffenen als bedeutsam eingeschätzt.<sup>131</sup>

Neben den Experteninterviews wurde zusätzlich eine repräsentative Zufallsstichprobe von Lehrenden an den ausgewählten Hochschulen in Form standardisierter schriftlicher Befragungen mit überwiegend geschlossenen Antwortvorgaben direkt befragt, um die Akzeptanz der Kriterien bei den Betroffenen und evtl. Auswirkungen auf das Arbeitsklima abschätzen zu können, wobei eine weitgehende Vergleichbarkeit der Ergebnisse untereinander zu gewährleisten war.<sup>132</sup> Diese Methode der Befragung birgt jedoch generell die Gefahr einer geringeren Tiefenschärfe in sich. Experteninterviews und standardisierte Befragungen wurden daher mit dem Ziel einer gegenseitigen Ergänzung eingesetzt. Dabei können Experteninterviews bei relativ freiem Erzählfluß eher die Nähe zum Untersuchungsgegenstand gewährleisten und bleiben durch Leitfaden bzw. gleichlautende "Stimuli" dennoch grundlegend vergleichbar.

Von einigen Autoren wird eine Durchführung von sich gegenseitig ergänzenden Untersuchungen auch als Methodentriangulation bezeichnet, wenn dadurch letztlich überprüft soll, ob eine Interpretation „die Konfrontation mit einer Serie komplementärer Testmethoden übersteht“ (vgl. Schröder-Lenzen 2003: 107). Beim Versuch einer solchen Methodentriangulation stellt sich jedoch immer die Frage, welche der Methoden validiert werden soll und welche jeweils als bereits valide gilt. Als klassisches Beispiel einer gelungenen Kombination wird häufig die Studie zu den Arbeitslosen von Marienthal angeführt (vgl. z.B. Engler 2003: 119), wo quantitative Datenanalysen u.a. mit qualitativen Beobachtungen und Gesprächen kombiniert und ergänzt wurden – ohne daß jedoch eine Methode zwingend als „valider“ angesehen wird als die andere. Jahoda u.a. (1960: 15) sprechen z.B. in Bezug auf ihre Untersuchungen zum Gesundheitszustand der Kinder von Arbeitslosen von „objektiven Indikatoren“ wie Gesundheitsstatistiken, (die sie aus Befragungen der Ärzte gewannen) und „subjektiven Äußerungen“, die in Einzel-Interviews gesammelt wurden.

---

<sup>130</sup> Vgl. auch Meuser/ Nagel 1994: 444, die Personalräte z.B. bei der Personalauswahl als Experten einordnen.

<sup>131</sup> Die Definition der „Experten“ erfolgte in Anlehnung an Schütz (1972) bzw. differenzierter nach Walther (1994) vgl. Meuser/ Nagel (2003: 484).

<sup>132</sup> Eine direkte Befragung wird auch bei Mitarbeiterbefragungen in Unternehmen als unumgänglich angesehen. Die tatsächliche Akzeptanz von Bewertungskriterien wäre hier eben nicht allein durch Befragung von Arbeitgeber- und Gewerkschaftsvertretern als „Experten“ feststellbar. (Vgl. z.B. Bartel-Lingg 1996)

Da in dieser Arbeit Ergebnisse von standardisierten schriftlichen Befragungen von Lehrenden und von Experteninterviews (u.a. mit ihren gewählten Vertretern) vorgelegt werden, erscheint es naheliegend, in ähnlicher Weise vorzugehen wie Jahoda u.a. und die Vor- und Nachteile der Methoden bei der Auswertung ebenfalls zu kombinieren und sich ergänzen zu lassen. Allerdings wird hier nicht davon ausgegangen, daß die Ergebnisse von Untersuchungen der einen Methode mit denen der anderen validiert i.S.v. überprüft werden könnten. Vielmehr sollen beide Methoden grundsätzlich gleichermaßen als valide oder nicht valide gelten. Die Frage nach der Validität der Ergebnisse entscheidet sich demnach hier nicht schon bei der Wahl der Methode (die natürlich gegenstandsangemessen erfolgen muß), sondern stärker bei der Durchführung und Auswertung.

Dies schließt jedoch in der Praxis nicht aus, daß man im Sinne einer „kommunikativen Validierung“ (vgl. Schröder-Lenzen (2003: 110) das „Forschungsobjekt“ nicht mehr lediglich als „Objekt“ auffaßt, sondern in Gesprächen durch das Bemühen um Verständigung und durch beobachtetes Verhalten z.B. die Schlüssigkeit von Interpretationen zu prüfen versucht, „denn mit dem reflexiven Subjekt kann man in Kommunikation treten und die Angemessenheit der Rekonstruktion der subjektiven Theorie im Dialog feststellen“. In der Praxis dieser Arbeit erfolgt jedoch keine Validierung im Sinne einer Überprüfung, sondern lediglich die Ergänzung der einen Perspektive mit einer anderen. Dies entspricht auch der Funktion der Experteninterviews zur Erlangung von „Kontextwissen“ im Gegensatz zur Überprüfung der Gültigkeit theoretischer Behauptungen, auf die bei der Durchführung und Auswertung noch zurückzukommen sein wird (vgl. Meuser/ Nagel 1994: 446).

### **Auswahl konkreter Gesprächspartner für die Experteninterviews**

Bei der Auswahl der Gesprächspartner für die Experteninterviews wurde in zwei Schritten vorgegangen: Zunächst waren geeignete Hochschulen und dann geeignete Gesprächspartner (an den Hochschulen) auszuwählen. Als Gesprächspartner wurden außerdem Vertreter der Gruppe der externen, nicht an den betreffenden Hochschulen beschäftigten, aber dennoch als Berater und über wissenschaftliche Publikationen auf diese einwirkende Personen aus der Hochschulforschung sowie Vertreter der Ministerien in die Überlegungen einbezogen.

Bei der Auswahl der Hochschulen für die Experteninterviews wurde davon ausgegangen, daß es bezogen auf das Arbeitsklima und die „Leitungskultur“ zwischen großen und kleinen Universitäten einerseits und Fachhochschulen andererseits häufig deutliche Unterschiede

gibt und deshalb möglichst jeder dieser Hochschultypen vertreten sein sollte. Außerdem wurde versucht, unterschiedlichen Landesgesetzgebungen wenigstens ansatzweise Rechnung zu tragen, indem Vertreter einer Hochschule sowie Vertreter eines Ministeriums aus einem anderen Bundesland einbezogen wurden. Dies deckt sich mit dem Prinzip des „maximal kontrastierenden Vergleiches“ bzw. des „theoretical sampling“ (vgl. Strauss 1987: 38), welches von Meuser/ Nagel (2003: 487) ausdrücklich als geeignete Vorgehensweise für Experteninterviews charakterisiert wurde. Außerdem ist es nach Merkens das Ziel einer Einordnung der zu untersuchenden „Fälle“, „die Ergebnisse auf andere, ähnlich gelagerte Fälle zu übertragen“ (vgl. Friebertshäuser/ Prengel 2003: 97). Nach Merkens kann dies über den Einbezug von für die Interpretation kritischen Fällen geschehen: „Das heißt, die Varianz der Ereignisse in der Stichprobe wird vergrößert. Wird die Hypothese hierbei nicht widerlegt, dann hat sie sich als robust erwiesen und die Wahrscheinlichkeit ihrer Angemessenheit ist erhöht worden.“ (Vgl. Merkens 2003: 99; Schmidt 2003: 545 sowie Lamnek 1995) Hier geht es jedoch nicht darum, möglichst *viele* verschiedene Fälle in die Untersuchung aufzunehmen. Vielmehr sollen – wie auch bereits von Morse (1994) beschrieben – möglichst stark „abweichende oder extreme Fälle“ einbezogen werden.

Eine erste Gruppierung der Interviewpartner erfolgte hier für das Themenfeld hochschulpolitische Reformvorhaben jedoch nicht wie bei Merkens beschrieben erst in einem zweiten Schritt, sondern bereits bei der Auswahl der Interviewpartner. Es war zu vermuten, daß Vertreter der Personalräte bezüglich der Hypothese der Akzeptanz bestimmter hochschulpolitischer Vorhaben z.B. einer stärkeren Leistungsorientierung skeptischer gegenüberstehen als Vertreter der Hochschulleitungen und daher mit dieser zu kontrastieren wären. Als repräsentativ im Sinne der häufig bei quantitativen Studien verwendeten Definition kann diese Auswahl dabei aufgrund der anderen Untersuchungsanlage und -ziele natürlich nicht gelten. An die Stelle der „Repräsentativität“ tritt hier also die Forderung nach inhaltlicher Repräsentation. Nach Glaser/ Strauss (1967) kann beim „theoretical sampling“ die Entscheidung, welche und wieviele Fälle in eine solche Untersuchung einbezogen werden müssen, nur theoriegeleitet abgesichert werden (vgl. Merkens 2003: 103). Wie sich bei der Auswertung der Experteninterviews nachträglich herausstellte, ergaben sich aber auch keine Anhaltspunkte für die dringende Notwendigkeit einer anderen Gruppierung der Experteninterviews.

Dabei besteht „wie bei allen qualitativen Studien unabhängig vom Typ die wesentliche Herausforderung darin, gute Informanten zu finden“ (vgl. Spradley 1979). Gute Informanten wären hierbei nach Morse (1994) mit folgenden Merkmalen zu beschreiben:

- Sie verfügen über das Wissen und die Erfahrung, deren die Forscher bedürfen,
- sie haben die Fähigkeit zu reflektieren,
- sie können sich artikulieren,
- sie haben die Zeit, interviewt zu werden und
- sie sind bereit, an der Untersuchung teilzunehmen.

Zusammenfassend ist zu formulieren, daß die Auswahl der Hochschulen und Personengruppen v. a. dazu diene, eine möglichst große Bandbreite von Situationen und Einschätzungen bei der Auswahl der Gesprächspartner für die Experteninterviews in die Überlegungen einzubeziehen, nicht jedoch, „Repräsentativität“ herzustellen. Hierbei war außerdem zu berücksichtigen, dass Ergebnisse der Experteninterviews ggf. als ergänzende Informationen zu Lehrendenbefragungen an den entsprechenden Hochschulen herangezogen werden sollen. Nachfolgende Übersicht ordnet die zehn ausgewählten Interviewpartner den jeweiligen Hochschulen und interviewten Personengruppen zu. Den Interviewpartnern wurde ein Großbuchstabe des Alphabetes zugewiesen.

Vertreter der Hochschulen:

Hochschultyp/ Personengruppe	Vertreter der Hochschulleitung	Gewerkschaftliche Vertreter/ Personalrat
Fachhochschule: Hochschule Zittau-Görlitz	B	G
kleine Universität: Europa-Universität „Viadrina“ Frankfurt/ Oder	N	<sup>133</sup>
große Universität: Technische Universität Dresden	M	S

Mit dem Thema befaßte Hochschulforscher/ hochschulübergreifende Institutionen:

Universität Magdeburg, Mitglied der Fakultät für Naturwissenschaften	R
Centrum für Hochschulentwicklung Gütersloh (CHE)	A
Hochschul-Informations-Systeme GmbH Hannover (HIS)	L
Senat für Wissenschaft und Kunst Berlin (SenWK)	P
Sächsisches Staatsministerium für Wissenschaft und Kunst (SMWK)	<sup>134</sup>

<sup>133</sup> Leider kam aufgrund fehlender Bereitschaft des Personalrates an der Europa-Universität „Viadrina“ Frankfurt/ Oder kein Gespräch mit diesem zustande (lt. Email aufgrund kurz bevorstehendem Amtswechsel). Damit war – vom fehlenden Zugang abgesehen – zugleich die fünfte Voraussetzung für einen „guten Informanten“ in Anlehnung an Morse (1994) nicht gegeben.

Unabhängig hiervon wurde, um nicht allein auf die Eindrücke und Informationen aus dem Gespräch mit einer Person (der Vertreter/in der Hochschulleitung) an dieser Hochschule angewiesen zu sein, jedoch zusätzlich ein internes, aufgrund fehlenden Einverständnisses nicht aufgezeichnetes Gespräch mit einer Mitarbeiterin des Controlling durchgeführt, um einen breiteren Einblick in die Situation an dieser Hochschule zu erhalten. Der Gesamteindruck dieses Gespräches deckte sich mit dem des aufgezeichneten Interviews.

<sup>134</sup> Das Gespräch mit Vertretern des SMWK wurde zwar durchgeführt; die Informationen konnten aber letztlich aus Gründen, die im entsprechenden Abschnitt zur Durchführung der Experteninterviews noch erläutert werden, nicht einbezogen werden. Da der verbleibende Vertreter eines Ministeriums große Nähe zur Hochschulforschung aufweist, wurden Hochschulforscher, hochschulübergreifende Institutionen und Ministerien zusammengefasst.

### 4.3 Durchführung der Experteninterviews

Die Experteninterviews erfolgten in grober Anlehnung an die von Meuser/ Nagel (1994) beschriebene Vorgehensweise. Die beiden Autoren gehören nach wie vor zu den sehr wenigen, die Experteninterviews nicht nur anwendeten, sondern auch umfassend methodisch reflektierten (vgl. Friebertshäuser/ Prengler 2003: 19 bzw. Meuser/ Nagel 2003: 482<sup>135</sup>). Die Interviews spielen im Forschungsdesign dieser Arbeit nicht die zentrale Rolle, sondern sollen zusätzliche Informationen und Hintergrundwissen liefern. Deshalb müssen sie hier eher eine „Randstellung“ einnehmen – um mit Meuser/ Nagel (1994: 445) zu sprechen. Eine detaillierte Auswertung und Positionierung der Experteninterviews im Zentrum eines Untersuchungsdesigns wäre sicher ein lohnendes Unterfangen. Allerdings würde dies den Rahmen der vorliegenden Arbeit bei Weitem sprengen und könnte eher Gegenstand einer weiteren, eigenständigen Arbeit sein.

Die forschungslogisch motivierte Funktion des Expertenwissens in dieser Untersuchung ist die Erlangung von „Kontextwissen“ (Meuser/ Nagel 1994: 446). Die Experteninterviews sind nur ein Erhebungsinstrument neben anderen. Ihre Ergebnisse sollen letztlich zur *Bestimmung* eines Sachverhaltes beitragen und nicht zur *Überprüfung der Gültigkeit theoretischer Behauptungen* über den Sachverhalt. Demzufolge werden die Texte nur partiell bezüglich bestimmter Forschungsfragen ausgewertet. Dies wurde auch bei sich stärker qualitativ einordnenden Auswertungen so gehandhabt, wenn diese nur auf bestimmte Themen fokussiert waren (vgl. „Problemzentriertes Interview“ bei Witzel 1995: 56, sowie ähnliche Ausführungen in Deeke 1995: 19). In vorliegender Auswertung wurde daher von vornherein keine detaillierte Lautverschriftlichung mit entsprechenden Transkriptionszeichen vorgenommen (vgl. Friebertshäuser/ Prengler 2003: 252). Lediglich einige stark betonte Wörter bzw. Zitate wurden kursiv gesetzt, da sich der Sinn mancher Passagen erst aus der Betonung erschließt. Ansonsten wurde überall dort, wo dies ohne größere Gefahr der Beeinträchtigung des Sinngehaltes möglich war (bei fehlenden Artikeln, falscher Satzstellung usw.), zur besseren Lesbarkeit eine grammatikalische Glättung vorgenommen. Wo es für das Verständnis des Sinngehaltes notwendig schien, Passagen stärker zu verändern,

---

<sup>135</sup> Andere Autoren, die zu ähnlichen Themen im Hochschulbereich Experteninterviews durchführten, gingen kaum auf deren methodische Probleme ein (vgl. Schröder 2004, Minssen/ Wilkesmann 2003, Orr 2001). Die hier angewandte Vorgehensweise wurde außerdem im Rahmen eines Workshops der Hans-Böckler-Stiftung im Jahre 2003 mit Ulrike Nagel diskutiert und damit ergänzend zu den hier zitierten Veröffentlichungen von 2003 und 1994 (wobei letztere deutlich ausführlicher methodische Probleme diskutiert) an aktuelle Erkenntnisse und spezifische Problemlagen angepasst.

wurde dies entsprechend kenntlich gemacht. In diesen Fällen wurden z.B. erläuternde Ergänzungen vorgenommen.

Grundlage der Experteninterviews war ein „offener Leitfaden“ (Meuser/ Nagel 1994: 448). Die Interviewten wurden jeweils aufgefordert, möglichst nach Individualebene (d.h. z.B. Ebene von Besoldungszulagen für einzelne Hochschullehrer) und der Institutionenebene zu differenzieren (Studiengänge, Fakultäten und Hochschulen). Der Leitfaden beinhaltete folgende Themen:

1. Einleitung: kurze persönliche Vorstellung (und des Promotionsvorhabens)
2. Ausgangslage („Rahmen“) ist der Wille der Politik, Alterszulagen in Leistungszulagen umzuwandeln. Die Frage ist daher wohl nicht mehr ob, sondern *wie* dies geschehen soll.
3. Diskussionspunkte:
  - 3.1 Einschätzung der bislang diskutierten Verfahren<sup>136</sup>/Leistungskriterien?
  - 3.2 Welche Verfahren/ Kriterien wären geeignet?
  - 3.3 Inwieweit Einbeziehung von Studenten-/ Absolventenbefragungen akzeptabel?
  - 3.4 Welche Gewichtung Grundgehalt/ Leistungsanteil favorisiert?
  - 3.5 Wie Anteile gewichten bzgl. Forschung/ Lehre/ Gremientätigkeit?
  - 3.6 Welche Art der Kopplung von Leistungsbewertung und -anreizen?
    - direkte Kopplung von Mittelvergabe und Leistungsbewertung,
    - Belohnung von Verbesserungen (oder befristet: des Bemühens),
    - verzögerte Kopplung von Evaluation und Mittelfreigabe (wie in den Niederlanden)?
  - 3.7 Wie groß und welche Art der „Spreizung“ der Zulagen?
    - Prinzip formelgebundener Kennzahlen als Orientierung an festgestelltem gewichteten Mittel der Leistungskennzahlen (Pseudogenauigkeit?) oder
    - Prinzip Einteilung in Leistungsgruppen mittels zuverlässiger stat. Verfahren wie Signifikanzanalysen (z.B. Spitzengruppe: +30%, Indifferente: +15% und Schlussgruppe: 0%)
    - oder andere Verfahren (Welche?)
4. Abschließend: Weitere Aspekte, die Ihnen im Rahmen dieses Themengebietes wichtig sind und die bisher evtl. noch nicht angesprochen wurden?

---

<sup>136</sup> Im Leitfaden ist wie in den Interviews von Verfahren und Kopplungsarten die Rede, um an das Alltagssprachverständnis anzuknüpfen. Eine Vereinheitlichung der Begriffe wie im 2. Kapitel erfolgte wg. der Zitate nicht.

In diesem Zusammenhang soll nochmals klargestellt werden, daß trotz der relativ klaren Strukturierung des Leitfadens eine teilstandardisierte schriftliche Befragung diesen Zweck nicht hätte erfüllen können (vgl. auch Deeke 1995: 18): Erstens stellte sich schon bei vorangegangener Sondierung des zu untersuchenden Feldes schnell heraus, daß dies bereits am Problem des Feldzuganges gescheitert wäre. (Die Experten hätten sich nicht die Zeit genommen, eine schriftliche Befragung zu beantworten und wären damit – wenn überhaupt – jedenfalls keine *guten* Informanten gewesen.) Zweitens sollten die Befragten nicht nur einfach eine Antwort geben, sondern auch Gelegenheit zu damit verbundenen Erläuterungen, Begründungen und ggf. Einschränkungen haben. Drittens ist in einem mündlichen, relativ offenen Gespräch feststellbar, ob die Zuschreibung der Expertenrolle auf die befragte Person auch berechtigt war. Und schließlich hätten (viertens) mehrere Mißverständnisse und unterschiedliche Begrifflichkeiten, die sich im Verlauf der Interviews herausstellten und geklärt wurden, ungeklärt bleiben müssen. Nach Deeke (1995: 18) ist gerade auch die Erarbeitung des kontextgebundenen Wissens „durch klärende Fragen und Gegenfragen, durch diskursive Rekonstruktion usw.“ ausdrücklich Gegenstand von Experteninterviews, auch wenn es nicht inhaltlich um subjektiv gemeinten Sinn geht. Eine teilstandardisierte schriftliche Befragung hätte daher nicht oder nicht im selben Ausmaß zur Aufhellung der hier interessierenden Sachverhalte beitragen können.<sup>137</sup>

Dennoch gilt es bei Experteninterviews wie bei allen mündlichen face-to-face-Interviews, einen möglichen Interviewereinfluß zu beachten (vgl. z.B. Cicourel 1970). Meuser/ Nagel (1994: 449 ff.) haben dies weiterführend in verschiedenen „Formen des Mißlingens“ typisiert. Die hier durchgeführten Interviews sind demnach im Sinne der Typisierungen<sup>138</sup> „d“ (tendenziell eher Interviewcharakter) oder „e“ (tendenziell eher Dialogcharakter) in der überwiegenden Zahl als „gelingen“ einzustufen.<sup>139</sup> Lediglich ein Interview (das im SMWK) muß einer Form des Mißlingens zugeordnet werden.

---

<sup>137</sup> Vgl. hierzu auch Plath (1995: 85), nach dem es bei Experteninterviews zumeist gerade um Realitätsausschnitte gehe, die komplex, teilweise undurchschaubar, untereinander vernetzt, eigendynamisch, polytelisch und bezüglich des zu erreichenden Zielzustandes offen sind, so daß eine angemessene Erfassung des Expertenwissens eines kommunikativen Prozesses bedarf.

<sup>138</sup> Von Meuser und Nagel (1994) werden insgesamt fünf Formen des „Mißlingens“ und des „Gelingens“ vorgestellt.

<sup>139</sup> Es wäre allerdings z.T. sehr schwer, klar zu unterscheiden, welche „Form des Gelingens“ letztlich vorlag, da in einigen Fällen bei Nachfragen der Interviewten Erläuterungen durch Interviewer zu bestimmten Sachverhalten notwendig waren, um ein gegenseitiges Verstehen zu ermöglichen. Dies gab auch ansonsten tendenziell eher als Interview einzuordnenden Fällen streckenweise Dialogcharakter.

Zu diesem Interview erschien nicht wie im Vorgespräch vereinbart eine Person, sondern drei Personen. Hierbei referierte der erst kurz zuvor auf diese Funktion berufene Referatsleiter ausführlich darüber, daß er „zwangsversetzt“ wurde und stufte sich selbst für den im Interview anzusprechenden Themenkreis als nicht ausreichend kompetent ein. Unglücklicherweise war die Gesprächsatmosphäre mit den aus Sicht des Interviewers durchaus als kompetent eingeschätzten Mitarbeitern durch die Anwesenheit des neuen Referatsleiters jedoch stark beeinträchtigt, so daß diese nur sehr zurückhaltend antworteten. Wie in einem solchen Fall (Form des Mißlingens: Typ „a“) von Meuser/ Nagel (1994: 449) empfohlen, wurde das Interview jedoch nicht abgebrochen, sondern versucht, für alle Beteiligten das Gesicht zu wahren und das Interview so planmäßig wie möglich zu Ende zu führen, um evtl. doch noch ein wenig zur Erhellung des Sachverhaltes beitragen zu können. Das Interview war jedoch für die Zielstellung dieser Untersuchung letztlich nicht aussagekräftig genug<sup>140</sup> und wurde von der weiteren Auswertung ausgeschlossen.

## **4.4 Auswertung der Experteninterviews**

### **4.4.1 Arbeitsschritte der Auswertung**

Die nachfolgend dargestellten Auswertungen wurden in folgender Systematik durchgeführt: Nach der Verschriftlichung waren die zentralen Passagen zu den jeweiligen Fragestellungen zu identifizieren. Diese wurden entweder zu der entsprechenden Sequenz paraphrasiert (vgl. Meuser/ Nagel 1994: 456) oder bei ausreichender Prägnanz als Originalzitat übernommen. Die Fundstelle ist nachfolgend in Klammern hinter dem jeweiligen Zitat mit Angabe der Seitenzahl aus dem Anhang mit den verschriftlichten Experteninterviews belegt. Die zentralen Passagen sind in Form einer tabellarischen Übersicht zu jedem Interview komprimiert auf 2 bis 3 Seiten zusammengestellt. Im darauffolgenden Arbeitsschritt wurde versucht, zu den jeweiligen Textpassagen die Kernaussage in Form eines Wortes oder weniger Worte als eine Art „Überschrift“ herauszuarbeiten (vgl. Meuser/ Nagel 1994: 457). Diese sollte möglichst in den Worten der jeweils Interviewten gefunden werden und wurde in der tabellarischen Übersicht durch Fettdruck markiert. Nachdem die ersten drei Arbeitsschritte durchgeführt wurden und die Interviews in Form von tabellarischen Übersichten vorlagen, wurde der thematische Vergleich vorgenommen. „Passagen aus

---

<sup>140</sup> Die aufgetretenen Probleme werden auch von anderen Autoren als für die öffentliche Verwaltung typisch beschrieben (vgl. z.B. Hägele 1995: 70) und als schwerwiegend eingeschätzt.

verschiedenen Interviews, in denen gleiche oder ähnliche Textpassagen behandelt werden, werden zusammengestellt und die 'Überschriften' vereinheitlicht' (vgl. Meuser/ Nagel 1994: 459 ff.). Im letzten Arbeitsschritt dieser Auswertung, der soziologischen Konzeptualisierung, wird „im Rekurs auf soziologisches Wissen das Gemeinsame im Verschiedenen' begrifflich gestaltet und in die Form einer Kategorie 'gegossen'“ (vgl. Meuser/ Nagel 1994: 462). Die Abstraktionsebene, auf der man sich bei diesem Auswertungsschritt bewegt, ist die der empirischen Generalisierung. Da das Erkenntnisinteresse auf dem Kontextwissen liegt, kann die Auswertung auf dieser Stufe abgebrochen werden (vgl. Meuser/ Nagel 1994: 466). Die einzelnen Arbeitsschritte der Auswertung werden nun zur besseren Übersicht noch einmal stichpunktartig aufgelistet:

1. Transkription
2. Paraphrasieren bzw. Zusammenstellen prägnanter Textpassagen
3. Herausarbeiten von Kernaussagen bzw. „Überschriften“
4. Thematischer Vergleich
5. Soziologische Konzeptualisierung

#### **4.4.2 Thematischer Vergleich**

Der thematische Vergleich wurde anhand der Zusammenstellung prägnanter Textpassagen, aber nicht in der Reihenfolge der Fragestellungen vorgenommen. Für das Verständnis der Sinnzusammenhänge erscheint eine andere Reihenfolge sinnvoller. Dies zeigte sich bereits bei der Durchführung einiger Interviews, in denen die Experten die avisierten Themen z.T. in anderer Reihenfolge als in der des Leitfadens ansprachen. Auch dort wurde dem Erzählfluß der Experten gefolgt.

Nachfolgend werden anhand ausgewählter prägnanter Textpassagen aus den verschriftlichten Experteninterviews die jeweiligen Positionen der Gesprächspartner illustriert. Hierbei wurden die Zitate zur besseren Gewährleistung der Übersichtlichkeit als eingerückter Text dargestellt. Die vollständige Zusammenstellung prägnanter Textpassagen wurde im Anhang dokumentiert.

## **Grundpositionen der Gesprächspartner**

Nach einer kurzen Selbstvorstellung und Beschreibung des Rahmens der Experteninterviews wurden die Interviewpartner gefragt: Was halten Sie von den bislang diskutierten Verfahren/Leistungskriterien? Dazu wurden sie gebeten, dies einerseits auf der Ebene von Besoldungszulagen für einzelne Hochschullehrer (Individualebene) und andererseits auf der Ebene von Studiengängen, Fakultätsebene, Hochschulebene (institutionelle Ebene) zu betrachten.

Die erste Frage stellte somit damit eine Art „Aufwärmfrage“ dar, da die Experten hierbei Gelegenheit hatten, sich zur Gesamtsituation zu äußern und sich selbst zu positionieren. Allerdings können die Antworten zu dieser Frage auch einer grundsätzlichen Einordnung der Interviewten und einer ersten Beschreibung der Grundpositionen dienen, die aber ggf. am Ende der Interviewauswertungen zu revidieren wäre. Diese sollte zugleich schlaglichtartig einen kleinen Einblick in die Grundstimmung der Handelnden im Gesprächszeitraum zeigen. Die Frage ermöglicht, wie sich im Nachhinein bei der Auswertung zeigte, zum Teil auch eine Differenzierung, in welcher Rolle sich die Gesprächspartner sehen: Eher als „Privatperson“ (mit eigener Meinung) oder eher als „Funktionär/in“ (als Vertreter/in der Hochschule, des Personalrates, eines Wissenschaftsgebietes oder Ministeriums). Überwiegend äußern sich die Experten wie erwartet in ihrer Rolle als Vertreter ihrer beruflichen Aufgabe (insbesondere die Vertreter/innen der Hochschulleitungen fast ausschließlich). An den Stellen, wo die Gesprächspartner/innen sich davon abweichend klar als Privatpersonen äußern, wird dies von ihnen selbst kenntlich gemacht, indem sie es explizit dazu sagen.<sup>141</sup> So äußert sich z.B. S (Vertreter/in des Personalrates) im Zusammenhang mit der Frage nach geeigneten Kriterien explizit als Privatperson und deutet an, daß sie/ er auch innerhalb des Personalrates z.T. eine Minderheitenmeinung vertritt und nicht in allen, aber wohl in einigen solchen Fällen auch die Diskussion innerhalb des Personalrates mit der „Privatmeinung“ beeinflusst hat. Daher erscheint dies als Hintergrund für die Einordnung der Äußerungen wichtig und wurde in die tabellarische Übersicht und bei der Interpretation der Interviews einbezogen.

---

<sup>141</sup> Es ist jedoch natürlich in Rechnung zu stellen, daß sich die Gesprächspartner (auch wenn sie es nicht explizit deutlich machen) nicht nur als Repräsentanten ihrer Organisation äußern, sondern ebenso als Personen. Sie sind deshalb nicht als „austauschbare Rollenträger“ zu sehen (vgl. Krafft/ Ulrich 1995: 28). Da sie jedoch auch als Personen und nicht nur als „austauschbare Rollenträger“ in das zu untersuchende Umfeld wirken und dies für das Umfeld ebenfalls häufig nicht klar zu trennen ist, soll hier die Wirkung von Funktion und Person gemeinsam betrachtet werden.

Die Interviews lassen sich bezüglich der Frage nach ihrer Grundposition grob in zwei Gruppen einordnen: Die erste Gruppe von Interviewten nahm eher eine Art Mittlerposition<sup>142</sup> ein, während sich die zweite Gruppe sehr viel klarer für stärkere Leistungsbewertung und Leistungsanreize aussprach. Keine/r der Interviewten lehnte dies geradeheraus ab.<sup>143</sup> Auffällig ist, daß sich die (interviewten) Vertreter der Personalräte, bei denen aufgrund von Beiträgen in gewerkschaftlichen Publikationen (vgl. z.B. Conrady 2003: 10 sowie Finkemeier 2003: 30<sup>144</sup>) allgemein eher eine skeptischere Haltung zu vermuten war, klarer für Leistungsbewertung und Leistungsanreize aussprachen als Vertreter der Hochschulleitungen, die sich eher eine Mittlerposition zuwiesen. Möglicherweise ist dies zufällig und spezifisch für die interviewten Gesprächspartner. Zumindest muß aber festgehalten werden, daß es auch diese Positionen gibt, die überraschend wenig mit der verbreiteten (Selbst-)Darstellung von häufig gewerkschaftlich orientierten Personalvertretungen übereinstimmen. Nachfolgend soll dies anhand möglichst prägnant formulierter Textpassagen illustriert werden.

So formuliert beispielsweise G (Personalrat) im Gegensatz zu N (Hochschulleitung):

G: „Leistungsbewertungen sollte man unbedingt machen, denn ansonsten gibt es ja überhaupt keine Motivation für Leute, denen es ein bißchen schwerer fällt, (auch)<sup>145</sup> einen Anreiz zu schaffen. Unter anderem für die, die ein bißchen lasch mit sich umgehen würden.“ (101)

N: „Also, was ich sehr schwierig finde ist, wenn man versucht das eine Modell, das sich als schlecht erweist, durch das andere Kontrastmodell zu ersetzen (...). Also von der völligen Unkontrolliertheit quasi - momentan ist es ja relativ egal, wie viele Absolventen man produziert - in die sozusagen ökonomisierte

---

<sup>142</sup> Dieser Begriff wurde bewußt gewählt, um eine möglichst treffende Beschreibung zwischen der Position von „Vermittlern“ und einer Mittlerposition (zwischen verschiedenen Standpunkten) zu finden, da eine Kombination beider Aspekte die Grundposition am besten beschreibt.

<sup>143</sup> Allerdings kam - wie bereits erwähnt - ein geplantes Gespräch mit einem Personalrat nicht zustande.

<sup>144</sup> So wird hier eingeschätzt, daß die Variabilisierung von tariflichen Gehältern von (in diesem Fall) Betriebsräten „in der Regel als unzumutbar“ zurückgewiesen werden. Die vielerorts bereits bestehenden leistungs- und erfolgsabhängigen Entgeltanteile seien „beinahe überall“ übertariflich. Aber hier gelte ebenso: „Die Erfahrungen mit übertariflichen variablen Entgeltsystemen sind zwiespältig.“ Die meisten Systeme sehen zwei Komponenten vor: Eine vom Betriebsergebnis abhängige Erfolgsbeteiligung (Gruppenbezogene Zulage) sowie eine auf die individuelle Leistung des Mitarbeiters bezogene Zusatzvergütung (Individualzulage). Hauptkritikpunkt der Betriebsräte ist: „Für die Aktionäre wird anders gerechnet als für die Belegschaft.“ Während für die Aktionäre die Bilanzen zählen, sei für die Mitarbeiter die Erreichung von Zielvorgaben maßgeblich, die die Unternehmensleitung festgesetzt hat. Diese seien „zum Teil sehr verschwommen“ und die Latte zudem oft sehr hoch gelegt. Bei aller Kritik wird jedoch eingeräumt: „Bislang kommt bei der erfolgsabhängigen Vergütung immer was raus für die Mitarbeiter.“

<sup>145</sup> Dies ist zu verstehen i.S.v.: für sich selbst.

Kontrolle und setzt damit auch Leistungs- bzw. überhaupt Handlungsanreize, die problematisch sind, finde ich.“ (55) „...ich kann mich in keines der beiden Lager schlagen“ (89).

Die interviewten Wissenschaftler aus dem Bereich der Hochschulforschung ordneten sich eher als Befürworter einer stärkeren Leistungsorientierung ein. A formuliert dies sehr deutlich und quantifiziert diese Einschätzung zusätzlich:

„Ja, ich finde das prinzipiell gut, daß man das macht. (...) Also auf einer Skala von 0 bis 10 mit 10 "Begeisterung", 0 "furchtbar", da ordne ich mich locker bei 7 oder 8 ein.“ (173)

### **Art der Kopplung von Leistungsbewertung und -anreizen**

Zu diesem Unterthema des Experteninterviews wurden die Gesprächspartner gebeten, Ihre Einschätzungen zur Art der Kopplung von Leistungsbewertung und -anreizen zu formulieren (sofern sie dies nicht im Verlauf des Gespräches bereits von sich aus getan hatten). Hierzu wurden ihnen einheitlich drei mögliche Arten genannt (und bei Erklärungsbedarf kurz Beispiele dazu erläutert): Direkte Kopplung von Mittelvergabe und Leistungsbewertung (z.B. formelgebundene Mittelzuweisungen), Belohnung von Verbesserungen (oder befristet: des Bemühens darum, also z.B. Zielvereinbarungen), zeitlich verzögerte Kopplung von Evaluation und Mittelfreigabe (Modell wie in den Niederlanden, wo die Finanzierung nach zweifacher negativer Evaluation eingestellt oder eingeschränkt werden kann).

In der ersten Nennung dieses Unterthemas wurde (sofern eine beispielhafte Erläuterung nicht notwendig war) versucht, zunächst bewußt eher abstrakt nur Grundprinzipien und noch keine konkreten Beispiele zu nennen, um die Befragten nicht vorschnell zu konkreten Erfahrungsberichten anzuregen. Denn an dieser Stelle ging es vorrangig um die Erfassung der grundsätzlichen Einschätzung unabhängig von evtl. vorhandenen konkreten positiven oder negativen Erfahrungen. Außerdem sollte nicht von vornherein der Fokus auf finanzielle Anreize gelenkt werden. Dennoch wurden von allen Gesprächspartnern vor allem diese thematisiert.

Im weiteren Verlauf der Interviews war es aber z.T. zum gegenseitigen Verständnis notwendig (z.B. bei Personalratsvertretern), konkrete Beispiele zu nennen. In diesen Fällen

wurde z.B. zur zeitlich verzögerten Kopplung das Modell von Evaluation und Mittelfreigabe in den Niederlanden erwähnt und erläutert. Dies führte im Gesprächsverlauf z.T. eher zu einem Dialogcharakter der Interviews. Hier war es nicht möglich, immer gleichlautende „Stimuli“ zu setzen, gleichwohl ist dies als ‚Form des Gelingens‘ einzustufen (vgl. Abschnitt zur Vorgehensweise bei der Auswertung bzw. Meuser/ Nagel 1994: 449).

Insgesamt nannten die meisten Gesprächspartner an erster Stelle die Belohnung von Verbesserungen (oder befristet: des Bemühens darum) als eine ihrer Einschätzung nach geeignete Möglichkeit. Oft wurden als Beispiel für diese Art der Kopplung Zielvereinbarungen genannt. Eine direkte Kopplung als alleiniges Modell wird eher abgelehnt, wobei dies im Interview mit B (Hochschulleitung) am prägnantesten begründet wurde:

B: „Also, daß die Starken noch mehr kriegen und die Schwachen noch weniger. Also das ist ein System, das die Ungleichheiten verfestigt, aber nicht hilft sie aufzulösen.“

Allerdings können sich viele auch eine Kombination von Modellen vorstellen, wie z.B. im Interview mit N (ebenfalls Hochschulleitung) dargelegt wurde:

N: „Ich finde auch, daß man da sozusagen (versch. Arten der Kopplung) kombinieren muß und das machen alle schlaun Modelle. Auch die schlaun Modelle des CHE, die es sogar versuchen, sozusagen das „Ist“ auch irgendwie zu berücksichtigen. Also zum Beispiel jetzt eine Uni, die schon dreißig Prozent Ausländer hat oder so, die kann natürlich ihren Ausländeranteil nicht noch beliebig steigern. Und wenn man da dann nur das Positive/ die Weiterentwicklung bemißt, dann kommt man irgendwann zu einem konträren Ziel. Und ich glaube, das muß man immer mit in Betracht ziehen (.d.h.:) auf welcher Schwelle sich das „Ist“ befindet und das einrechnen in die Verbesserungen.“

Im Interview mit S (Personalrat) wird auch ein möglicher Weg thematisiert, wie eine Kombination aussehen könnte:

„Warum nicht bestimmte Verbesserungen belohnen und trotzdem die Mittelvergabe und Leistungsbewertung koppeln. Das war ja das, wo wir vorhin diskutiert haben. Also eine bestimmte Leistung (die zuvor als Ziel vereinbart wurde - Anm. des Verf.), die kriegt etwas besonderes, die wird als besonders hervorgehoben und das Normale ist ansonsten die Kopplung an die Leistungsbewertung. Das kann man doch alles machen.“

Im Interview mit A (Hochschulforschung) wird dies sehr ähnlich formuliert, wobei als zusätzlicher Aspekt eingeführt wird: „...manchmal kann man vielleicht auch Leistungen noch nicht so richtig abschätzen“ (und sollte daher zunächst stärker das Bemühen darum per Zielvereinbarung mit Leistungsanreizen verbinden, wie dann weiter besprochen wurde).

### **Art der „Spreizung“ der Zulagen**

Zur Art der „Spreizung“ der Zulagen wurden die Experten gefragt, wie groß und welche Spreizung sie evtl. für sinnvoll hielten. Hierzu wurden wiederum mögliche Varianten genannt. Erstens war es das Prinzip formelgebundener Kennzahlen als Orientierung an festgestelltem gewichteten Mittel der Leistungskennzahlen (wobei hier eruiert werden sollte, ob dies nicht eine Pseudogenauigkeit suggeriert) oder zweitens das Prinzip der Einteilung in Leistungsgruppen evtl. mittels zuverlässiger statistischer Verfahren wie Signifikanzanalysen (z.B. Spitzengruppe: +30%, Indifferente: +15% und Schlussgruppe: 0%) oder drittens andere Verfahren (Wobei hier die Frage wäre, welche das sein könnten bzw. sollten?)

Diese Frage zielte auf eine Einschätzung der Vor- und Nachteile der beiden Prinzipien formelgebundene Mittelzuweisung (ohne weitere Festlegungen zur Höhe) versus Einteilung in Leistungsgruppen ab, wobei auch die Art und Weise der Abgrenzung verschiedener Leistungsgruppen gegeneinander thematisiert werden sollte (z.B. mittels in der Fachwelt als zuverlässig eingeschätzter statistischer Verfahren wie Signifikanzanalysen). Soweit es sinnvoll erschien, sollten die Gesprächspartner zudem angeregt werden, andere ihnen bekannte Verfahren vorzustellen oder eigene Vorstellungen zu erläutern.

Insgesamt favorisierten die meisten Gesprächspartner eine Einteilung in Leistungsgruppen. Allerdings wollten sich mehrere Gesprächspartner zu dieser Frage nicht äußern, da sie entweder wie M (Hochschulleitung) diese Diskussion für „noch nicht spruchreif“ hielten oder wie N (ebenfalls Hochschulleitung) solche Prinzipien grundsätzlich ablehnten. Auffällig ist denn auch, daß sich vor allem Vertreter der Hochschulleitungen hierzu nicht positionieren wollten (obwohl ihnen vollständige Anonymität zugesichert wurde). Vertreter der Personalräte wie S befürworteten klar eine Einteilung in Leistungsgruppen bzw. befürworteten dies wie G (ebenfalls Personalrat) zumindest für Verantwortungsträger:

G: „Individuell auf jeden Fall für die, die Verantwortung tragen. Und die anderen, denen sollte man die Grundzuführung geben.“ (117)

Die interviewten Hochschulforscher favorisierten einhellig Leistungsgruppen bzw. stehen dem aufgeschlossen gegenüber. So favorisiert R drei bis vier Leistungsgruppen, während A das Stufenmodell des CHE beschreibt:

R: „3 oder 4. Mehr, würde ich sagen, sind überdifferenziert. Also, bisher haben wir ja 2: C3 und C4. Vielleicht noch C2. (...) 4 Gruppen, denke ich, müßten reichen...“ (167)

N: „Ich glaube, daß das viele Probleme oder potenzielle Probleme mit einer Klappe erschlägt. Das ist vielleicht auch nicht das Nonplusultra, vielleicht kommen ja da auch noch raffiniertere Lösungsvorschläge, aber ich glaube, das ist eigentlich ein guter, machbarer Ansatz, daß man sagt, man hat also kein, es ist nicht alles free-floating, sondern man definiert bestimmte Level (...)“ (175)

### **Gewichtung des Leistungsanteils im Verhältnis zur Gesamtfinanzierung**

Die Frage, welche Gewichtung des Leistungsanteils die Gesprächspartner favorisieren, wurde zunächst allgemein gestellt und dann gebeten, dies nach Individualebene (z.B. Besoldungszulagen für einzelne Hochschullehrer) und institutioneller Ebene (Studiengänge bzw. Fakultäten bis hin zur Hochschulebene) ggf. noch zu differenzieren.

Hierbei bleibt zu ergänzen, daß auf der Individualebene mit der Umsetzung der Reform der Besoldungszulagen für Professoren bereits relativ konkrete Vorstellungen und Diskussionen existieren und damit Grundgehalt und Leistungsanteil von der Begrifflichkeit her den Gesprächspartnern relativ klar waren. Analog lassen sich auf der institutionellen Ebene der Anteil von Grundzuführung und leistungsbezogener Mittelzuweisung betrachten. Allerdings gab es in den Interviews z.T. bei der Zuordnung zu den genannten Ebenen Verständnisschwierigkeiten, die in diesen Fällen erst durch Erläuterungen ausgeräumt werden mußten. Obwohl aus diesem Grunde nicht bei allen Gesprächen von gleichlautenden Stimuli ausgegangen werden kann, ließ sich in allen Fällen nach relativ kurzer Zeit nach Einschätzung des Verfassers gegenseitiges Verständnis herstellen, so daß hier von einer guten Verwertbarkeit der Ergebnisse ausgegangen werden kann.

Die Ergebnisse fielen ähnlich den bereits eingangs der Auswertung formulierten Grundpositionen aus, wobei sich die Interviewten wieder grob in die Gruppe der Vertreter der Hochschulleitungen und die Gruppe der übrigen einteilen ließen.

Die interviewten Vertreter der Hochschulleitungen äußerten sich überwiegend nicht konkret zur Gewichtung des individuellen Leistungsanteils. Bezogen auf die institutionelle Ebene wird betont, daß die mögliche Gewichtung auch von der absoluten Höhe der Grundausstattung (N) bzw. von „finanzpolitischen Rahmenbedingungen“ (M) abhängt. Letztere/r Vertreter/in der Hochschulleitung führt zudem aus:

M: „Der Leistungsanteil sollte größer gleich 15% sein. Dies hängt aber von finanzpolitischen Rahmenbedingungen ab, die derzeit (und wohl auch auf absehbare Zeit) komplett unter der Prämisse stehen, daß Anreizsysteme ‚kostenneutral‘ zu sein haben“.

Die beiden Personalvertreter/innen äußerten sich wiederum konkreter und ohne Vorbedingungen: 20% bzw. 25% werden als geeignete Größenordnungen für den individuellen Leistungsanteil angesehen. Auf der institutionellen Ebene sollte man 20% Leistungsanteil nicht überschreiten (G). Hier äußert G spontan und klar, bezogen auf die Individualebene:

G: „75% als Grundgehalt.“ (113)

Bezogen auf die institutionelle Ebene (Studiengänge bzw. Fakultäten) schätzt G ein:

G: „Und da sollte man also achtzig Prozent nicht unterschreiten. Achtzig Prozent Grundzuführung sollte es geben.“ (114)

Ein/e andere/r Personalvertreter/in (S) lehnt gruppenbezogene Leistungsanteile – was auf der institutionellen Ebene automatisch der Fall wäre – grundsätzlich ab. Lediglich in bestimmten Fällen, z.B. bei Zielvereinbarungen, sei dies vorstellbar. Für die individuelle Ebene hält S jedoch 20% Leistungsanteil durchaus für angemessen.

S: „Dann meine ich, wenn das eingehalten ist, daß das Grundgehalt automatisch so hoch ist, daß er davon existieren kann. So daß der Leistungsanteil wirklich ein Anreiz ist für ihn, irgendwie in den Genuß von etwas Besonderem zu kommen.“ (134) „Ich könnte mir durchaus vorstellen, daß der Leistungsanteil etwa 1/5 sein könnte.“ (135)

S führt dann im Rahmen einer Retrospektive auf die DDR-Zeit hierzu weiter aus:

S: „Und unter dem Strich stand, dieser erhält tausend Mark und jener (kriegt) hundert Mark. (Wenn ich das jetzt einmal ganz extrem formuliere:) Es hat keiner etwas weggenommen bekommen. Aber es ist für jeden etwas draufgelegt worden. (...) Aber es war ein enormes Erlebnis für jeden, der dort saß. Der Eine hat einen roten Kopf bekommen und der Andere hat gestrahlt. Sie verstehen, wie ich das meine, oder? Es hat jeder irgendwie etwas daraus gelernt. Und im Endeffekt hat das vorwärts gebracht.“ (126) „Und ich könnte mir durchaus vorstellen, daß so etwas auch heute den Effekt bringen würde wie damals. (...) Es gab ja zwei verschiedene Sachen. Es gab die sog. Leistungsstufen, wo also während des nächsten Jahres jeder zum Gehalt eine bestimmte Stufe zusätzlich anerkannt bekam und es gab die Prämie. (...) Diese einmalige Auszahlung konnte auch manchmal für eine einmalige ganz enorme Leistung gezahlt werden.“ (127)

S favorisiert grundsätzlich Einzelzulagen, in bestimmten Fällen sind aber auch Gruppenzulagen vorstellbar, z.B. bei Zielvereinbarungen für Studiengänge/ Fakultäten (132):

S: „Wenn solche Leistungszulagen in irgendeiner Art und Weise einen Anreiz haben sollen, müßten sie eigentlich (...) bis auf den Einzelnen runter gehen. Wobei ich mir durchaus auch vorstellen kann, daß es für bestimmte Personengruppen solche Anreize geben kann“ (127)

Vertreter der Hochschulforschung äußerten sich für die individuelle Ebene in ähnlicher Weise wie die Personalvertreter/innen. Demnach sollte der Leistungsanteil ebenfalls ein Fünftel bis ein Viertel bzw. 20% bis zu 25% betragen, wobei sie teilweise (R und L) von einer schrittweisen Anhebung auf zunächst etwa 10% und erst in einem weiteren Schritt auf 20% ausgehen. 10% werden jedoch einhellig als ein auf die Dauer zu geringer Anreiz gesehen, da ein „wettbewerbsfähiges“ (M) bzw. den Lebensunterhalt (und den Status) sicherndes Grundgehalt (R, P) vorausgesetzt wird. Für die institutionelle Ebene wird nicht von allen Hochschulforschern differenziert, wobei nach dem Eindruck des Verfassers davon auszugehen ist, daß hier (zumindest) dieselbe Größenordnung geeignet scheint. Diejenigen, die sich explizit dazu äußerten, nennen z. T. sogar deutlich höhere Anteile wie 33 bis 50% (R).

R: „Also, das Grundgehalt: 90 oder 80 Prozent und das flexible dann 10 oder 20 Prozent.“ (152) Auf der Ebene der Studiengänge seien es: „50 zu 50 oder  $\frac{2}{3}$  zu  $\frac{1}{3}$ .“ (162)

L: „Jedenfalls sollte es einen Grundanteil geben, der Stabilität sichert. Erfahrungen gibt es bisher nur mit eher geringen Anteilen leistungsbezogener Finanzierung an Gesamtfinanzierung, die in der Größenordnung von 10% liegen. Im nächsten Schritt anzustreben wäre ein höherer Anteil von etwa 20%.“ (213)

### **Gewichtung der Anteile für Lehre, Forschung und weitere Aufgaben**

Im Unterschied zur Frage nach der Gewichtung des Leistungsanteils, die sich auf die Gesamtfinanzierung bezieht, ging es in der Frage zur Gewichtung von Lehre, Forschung und weiteren Aufgaben um die Gewichtung *innerhalb* des Leistungsanteils.

Die Vertreter/innen der Hochschulleitungen wollten sich auch hier kaum auf konkrete Gewichtungen festlegen. Daher wurde vom Verfasser in mehreren Gesprächen der Begriff des „Korridors“ eingebracht, d.h. eines Bereiches zwischen einem Maximum und einem Minimum, in dem sich aus Expertensicht die Anteile bewegen sollten. An der kleinen Universität wurde im Ergebnis ein Korridor für akzeptabel gehalten:

N: „Ich finde, eine Uni und eine Professur an der Uni hat sich darauf eingelassen und eigentlich auch einzulassen, daß die Lehre die Hälfte der Tätigkeit ausmacht und ich würde es fast zu defensiv finden zu sagen, man nimmt die Lehre raus. Also eine Universität ist eben eine Hochschule und kein Max-Planck-Institut. (...) Ich meine klar, man könnte natürlich pragmatisch sagen: Gut, wenn einer nicht lehren kann, dann muß man sich etwas überlegen. Ich würde aber nicht als Institution sozusagen nachgeben wollen, (denn) das finde ich eine ganz schwierige Kiste, weil das auch zwischen den Institutionen sehr starke Rivalitäten produziert.“ (85) „Ich würde schon sagen Korridor, aber beides muß sehr prominent vertreten sein... (bei der Gewichtung von Lehre und Forschung).“ (86)

An der großen Universität wird zwar eine Gleichgewichtung der Finanzierungsanteile von Forschung und Lehre favorisiert. Es werden aber mit mindestens 25% in erheblichem Ausmaß Finanzierungsspielräume für weitere Aufgaben bzw. „zusätzliche Ziele wie z.B. Wissenstransfer, Berufseinstiegsunterstützung u.a.“ (M) als sinnvoll erachtet. Dies würde

bedeuten, daß Lehre und Forschung mit je maximal 37,5% in die Finanzierung eingehen. Von der Hochschulleitung der in die Interviews einbezogenen FH wird eine Gleichgewichtung von Lehre, Forschung und Weiterbildung favorisiert (d.h. je rd. 33%).

Von den Personalvertretungen wird in einem Fall (S) zunächst eine rechtliche Grundlage eingebracht, wonach z.B. für ein Gehalt nach BAT 2a mindestens 50% der Arbeitszeit insgesamt für Forschung und Lehre aufgewendet werden müssen.

S: „...das muß ein Mitarbeiter bringen. Er muß mindestens 50 Prozent Forschung und Lehre machen, sonst kann er die 2a nicht kriegen. Das sind ja für uns dann immer so die Prämissen.“ (138)

Unabhängig von derzeit vorhandenen juristischen Hürden wird aber von S ein Korridor von z.B. 30 bis 70% für „begründete Forscher“ bzw. „begründete Lehrer“ als durchaus sinnvoll angesehen: „Warum nicht!“ (138). Von der Personalvertretung an der FH wird (analog des sehr viel höheren Lehrdeputates) eine Priorisierung der Lehre favorisiert: Hier sollte die Lehre für die Individualebene mit einem Gewicht von 60% eingehen. Für die institutionelle Ebene sollte es sogar mit 70% zugunsten der Lehre gewichtet sein:

G ist „nach wie vor der Meinung, die Lehre ist das, wofür es sich eigentlich ...<sup>146</sup>. Und die soll mindestens zu sechzig Prozent dazu beitragen. Zwanzig Prozent Forschung und zwanzig Prozent Beteiligung an Gremientätigkeit.“ (114) Bezogen auf die institutionelle Ebene meint G: „*Siebzig – dreißig*. Siebzig Prozent Lehre und dreißig Prozent Forschung.“(115)

Bei den Hochschulforschern wird grundsätzlich von einer gleich starken Gewichtung von Lehre und Forschung als gleichermaßen wichtige Aufgaben der Hochschule ausgegangen (wobei diese alle einen universitären Hintergrund haben). Dieser Gedanke findet sich bei allen Hochschulforschern im Verlauf der Gespräche. Hierbei werden jedoch z.T. weitere wichtige Aufgaben wie z.B. die Nachwuchsförderung (A) gesehen, die in einem Atemzug mit Lehre und Forschung genannt wurden. Von anderen Gesprächspartnern wurde dies vermutlich implizit der Lehre oder der Forschung zugeordnet. Zwei der interviewten Hochschulforscher, die diesen Aspekt noch einmal explizierten, halten eine stärkere Variabilität jedoch durchaus für sinnvoll. In einem Fall wird vorgeschlagen, die

---

<sup>146</sup> unverständlich (1 Teilsatz) vermutlich: lohnen soll

Detailregelung hierzu den Hochschulen selbst zu überlassen (A). Im anderen Fall wird das Gedankenspiel einer wettbewerblichen Verteilung des Lehrdeputates mittels Punktbewertung eingebracht (P), wobei von einer Größenordnung der Variabilität des Lehrdeputates von 12 statt 8 für „begnadete Lehrer“ bzw. von 4<sup>147</sup> statt 8 für „begnadete Forscher“ die Rede ist.

P: „Wir haben ja auch ein interessantes Modell gerade in der Diskussion von Niedersachsen in die Debatte geworfen, ob denn dieser Wert des Lehrdeputats wirklich so fix sein muß. Also, ob es nicht Hochschullehrer/ Hochschullehrerinnen geben kann, die zehn Stunden lehren statt acht oder zwölf statt acht und andere dann eben nur sechs oder vier.“ (236) „Und, daß dann auch gesagt wird, *Okay, der macht ja mehr Lehre*. Und in der Punktbewertung wird das angerechnet. Und der kann nicht so viel publizieren und so viele Drittmittel einwerben. Der kann ja dadurch wieder substituieren. Das finde ich ganz o.k. Ich würde das jetzt nicht unter dem Titel Lehrprofessur und Forschungsprofessur<sup>148</sup>. Ich denke schon, daß die Lehre sich aus der Forschung speisen sollte oder aus der Beteiligung an der Forschung. So etwas kann man freiwillig machen. Man kann aber die Lehrdeputate auch wettbewerblich verteilen. Man kann sagen, der Durchschnitt ist acht Stunden. Aber, ob man da jetzt genau acht Stunden macht oder mehr oder weniger, das entscheidet sich danach...“ (236)

Voraussetzung sei aber, daß eine exzellente Aufgabenerfüllung in der Lehre in demselben Umfang anerkannt und honoriert wird wie in der Forschung, um nicht „Lehrmuffel“ durch ein geringeres Lehrdeputat und damit größere Freiräume zu „belohnen“.<sup>149</sup>

---

<sup>147</sup> Dies gilt für Universitäten und ist bereits jetzt z.T. schon als Lehrdeputatsermäßigung bei der Übernahme von Gremienfunktionen der Fall, wobei die Erfüllung der Lehraufgaben dann von anderen Lehrenden der Institution mit höherem Lehrumfang wieder ausgeglichen werden muß. Für FH ist das doppelte Lehrdeputat üblich.

<sup>148</sup> Einschub: fassen

<sup>149</sup> Dieser Themenbereich wurde inzwischen auch mehrfach in öffentlichen Diskussionen aufgegriffen, so von Altbundeskanzler Helmut Schmidt in der „Zeit“ (Nr. 40/ 2001) oder vom Leiter des Centrums für Hochschulentwicklung (CHE), Detlef Müller-Böling, in der Frankfurter Rundschau vom 03.02.2004, Seite 31.

## **Geeignete Kriterien aus Sicht der Experten**

Diese Frage wurde gleich nach der Eingangsfrage (zur Grundposition) als zweite Frage im Gesprächsleitfaden plazierte und möglichst offen gehalten, um alle grundsätzlich in Frage kommenden Kriterien in die Diskussion einzubeziehen.

Eine Platzierung weiter hinten im Leitfaden bzw. zu einem späteren Zeitpunkt im Gespräch hätte evtl. die Art oder die Nennung bestimmter Kriterien vorbeeinflusst. Auch hier wurde gebeten, die Kriterien für die Individualebene und die institutionelle Ebene zu differenzieren.

Die Vertreter/innen der Hochschulleitungen bezogen neben Aspekten der Lehrqualität wie didaktischer Kompetenz (B), Aktualität (B, M) und Weiterentwicklung von Studiengängen (N) bzw. den curricularen Aspekt (M) auch Aspekte der Studienbedingungen wie Ansprechbarkeit der Hochschullehrer (B), Internationalisierung (B, M) und Gender-Mainstreaming (B) sowie Aspekte des Überganges vom Studium in den Beruf wie Praxiskontakte und Arbeitsmarktchancen (B) als Kriterien in ihre Überlegungen ein. Auch Aspekte aus Berufungsverfahren gelten als geeignet (M). Außerdem wurde betont, daß nicht Absolutwerte, sondern Relationen und Vergleiche erst wirklich aussagekräftig sind (N als Vertreter/in der Hochschulleitung der kleinen Universität), und daß immer sowohl quantitative als auch qualitative Kriterien einbezogen werden sollten:

N: „Relation ist immer das Vorher, wo diese Frage, ob jemand besonders gut oder besonders schlecht ist, einfach gar keine Rolle spielt. Diesen Vergleich dazu finde ich sinnvoller. (...)“ (63)

Der Aspekt, daß immer sowohl quantitative als auch qualitative Kriterien einbezogen werden sollten, und der Aspekt des Vergleiches wird auch von M (Vertreter/in der Hochschulleitung der großen Universität) aufgegriffen und am Beispiel der Internationalisierung erläutert. Schließlich wird von M darauf hingewiesen, daß auch gesellschaftliche Rahmenbedingungen, wie die „außeruniversitären Besoldungsskalen“ in die Überlegungen einbezogen werden müssten.

M: „Es ist sicherlich richtig, also Köpfe allein zählen. Das ist ein relativ grobes Instrument, aber es ist als solches auch nicht verzichtbar. Denn wenn wir an der Universität einen Anteil von zwei Prozent bei den (ausländischen) Dozenten haben, dann ist das eben, auch im kritischen Vergleich, ganz außerordentlich wenig – viel

zu wenig. (...) Im übrigen wird es ja sicherlich auch so sein, daß der Erfolg einer solchen Internationalisierungsstrategie nicht nur von Marketing abhängt, sondern eben auch von inhaltlich qualitativen Gesichtspunkten. Und dazu gehört der curriculare Aspekt unbedingt dazu. Dazu gehört auch der sprachliche Aspekt. Dazu gehört sicherlich auch das Umfeld, (das) was die jetzt als *Noneducational Services* bezeichnen.“ (93)

„Ja, Sie wollten das gerne noch ausdifferenziert haben für die Ebene der Besoldung. Das ist nicht nur eine Frage, die an die Hochschulen zu stellen ist, sondern an den öffentlichen Dienst generell. Das hat natürlich alle Probleme der Meßbarkeit und der Vergleichbarkeit von Leistungskriterien und mitunter spielen da auch noch gesellschaftliche Rahmenbedingungen eine Rolle. Es gibt Fächer, in denen – nehmen wir die ingenieurwissenschaftlichen Fakultäten – die außeruniversitären Besoldungsskalen ganz anders sind, als wenn man sich im Bereich der Sozial- und Geisteswissenschaften befindet. Also man kann dies auch nicht als isoliertes Problem der Universitätsbesoldung oder auch der Besoldung des öffentlichen Dienstes sehen, sondern man muß die gesellschaftlichen Rahmenbedingungen (vielleicht) noch stärker bedenken, als das bis jetzt der Fall gewesen ist. (...) Also man kann das noch kleinräumig ausdifferenzieren, aber grundsätzlich wollte ich darauf abheben, daß – weil hier wieder Beamte etwas für Beamte erdacht haben – dieser Aspekt vielleicht nicht ganz – und auch der internationale Aspekt, wenn man Mobilität fördern will – in der Weise mitbedacht worden ist, wie man dies für wünschenswert halten kann.“ (93) „Also mit diesem sowohl Steuerungs- als auch Bewertungsinstrumentarium stehen wir relativ am Anfang, wenn man das vergleicht mit (etwa) Dingen, die man schon immer gemacht hat. Ich meine, weil man sie im Berufungsverfahren braucht, nämlich die Bewertung einer individuellen Lehr- und Forschungsleistung. Dann können wir da sehr viel mehr, auch wenn es eine Reihe von Kollegen gibt, die immer abstreiten, daß man da (ja) vergleicht. Aber das tut man natürlich, denn sonst könnte man auch eine Münze werfen, wen man beruft. Das heißt, wir haben die Kriterien, man muß sie ein bißchen klarer stellen.“

Die Personalvertreter/innen beziehen sich v.a. auf die Individualebene, denn Gruppenprämie „das bringt überhaupt nichts. Und ist auch schwer vergleichbar“ (G). Hierbei wird von S auf die Beurteilungskriterien verwiesen, die für Beamte bereits in Form von

Dienstbeurteilungen festgeschrieben wurden und die man für Angestellte ebenfalls sukzessive einführen will. An dieser Stelle wird auch explizit deutlich gemacht, was „private“ Meinung des/r Experte/in ist und was die „allgemeine Meinung“ des Personalrates ist. So äußert sich S im Zusammenhang mit der Frage nach geeigneten Kriterien explizit als Privatperson und deutet an, daß sie/ er auch innerhalb des Personalrates bei dieser Frage eine Minderheitenmeinung vertritt und nicht in allen, aber wohl in einigen solchen Fällen auch die Diskussion innerhalb des Personalrates mit der „Privatmeinung“ beeinflusst hat. Daher erscheint dies als Hintergrund für die Einordnung der Äußerungen wichtig und muß bei der Interpretation der Interviewaussagen einbezogen werden:

S: „Das Jüngste sind die ganzen Beurteilungskriterien, die für Beamte in Form von Dienstbeurteilungen festgeschrieben sind und die man eben jetzt für Angestellte auch sukzessive einführen will. Den Kriterien hat der Hauptpersonalrat bereits zugestimmt – gezwungenermaßen. Er hat es lange hinausgezögert. Weil ja meine Meinung, die meine private ist, daß ich also seit langem was für Leistungsanreize übrig habe, durchaus nicht die allgemeine Meinung ist. Es kann ja dann auch bedeuten, denn manche können sich sehr gut einschätzen und wissen demzufolge, daß sie vielleicht dadurch doch etwas schlechter wegkommen würden. Und oft sitzen diese Leute an den Schalthebeln.“ (124) „Was für meine Begriffe notwendig wäre, ist, mal zu sagen: *Das ist der ideale Professor. Das ist der ideale Mitarbeiter.* (...) Und für diesen idealen, oder durchschnittlichen vielleicht besser (...) Also, was kann ich von einem Professor verlangen?“ (125) Dann folgt eine Retrospektive: „Ich gehe einmal in „alte Zeiten“ zurück. Jedes Jahr im September ungefähr (...) gab es in den Arbeitsgruppen, Lehrkollektiven große Diskussionsrunden, weil es am 7. Oktober immer Prämie gab. (...) Da bekam vornweg jeder drei, vier Blatt. Dort standen harte Zahlen. Wieviel Lehrveranstaltungen welcher Art. Wieviel Belegarbeiten usw. *Was wurde jeweils in der Forschung gemacht?* Wobei Forschung qualitative Ziele sind, Erfüllung qualitativer Ziele – im Endeffekt. Das ist nichts quantitatives mehr. Da stand zwar im Endeffekt auch etwa: *So und soviel Zeit habe ich für Forschung aufgebracht.* Aber ansonsten ist das qualitativ, weil der Chef ja gesagt hat, ob es gut oder (ob es) schlecht ist, ob die Ziele erreicht sind oder nicht.“ (126) Insgesamt sieht S qualitative und quantitative Kriterien beide als notwendig an: „Leistung, die dann quali- und quantitativ bewertet werden müßte - das ist sicherlich eine Einheit“ (128).

Außerdem wird als ein bisher noch nicht erwähntes Kriterium von G das Problem der Abstimmung von Lehrveranstaltungen genannt, das sicherlich den Aspekten der Studienbedingungen zuzuordnen ist.

G: „Also, was mir nicht gefällt ist, daß da zu wenig Abstimmungen in den Lehrveranstaltungen, in den angrenzenden Lehrveranstaltungen miteinander durchgeführt werden. Das heißt also, es gibt Überschneidungen zu bestimmten Problemen. Oder andere Dinge fallen dann unter den Tisch, die mitkommen müßten.“ (106)

Die Hochschulforscher/innen sprechen einhellig die Notwendigkeit quantitativer *und* qualitativer Kriterien an und sind optimistisch, daß diese einbezogen werden können. Als eher quantitative Kriterien wird z.B. die Anzahl der Studierenden (R) gesehen, oder wie viele Absolventen (A) bzw. wieviel Prozent (der Studierenden) erfolgreich eine formale Qualifikation erreichen (R). Dies wird z.T. auch als Absolventenquote bezeichnet (P). Hierbei würde allerdings vernachlässigt, was diejenigen dann tatsächlich wissen und außerdem könne passieren, daß bei nur wenigen solchen Indikatoren versucht wird, genau diese auf Kosten der Ausbildungsqualität zu maximieren:

R: „Also, was man bestimmt machen kann, ist qualitative und quantitative Indikatoren von Lehrqualität zu bestimmen. Also, so und soviel Prozent erreichen erfolgreich eine bestimmte formale Ausbildung, also eine Qualifikation. (...) Und dann müßte man noch einen objektiven Leistungstest haben. Was die, die erfolgreich abgeschlossen haben, dann *tatsächlich* wissen. (...) Und ein weiteres Problem ist, wenn man die Qualität der Lehre nicht symmetrisch abdeckt, also komplett dieses Phänomen gleich abdeckt, sondern nur bestimmte Indikatoren heranzieht. Dann könnte es sein, daß, wenn die Leistung in diesen Indikatoren Rückwirkung hat auf die Steuerung der Hochschulen, (dann) die Studiengänge oder die Institute oder die Professoren versuchen, diese Indikatoren zu maximieren. Nehmen wir einmal an, die Anzahl der Promotionen, pro Promotion würde man Geld bekommen an der Hochschule. Dann würde man die Promotionsquote eventuell auf Kosten der Qualität der Promotion erhöhen. Eine Ebene darunter, (auf) der (Ebene der) normalen Studienabschlüsse. Also wäre zumindest denkbar. Man müßte dann nicht nur die Anzahl messen, also nicht nur einen Indikator heranziehen, sondern wirklich die

Qualität symmetrisch abdecken, also die verschiedenen Indikatoren. Z.B. messen, wie gut diese Leute dann sind. (147)

Auch die Studiendauer wird angesprochen, wobei diese nur als eingeschränkt geeignet angesehen wird (R). Auf jeden Fall wären dann Umfeldbedingungen einzubeziehen, z.B. der Anteil von Frauen mit Kindern:

R: „Studiendauer würde ich schon als ein relevantes Kriterium heranziehen für die Qualität der Lehre. (...) hängen natürlich nicht nur von Qualitätsmerkmalen (ab), sondern auch von *Umfeldbedingungen* ab. Also, Studienzeiten: Frauen mit Kindern werden einfach länger brauchen zum Studieren. (...) Also, so etwas müßte man zumindest mitbeachten als Gewichtungsfaktor. *Anzahl der Kinder* müßte dann einen Bonus rechtfertigen können. (149)

Zudem müsse man sich die Frage stellen, inwieweit die Gründe für eine lange Studiendauer mit der Leistung zu verknüpfen sind. Dabei sind externe Rahmenbedingungen und individuelle Gründe der Studierenden hinzuzuziehen, da diese höchstens teilweise beeinflußbar sind.

A: „Studiendauer ist so ziemlich problematisch, muß ich eigentlich sagen, weil ich mich frage, inwieweit sind tatsächlich die Gründe für eine lange Studiendauer tatsächlich mit der Leistung des einzelnen Professors zu verknüpfen?<sup>150</sup> Also ich meine, prinzipiell ist es doch so, es gibt externe Rahmenbedingungen, die sind an den Hochschulen teilweise nicht besonders günstig. Aber die liegen sehr wenig, glaube ich, oder höchstens teilweise in dem, was ein einzelner Professor beeinflussen kann. Also müßte man schon mal die Voraussetzung haben, daß die Bedingungen, unter denen diese Leute arbeiten, für alle, die man sich da anschaut, gleich sind. Aber dann gibt es ja auch, sagen wir mal individuelle Gründe, die dazu führen, daß man vielleicht sein Studium verlängert.“ (176)

Also müßte man die Voraussetzung haben (oder z.B. durch Gewichtungsverfahren schaffen), daß die (Ausgangs)Bedingungen für alle gleich sind (A). Dennoch wäre die Studiendauer kein alleiniges oder zentrales Kriterium, sondern nur im Rahmen eines

Kriterienbündels geeignet (L). Außerdem wird als noch ungelöstes Problem angesprochen, ob man individuelle Talente (oder Ziele) oder den Grad der Erreichung bestimmter Standards als Maßstab nutzt. Letzteres sei zwar im Augenblick das gängige Modell. Dabei würde aber die individuelle Leistungssubstitution unterschiedlicher Menschen nicht beachtet (P). Schließlich sollten unterschiedliche Wirkungsintervalle bei den Kriterien beachtet werden (z.B. bei Absolventenquoten ein bis fünf Jahre) um zu sehen, ob es da im Zeitverlauf Entwicklungen gibt:

P: „...Bereich, den ich für wichtig halte, der sowohl die institutionelle als auch die individuelle Ebene betrifft. Das sind unterschiedliche Wirkungsintervalle. Bestimmte Dinge wie Absolventenquoten (oder so), die kann man innerhalb eines Jahres oder nach fünf Jahren, wenn man verlässlichere Daten haben möchte, so erheben und vergleichen und in das Verhältnis setzen zwischen verschiedenen Fachbereichen, aber auch innerhalb eines Fachbereiches, um zu sehen, wenn man das mithilfe eines Zeitstrahles macht, ob es da Entwicklungen gibt. (...) Also, insbesondere solche Sachen wie *der Erfolg der Absolventen auf dem Arbeitsmarkt*.“ (221)

Als eher qualitatives Kriterium (bei dem ebenfalls Wirkungsintervalle zu beachten seien) wird von mehreren Hochschulforschern der Erfolg bzw. die Chancen der Absolventen auf dem Arbeitsmarkt gesehen (P, R). Dies könnte das Problem unerwünschter Versuche der Maximierung einzelner quantitativer Indikatoren vermindern. Dabei wäre aber wiederum die Abhängigkeit beruflicher Chancen von der Arbeitsmarktsituation zu beachten (R), was über Arbeitsmarktstatistiken geschehen könnte.

### **Studenten- und Absolventenbefragungen als Kriterien**

Während die vorhergehende Frage möglichst offen gehalten wurde, um den Gesprächsverlauf zu diesem Thema nicht zu beeinflussen, ging es in dieser Frage zur Akzeptanz von Studenten- und Absolventenbefragungen als Kriterien darum, die Einschätzung der Experten hierzu explizit zu erfragen, sofern diese es nicht schon von sich aus angesprochen hatten.

Insgesamt kann vorausgeschickt werden, daß die Einbeziehung von Absolventenbefragungen zum beruflichen Verbleib bereits von mehreren Gesprächspartnern (B, R) von sich aus als Kriterium in die Überlegungen einbezogen wurden. Die didaktische Qualität der

---

<sup>150</sup> Hier wurde vom Interviewpartner offenbar die individuelle mit der institutionellen Ebene vermengt. In den

Lehre (B), die Ansprechbarkeit von Lehrenden (B, P) und die Abstimmung von Lehrveranstaltungen (G), was überwiegend mittels Studentenforschungen bewertet wird, wurden ebenfalls mehrfach von sich aus als Kriterium einbezogen. Allerdings wurden hierbei nicht explizit Studentenforschungen genannt.

Auf die explizite Frage, wie sie Studentenforschungen als Kriterium bewerten, schätzen die Vertreter/innen der Hochschulleitungen ein, Studenten können die Lehrqualität „sehr wohl beurteilen“ (B), und es gebe grundsätzlich, so N, „eine sehr große Zustimmung zu studentischen Befragungen“:

N: „...es ist unterschiedlich zwischen den Fachbereichen. Grundsätzlich gibt es aber eine sehr große Zustimmung zu studentischen Befragungen. Die Frage ist nur, welche Relevanz haben sie im Endeffekt, also werden sie eingesetzt, um zum Beispiel leistungsbezogene Mittelzuwendungen zu machen, was wir innerhalb der Uni nicht machen...“ (67) „es gibt eigentlich keine Resistenz gegenüber Studentenforschungen, außer bei Einzelnen.“ (67)

In ähnlicher Weise äußert sich M, da dies in den USA bereits seit langem der Fall sei und „es bringt das Hochschulsystem nicht zum Zusammenbruch“ (95). Für die Verwendung nicht nur zur Leistungsbewertung, sondern auch für Mittelzuweisungen wären aber Probleme der Vergleichbarkeit zu berücksichtigen, z.B. zwischen Pflichtveranstaltungen und Wahlveranstaltungen (M) und es werden (N) z.T. Akzeptanzeinschränkungen formuliert:

N: „Widerstände, die so geäußert werden, sind in der Regel (...), daß man praktisch jetzt die Studierendenbefragungen zum Maß aller Dinge macht...“

Absolventenforschungen gelten als Kriterium mindestens genauso geeignet, da es ehemalige Studierende sind und diese zudem schon einige Zeit Abstand zum Studium haben. Es werden nicht wie bei Studierendenbefragungen Einschränkungen der Verwendung formuliert. Ein/e Vertreter/in der Hochschulleitung (B) formulierte hierzu explizit Fragen, was Absolventenforschungen inhaltlich erfassen sollten:

---

bisher diskutierten Konzepten wird dies jedoch meist getrennt angesprochen.

B: „Also z.B.: Wo kommen die unter, wie gut kommen die unter? Kommen die in ihrem Fachgebiet<sup>151</sup>, unter oder woanders usf.? Gibt es da auch über solche Kontakte von Professoren und Professorinnen gute Praxisstellen, gute Diplomarbeiten usw.? Also ist das eine unmittelbare Wirkung eines Engagements eines Professors oder einer Professorin?“ (13)

Die interviewten Personalvertreter/innen sehen die Einbeziehbarkeit von Studentenerhebungen als Leistungskriterium für die Verteilung finanzieller Mittel etwas skeptischer. Als Bedenken genannt werden sowohl Probleme der rechtlichen Absicherung (fehlende Datenschutzverordnung zum Erhebungszeitpunkt in Sachsen), als auch der Einwand, die Ergebnisse könnten evtl. durch Einflüsse verzerrt sein, die nichts mit der Lehrqualität zu tun haben (G, S).

S: Es ist (...) sicherlich das am meisten subjektiv befrachtete Gebiet. Und das muß man berücksichtigen. Ich kann unmöglich, das ist meine Meinung, auf der Grundlage einer Beurteilung eines Studenten oder eines solchen Fragebogens sowohl eine Lehrveranstaltung (verdammten), als auch einen Lehrenden verdammten. (...) Also ich übertreibe jetzt. Das ist ja hier gestattet. Ich weiß, daß das gar nicht passieren kann.“ (130)

Absolventenerhebungen werden wie von den Vertreter/innen der Hochschulleitungen auch von den interviewten Personalvertreter/innen wie z.B. S stärker akzeptiert:

S: „Ich halte eigentlich viel von Absolventenerhebungen, wenn man sie ein Jahr später macht. Und zwar deshalb, weil die Leute dann erst einmal mitbekommen, was eigentlich notwendig gewesen ist in ihrem Studium. Was sie vielleicht als Studenten gar nicht so gesehen haben und was sie jetzt durch das Messen an der Praxis erst einmal merken, daß sie dort eigentlich viel zu wenig Wert draufgelegt haben.“ (130)

Dies müsse aber „gemeinsam mit Fachleuten“ geschehen, damit „nicht bloß *einmal* jemand angerufen wird für die Studieninhalte, sondern *vielfach* Leute, um die es eigentlich geht“, sich beteiligen können (G).

---

<sup>151</sup> Vgl. hierzu die Ausführungen zu Kriterien beruflichen Erfolges im Abschnitt zu Absolventenerhebungen, insbes. zur sogenannten Ausbildungsadäquanz.

Die interviewten Hochschulforscher sehen Studentenbefragungen insgesamt als ein geeignetes Kriterium an. Dabei werden sie z.T. als der wichtigste Indikator für einen bestimmten Aspekt gesehen, z.T. „nur“ als ein Punkt unter anderen.

Für Indikatoren, die Qualität messen (sollen) bzw. „qualitative Indikatoren“ wird eingeschätzt, bei der Qualität von Lehrveranstaltungen würde „das Studierendurteil auf jeden Fall das wichtigste sein“.

R: „Jetzt kann man natürlich (noch) qualitative Indikatoren heranziehen oder Indikatoren, die irgendwie Qualität messen. Und ein wichtiges Kriterium wäre natürlich die Qualität von Lehrveranstaltungen. Das sollte man auch heranziehen für Gehaltsverhandlungen, bzw. Gehaltsbemessungen, wie gut aus Sicht der Studierenden z.B. Strukturierung oder Lehrkompetenz oder der selbsteingeschätzte Lernerfolg durch eine Veranstaltung ist. (...) das Studierendurteil wird auf jeden Fall das wichtigste sein. (152)

Andere Hochschulforscher/innen formulieren hierzu:

A: „Frage, ob Lehrevaluation mit einbezogen werden soll, das denke ich schon. Ich denke "Ja", man sollte das mit einbeziehen, aber als einen Punkt unter anderen. (...) Es gibt in Baden-Württemberg ein Papier, im Prinzip eine Art Vorstufe des, was vielleicht mal Verordnungsentwurf wird, und da ist es so, daß man im Prinzip sowas wie Parameter vorgibt oder Themen. Man sagt also: Leistungen beziehen sich auf (und dann ist gesagt) Lehre, Forschung, Nachwuchsförderung usw. Und da (...) ist Lehrevaluation auch explizit genannt.“ (184)

In ähnlicher Weise wird dies auch von P formuliert:

P: „Studentenbefragungen halte ich für sehr wichtig, wenn man zugleich ihre Grenzen in Rechnung stellt. Studentenbefragungen sind in meiner Sicht Zufriedenheitsmessungen – in erster Linie. Studenten werden als „Experten ihrer Situation“ befragt, nicht als Experten für das Fach (...). Und insofern ist es eine Expertise, die da rauskommt. Und sie bewertet die Lehre nicht danach, ob die wirklich gut ist oder nicht, sondern bewertet sie danach, ob sie für mich als Studierender gut ist oder nicht. Und das hat viel damit zu tun, welche Erwartungen der oder die einzelne an das Studium stellt. (...) deshalb, denke ich, haben, bzw. müssen Lehrbewertungen zwei Komponenten haben. Also, die können über Punkte u.ä. durchaus

in Bewertungen einfließen, aber viel stärker sollte bei studentischen Lehrbewertungen im Vordergrund stehen, daß die Bewertungen eine solche Befragungsqualität haben, daß die Lehrenden daraus selbst etwas für die Entwicklung der Qualität der Lehre ziehen können.“ (224)

Als Einschränkungen werden formuliert: Erstens: Ergebnisse von Studentenforschungen seien von Erwartungshaltungen abhängig (P). Zweitens: Studentenforschungen seien dann etwas problematischer, wenn sie Selbsteinschätzungen darstellen, z.B. bezüglich der sogenannten sozialen Kompetenzen, die erst bei Absolventen mit den Anforderungen im Arbeitsleben kontrastiert werden können (L).

Obwohl die Hochschulforscher bereits die Studentenforschungen deutlich positiver einschätzten, äußern sie sich wie R zu Absolventenforschungen überwiegend noch einmal deutlich positiver:

R: „Das ist *natürlich* ein weiteres Kriterium. (...) Also, ich fände es gut, wenn man die Absolventen ein Jahr nach Abschluss ihres Studiums befragen würde, 1.) *ob sie berufstätig sind* und 2.) *wie sie retrospektiv die Qualität der Ausbildung beurteilen*. Und vielleicht noch 3.) *inwieweit sie Inhalte dessen, was sie gelernt haben jetzt in ihrem beruflichen Feld anwenden können*. Also, da hätten wir drei Indikatoren, die für aussagekräftig erachtet werden können“ (154).

Ein Hochschulforscher (P) sieht Absolventenforschungen als Kriterium weniger akzeptabel an als Studentenforschungen:

P: „Da bin ich gespalten. Also, einerseits finde ich es gut, daß man Lehrinhalte daraufhin abfragen sollte, inwiefern der einzelne das Gefühl hat, die haben etwas gebracht. Ich bin aber sehr skeptisch in Bezug auf die Antworten, die dann regelmäßig kommen: *Ja, unsere Ausbildung, die war nicht praxisbezogen genug*. ... Der Magisterstudiengang soll mich in die Lage versetzen, mich in einer vergleichsweise kurzen Zeit in ein beliebiges Gebiet zum Spezialisten machen zu können. (...) Also, zum Generalisten und Spezialisten meine ich nur, daß das didaktische Modell so sein müßte, daß man eine generalistische Ausbildung erhält und daneben eine Spezialausbildung. Denn nur, wenn man einmal eine Tiefenbohrung gemacht hat, weiß man, wie man dann eine Tiefenbohrung im Berufsleben anstellt.“ (228)

Es überwiegen bei den interviewten Hochschulforschern jedoch Einschätzungen wie die folgende, wobei der Zusammenhang zu den Studentenbefragungen noch einmal hergestellt wird: Absolventenbefragungen seien eine sehr gute Ergänzung zu den bisher verwendeten Indikatoren. Sie hätten jedoch den Nachteil, daß sie sehr aufwendig und Ergebnisse auf eine Situation bezogen sind, die sehr lange zurückliegt. Daher seien zur Einschätzung der aktuellen Situation Studentenbefragungen unverzichtbar. Von den drei genannten Befragtengruppen (Lehrende, Absolventen und Studenten) seien sie sogar am besten geeignet. Die hierfür genannten Argumente sollen daher noch einmal zusammengefaßt werden:

Von vielen Gesprächspartnern wird grundsätzlich die Argumentation geführt, daß eher „qualitative Indikatoren“ gefunden werden sollten in dem Sinne, daß der Aspekt der Qualität stärker erfaßt wird. Beispielsweise wurden hierfür die didaktische Qualität, die Ansprechbarkeit von Lehrenden (ergänzend zu Betreuungsrelationen) und die Abstimmung von Lehrveranstaltungen genannt.

Während einige Gesprächspartner keine konkreten Operationalisierungen nennen, beziehen andere hierfür Absolventen- und Studentenbefragungen konkret in ihre Überlegungen ein. Schließlich seien Studenten „Experten ihrer Situation“ und könnten daher die Qualität der Lehre am besten einschätzen. Es werden allerdings auch von diesen Befürwortern Aspekte genannt, die berücksichtigt werden sollten. So wird die Berücksichtigung des Einflusses unterschiedlicher Motivation der Studierenden gefordert, da diese eine „Verzerrung“ der Ergebnisse zur Folge haben könne. Für Absolventenbefragungen wurde gefordert, daß die Absolventen durch das Messen an der Praxis Gelegenheit haben sollten zu bemerken, ob sie evtl. im Studium auf bestimmte Aspekte zu wenig Wert gelegt haben. Es wurden meist auch Möglichkeiten angesprochen, die Befragungen so zu gestalten, daß die angesprochenen Probleme gelöst und die Grenzen der Aussagekraft der Ergebnisse in Rechnung gestellt werden. So sollten Absolventenbefragungen frühestens ein Jahr nach Studienabschluß durchgeführt werden. Bei Studierendenbefragungen sollte immer die Motivation (z.B. zum Besuch einer Lehrveranstaltung) mit erhoben und bei der Auswertung der Ergebnisse berücksichtigt werden. Hierzu wird auch auf einschlägige Fachpublikationen verwiesen, in denen sich weitere Möglichkeiten zur Lösung solcher Probleme finden – wie auch Ergebnisse von Analysen zur Zuverlässigkeit und Vergleichbarkeit von Studentenbefragungen.

#### **4.4.3 Soziologische Konzeptualisierung**

Bei der soziologischen Konzeptualisierung soll im Rekurs auf soziologisches Wissen ´das Gemeinsame im Verschiedenen´ begrifflich gestaltet und in die Form einer Kategorie ´gegossen´ werden. Die Abstraktionsebene, auf der man sich bei diesem Auswertungsschritt bewege, sei die der ´empirischen Generalisierung´ (vgl. Meuser/ Nagel 1994: 462).

Hierbei ist aber selbstverständlich zu beachten, daß es sich nicht um Ergebnisse im Sinne einer Hypothesenprüfung in demselben Sinne wie in der quantitativen empirischen Sozialforschung handeln kann, sondern eher um ein vertiefendes Eintauchen in den Forschungsgegenstand zum besseren Verständnis und ein Aufzeigen der möglichen Bandbreite angesprochener Aspekte bzw. um auffällige Gemeinsamkeiten in den Äußerungen der interviewten Gesprächspartner. Für dieses Erkenntnisinteresse (i.S. des „Kontextwissens“ nach Meuser/ Nagel) soll nachfolgend versucht werden, die Ergebnisse der Experteninterviews zu verallgemeinern und dabei ggf. Thesen zu formulieren, die - soweit möglich - an anderer Stelle einer empirischen Überprüfung im Sinne der quantitativen empirischen Sozialforschung unterzogen werden können. Einige der hier anhand der Auswertung der Experteninterviews formulierten Thesen mögen eventuell provokativ wirken. Der Verfasser ist jedoch der Überzeugung, daß gerade diese Thesen - die ja offenbar auch für die interviewten Gesprächspartner wichtige Themen sind - die Diskussion auch an anderen deutschen Hochschulen insgesamt bereichern könnten.

#### **Hochschulleitungen sehen sich eher als ´Mittler´**

Zur grundsätzlichen (Selbst-)Einordnung der Gesprächspartner/innen läßt sich festhalten, daß sich die interviewten Vertreter/innen der Hochschulleitungen eher als ´Mittler´ und weniger als ´Befürworter´ von Reformansätzen sehen, während sich die interviewten Hochschulforscher/innen und auch die Personalvertreter/innen offenbar stärker als ´Befürworter´ sehen. Dies zeigt sich auch bei mehreren anderen untersuchten Aspekten, z.B. bei der Art der Spreizung von Zulagen (auffällige Zurückhaltung der Hochschulleitungen) und bei der Einschätzung zur Stärke der Gewichtung des Leistungsanteils im Verhältnis zur Gesamtfinanzierung. Hierbei ist in Rechnung zu stellen, daß sich alle interviewten Personen bereits mit Systemen der Leistungsbewertung und Anreizsystemen auseinandergesetzt haben, die Zurückhaltung also wohl nur zu einem geringeren Teil aus der Unkenntnis einzelner Sachverhalte resultieren kann. Gerade Vertreter der Personalräte, die sich noch

vergleichsweise wenig mit dem Thema beschäftigt hatten, waren weniger zurückhaltend. Allerdings klang in einigen Gesprächen an, daß diese Themen sehr heikel seien und daher trotz der Zusicherung, bei der Auswertung vollständige Anonymität zu wahren, Bedenken hierzu bestanden. Daher hat sich ein/e Vertreter/in des Personalrates im Zusammenhang mit der Frage nach geeigneten Kriterien explizit als Privatperson geäußert, gleichwohl aber angedeutet, daß sie/ er innerhalb des Personalrates wohl in einigen solchen Fällen auch die Diskussion mit der „Privatmeinung“ beeinflusst hat. Dies muß bei der Interpretation der Ergebnisse einbezogen werden. Vor diesem Hintergrund läßt sich mit der entsprechenden Vorsicht und in der gebotenen Zurückhaltung formulieren, daß in Fällen, wo die Personalvertretung von relativ starken Persönlichkeiten ausgeübt wird (die sich nicht nur an den gegebenen oder vermuteten Meinungen der großen Masse des Personals orientieren, sondern trotz evtl. vorhandener Akzeptanzprobleme die Situation an der Hochschule mitzugestalten versuchen), Vertreter des Personalrates weniger als Mittler und stärker als Befürworter wirken können als Vertreter der Hochschulleitung. Dieses Ergebnis kann auch als Indiz für eine große Bedeutung des Individuums bzw. einzelner Personen in Reformprozessen gesehen werden. Denn sollte die hier vorgefundene Konstellation auch an einer größeren Zahl anderer Hochschulen vorhanden sein, dürfte dies nicht ohne Einfluß auf die Institutionalisierung von Leistungsbewertungs- und Anreizsystemen bleiben. Eine Schlussfolgerung aus einer solchen Konstellation könnte lauten, bei anstehenden Reformvorhaben nicht allein auf die Hochschulleitungen zu bauen, sondern auch andere Gruppen stärker als bisher einzubeziehen, wenn sie an deren Gestaltung mitwirken wollen.

### **Stärkere 'Leistungsorientierung' kann durchaus Akzeptanz finden, wenn der Gedanke der Fairneß und Leistungsgerechtigkeit ausreichend Berücksichtigung findet**

Der Grundgedanke einer stärkeren Leistungsorientierung in dem Sinne, 'Leistung soll sich auch lohnen' findet trotz relativer Zurückhaltung der Hochschulleitungen (die v.a. bei konkreten Aspekten auftrat) bei allen Gesprächspartnern einschließlich der Hochschulleitungen starke Akzeptanz. Dies zeigt sich deutlich bei den Äußerungen zur Art der Kopplung von Leistungsbewertung und -anreizen. So wurde an erster Stelle der Gedanke der 'Belohnung von Verbesserungen' (Bonusmodell) als geeignet angesehen und als Beispiele hierfür oft Zielvereinbarungen genannt. Der Gedanke einer direkten Kopplung im Sinne formelgebundener Mittelzuweisung wurde dagegen - zumindest als alleiniges Modell

– meist als weniger geeignet angesehen. Als Begründung wurde genannt: Dieses System würde dafür sorgen, daß 'die Starken noch mehr und die Schwachen noch weniger' bekämen und damit Ungleichheiten noch verfestigen. Hierbei spielt also der Fairneß- bzw. Gerechtigkeitsgedanke eine bedeutende Rolle. Auch in anderen Publikationen wurden der Fairneß-Gedanke (vgl. Kempen 2004, Jackson 1997: 103, Karpen 1991: 25ff.) bzw. der Gedanke der Leistungsgerechtigkeit (vgl. Felger/ Paul-Kohlhoff 2004: 89)<sup>152</sup> als wesentliche Aspekte für die Akzeptanz und Wirksamkeit von Anreizmodellen genannt. Als Arbeitshypothese könnte man also ableiten: Leistungsorientierung kann Akzeptanz finden, wenn der Fairneß- bzw. Gerechtigkeitsgedanke ausreichend Berücksichtigung findet.

Die Akzeptanz des Leistungsgedankens zeigt sich außerdem bei der Gewichtung des Leistungsanteils, da mit bis zu 20 bzw. 25% überwiegend relativ hohe Gewichtungen des Leistungsanteils im Vergleich zur bisher z.T. schon existierenden bzw. konkret geplanten Gewichtung als geeignet genannt werden.

Daß bei der Art der Spreizung von Zulagen bzw. Leistungsanteilen in den Äußerungen der Hochschulleitungen Zurückhaltung herrschte, steht dieser Einschätzung nicht unbedingt entgegen. Zeigt doch gerade dieser Aspekt, daß Zustimmung keineswegs unreflektiert - evtl. aufgrund „sozialer Erwünschtheit“ (die bei Experten aus der Wissenschaft ohnehin tendenziell weniger ausgeprägt sein dürfte) geäußert wird, sondern die Grenzen der Überschreitung des momentan als allgemein akzeptiert Angesehenen sehr wohl deutlich gemacht werden.

Dies deckt sich auch mit der absoluten Betrachtung der Ergebnisse zur Selbsteinordnung der Gesprächspartner, da sich niemand als Gegner einer stärkeren Leistungsorientierung einstuft, sondern alle diesem Gedanken mehr oder weniger stark zustimmen. Natürlich ist hierbei zu beachten, daß gerade in den Unterschieden des Grades der Zustimmung auch Unterschiede der Akzeptanz zum Ausdruck kommen. Aber eine solche Differenzierung muß in diesem Versuch der Generalisierung in Anlehnung an Meuser/ Nagel eher vernachlässigt werden.

---

<sup>152</sup> Leistungsgerechtigkeit wäre hierbei nach Felger/ Paul-Kohlhoff (2004: 89) abzugrenzen von sozialer Gerechtigkeit (z.B. in Form von Kinderzuschlägen als Entgeltbestandteile, die aber kaum noch eine Rolle in der Weiterentwicklung von Entlohnungen spielen) Unternehmenserfolgsgerechtigkeit (durch Unternehmenserfolgsbeteiligungen als Entgeltanteile) und Anforderungsgerechtigkeit (Arbeitnehmer mit höheren Anforderungen der Tätigkeit wären danach besser zu entlohnen). Eine umfassende Diskussion des Gerechtigkeitsgedankens soll hier nicht geführt und statt dessen auf die Diskussion um Rawls (2003) verwiesen werden

### **Einheit von Forschung und Lehre könnte weniger starr gesehen werden**

Die Gewichtung der Anteile für Forschung, Lehre und weitere Aufgaben innerhalb des Leistungsanteils war ein weiterer Aspekt, der in den Interviews ausführlich thematisiert wurde. Dabei wurde – bezogen auf die Ebene der gesamten Hochschule – sehr wohl an einem relativ gleichen Gewicht von Forschung und Lehre festgehalten. Auf der Individualebene wurde jedoch mehrfach für eine weniger starre Sicht der Einheit von Forschung und Lehre plädiert, was sich v.a. an der Festlegung der Lehrdeputate manifestierte. So wird eine Art Korridor von 70% Gewichtung zugunsten der Lehre innerhalb des Leistungsanteils für 'begnadete Lehrer' (und lediglich 30% für die Forschung bzw. weitere Aufgaben) oder 70% zugunsten der Forschung innerhalb des Leistungsanteils für 'begnadete Forscher' als geeignet angesehen. Allerdings wurde als Voraussetzung betont, daß Lehrleistungen gleichermaßen honoriert werden müssen wie die in der Forschung. Außerdem wurde mehrfach formuliert, daß auch die weiteren Aufgaben, wie Wissenstransfer, Berufseinstiegsunterstützung, Weiterbildung und Nachwuchsförderung künftig ein größeres Gewicht von z.B. 25% oder 33% innerhalb des Leistungsanteils insgesamt erhalten sollten, so daß für Forschung und Lehre ggf. nur noch je maximal 37,5% bis 33% verblieben.

### **Wenn Kriterien quantitative *und* qualitative Aspekte enthalten, wären unerwünschte Nebenwirkungen zu vermindern**

Welche Kriterien aus Sicht der Experten geeignet erscheinen, wurde wie im Leitfaden vorgesehen, mittels zweier Fragen getrennt erhoben. Die erste Frage wurde in der Formulierung bewußt offen gehalten, um möglichst viele in Frage kommende Kriterien zu erfahren. Die zweite Frage war konkret zur Vertiefung zweier ganz bestimmter Kriterien gedacht, da diese in den konzeptionellen Überlegungen des Verfassers eine bedeutende Rolle spielen und daher deren Einschätzung bei den Experten eruiert werden sollte.

Zur ersten, offenen Frage wurden von den Interviewten v.a. Aspekte der Lehrqualität und Didaktik, Aspekte der Studienbedingungen/ Studienorganisation und Aspekte des Überganges in den Beruf genannt. Ergänzend wurden in einem Fall auch Aspekte aus Berufungsverfahren genannt, z.B. vergleichende Gutachten, die sich als geeignet erwiesen hätten und daher in anderen Zusammenhängen ebenfalls Anwendung finden könnten.

Generell sollten – das wird von mehreren Gesprächspartnern betont – nicht Absolutwerte, sondern Relationen und Vergleiche bzw. die Entwicklung im Zeitverlauf (‘Wirkungsintervalle’) betrachtet werden, da diese aussagekräftiger seien.

Außerdem sollten nach Überzeugung der Gesprächspartner immer quantitative *und* qualitative Aspekte enthalten sein, um unerwünschte Nebenwirkungen zu vermindern, die durch den mehrfach befürchteten Versuch der Maximierung bestimmter quantitativer Kriterien entstehen könnten. Dies wurde auch in anderen Publikationen bereits beschrieben (vgl. Abschnitt zu Anreizmodellen). Quantitative Kriterien – hier wurden v.a. Absolventenanzahl bzw. -quote und Studiendauer genannt – seien zwar häufig leichter verfügbar, werden aber nur als eingeschränkt geeignet angesehen.

### **Als qualitative Kriterien sind auch subjektive Einschätzungen von Studenten und Absolventen geeignet**

Qualitative Kriterien im Sinne von Kriterien, die die Qualität eines Gegenstandes erfassen (wie z.B. subjektive Einschätzungen der Lehrqualität und Studienbedingungen sowie rückblickend zum Wert des Studiums), seien zwar (noch) nicht überall in ausreichendem Maß und nur mit etwas höherem Aufwand verfügbar (zu machen). Sie werden jedoch als sehr wichtige und sehr wohl geeignete Kriterien eingeschätzt, wenn diese im Zusammenwirken mit Fachleuten erhoben werden und hierbei die Umfeld- bzw. Rahmenbedingungen einbezogen werden. Aspekte, die nicht oder kaum durch die Lehrenden bzw. den Studiengang beeinflusst werden können und damit als potentiell die Ergebnisse „verzerrende“ Einflußfaktoren zu betrachten sind, müßten allerdings bei der Auswertung kontrolliert werden.

Die Antworten auf die erste, offene Fragestellung lassen zusammen mit den Antworten auf das teilweise nötige „Nachhaken“ (mittels der zweiten, expliziten Fragestellung zur Akzeptanz von Studenten- und Absolventenbefragungen als Kriterium) zusammenfassend folgenden Schluß zu: Studentenbefragungen sind bei den interviewten Experten als wichtiges Kriterium zur Einschätzung des Aspektes der Lehrqualität überwiegend akzeptiert, werden aber in einigen Fällen mit Einschränkungen zu ihrer Nutzung (als Indikator zur Verteilung finanzieller Mittel) versehen. Absolventenbefragungen werden fast einhellig stärker akzeptiert und hierzu weniger Einschränkungen formuliert (z.B. lediglich Befragung erst ein Jahr nach Studienabschluß).

## 4.5 Fazit

Kernaussagen, die sich aus der Auswertung der Experteninterviews ableiten lassen, sollen abschließend noch einmal zusammenfassend genannt werden:

- Stärkere Leistungsorientierung kann durchaus Akzeptanz finden, wenn der Gedanke der Fairneß und Leistungsgerechtigkeit ausreichend Berücksichtigung findet.
- Einheit von Forschung und Lehre könnte weniger starr gesehen werden.
- Wenn Kriterien quantitative *und* qualitative Aspekte enthalten, wären unerwünschte Nebenwirkungen zu vermindern.
- Als qualitative Kriterien könnten auch subjektive Einschätzungen von Studenten und Absolventen geeignet sein.

Diese Aussagen wurden teilweise (wenngleich meist in abgeschwächter Form) auch bereits in wissenschaftlichen oder hochschulpolitischen Publikationen formuliert. Unerwartet ist aber, daß nicht nur mit dem Thema befasste Hochschulforscher, sondern auch Personalvertreter Leistungsbewertungen und -anreize akzeptieren. Dies gilt bis zu einem gewissen Grad (der vermutlich die wahrgenommene Akzeptanzschwelle markiert) auch für Vertreter von Hochschulleitungen. Auch Voraussetzungen und Bedingungen, die für deren Einführung bzw. für die Akzeptanz bestimmter Kriterien (wie z.B. Nutzung der Ergebnisse von Absolventenbefragungen) erfüllt sein sollten, erscheinen umsetzbar.

Abschließend bleibt als Desiderat dieser Auswertungen zu formulieren, daß die Kombination der Ergebnisse der Experteninterviews und von Lehrendenbefragungen an den einbezogenen Hochschulen im Sinne der angestrebten gegenseitigen Ergänzung unterschiedlicher Perspektiven einen weiteren Beitrag zu dieser Diskussion liefern könnte. Während die Experteninterviews hierbei eher die Sicht von „Akteuren“ wiedergeben, sollen die Lehrendenbefragungen eher die Sicht der „Betroffenen“ erfassen. Aufgrund der unterschiedlichen Perspektiven und Erhebungsmethoden ist dabei kein direkter Vergleich der Ergebnisse von Experteninterviews und Lehrendenbefragungen im Sinne einer Überprüfung möglich. Die Ergebnisse der Experteninterviews können aber nachfolgend im Sinne einer explorativen Erkundung des Forschungsfeldes und der Möglichkeiten zur Weiterentwicklung von Leistungsbewertungen und Leistungsanreizen genutzt werden. Wenn die befragten Experten aus hochschulpolitischen, personalpolitischen oder fachwissenschaftlichen Erwägungen heraus bestimmte Möglichkeiten zur Weiterentwicklung als sinnvoll (oder zumindest einer

Überprüfung wert) befinden und sie nicht von vornherein aus sachlichen Gründen verwerfen, wird eine weitere wissenschaftliche Untersuchung als lohnenswert angesehen.

Ein Beispiel hierfür wäre, daß qualitative Kriterien aus Perspektive der Experten grundsätzlich als sinnvolle Ergänzung der bisher überwiegend verwendeten quantitativen bzw. „objektiven“ Kriterien angesehen werden, um unerwünschte Nebenwirkungen zu vermeiden. Nach Experteneinschätzungen könnten auch subjektive Bewertungen als qualitative Kriterien geeignet sein. Im nachfolgenden fünften Kapitel sollen daher verschiedene der angesprochenen Kriterien genauer auf ihre Eignung untersucht werden. Ob diese – wenn sie denn grundsätzlich geeignet sind – auch bei den Betroffenen Akzeptanz finden bzw. wie diese Akzeptanz erlangt werden könnte, wird im darauffolgenden sechsten Kapitel diskutiert und auf diese Weise die Perspektive der Betroffenen einbezogen.

## 5. Kriterien für Leistungsbewertung und Leistungsanreize

*In der deutschen Diskussion um Leistungsbewertung und Leistungsanreize wird wie auch in anderen Ländern häufig auf die Wichtigkeit der Kriterien bzw. Indikatoren hingewiesen. Im Abschnitt zu Leistungsanreizmodellen wurde bereits herausgearbeitet, daß neben der grundsätzlichen Angemessenheit v.a. Transparenz und Nachvollziehbarkeit wichtig sind. Dies gilt natürlich auch für die in solchen Modellen Verwendung findenden Indikatoren. Außerdem werden an diese weitere Anforderungen gestellt. Beispielsweise sollen mit ihnen umfassend alle wesentlichen Dimensionen des betrachteten Gegenstandes erfasst werden und sie sollen methodischen Ansprüchen wie Validität und Zuverlässigkeit genügen. Nachfolgend wird zunächst ein Überblick über solche Anforderungen gegeben. Anschließend wird diskutiert, inwieweit bisher häufig Verwendung findende sogenannte „objektive“ Kennziffern oder Indikatoren solchen Anforderungen genügen. In weiteren Abschnitten werden Möglichkeiten zur Erhebung „subjektiver“ Indikatoren vorgestellt und diskutiert, inwiefern diese den genannten Anforderungen genügen können. Hierbei kann keine umfassende und detaillierte methodische Diskussion beispielsweise einzelner Frageformulierungen geführt werden. Vielmehr soll v.a. ein Überblick gegeben werden und die Frage erörtert werden, ob solche z.B. mittels Befragungen zu erhebenden „subjektiven“ Indikatoren grundsätzlich geeignet sein können.*

Gliederung:

### 5. Kriterien für Leistungsbewertung und Leistungsanreize

#### 5.1 Kriterien, Kennziffern und Indikatoren

5.1.1 Unterscheidung zwischen Kriterien, Kennziffern und Indikatoren

5.1.2 Anforderungen an Kriterien, Kennziffern und Indikatoren

#### 5.2 Die Eignung „objektiver“ Kennziffern und Indikatoren

5.2.1 Objektive Kennziffern und Indikatoren - Ersatz subjektiver Bewertung?

5.2.2 Studenten- und Anfängerzahlen als Orientierungshilfe, nicht als Leistungsindikator

5.2.3 Anzahl (abgelehnter) Bewerber – Indiz für Attraktivität des Studienganges?

5.2.4 Betreuungsrelation als ein Indikator für Überfüllung

5.2.5 Durchschnittliche Fachstudiendauer - ein Indikator für Lehrqualität?

5.2.6 Studierbarkeit – besseres Kriterium als die Studiendauer?

5.2.7 Abschlußnoten: Kein Kriterium für Studienerfolg, aber Orientierungshilfe

5.2.8 Erfolgsquote bei Prüfungen: Qualitätskriterium eines Studienganges?

5.2.9 Absolventenquote als Indikator für „Produktivität“?

5.2.10 Fazit zur Aussagekraft „objektiver“ Kennziffern und Indikatoren

#### 5.3 Subjektive Bewertungen als Leistungskriterien?

5.3.1 Die Bedeutung subjektiver Bewertungen für die Qualitätsorientierung

5.3.2 Möglichkeiten zur Erhebung subjektiver Bewertungen

#### 5.4 Die Eignung von Studentenbefragungen zur Qualität der Lehre

5.4.1 Wie kann Lehrqualität mittels Studentenbefragungen „gemessen“ werden?

5.4.2 Die Datenqualität von Ergebnissen studentischer Bewertungen der Lehre

5.4.3 Vergleichbarkeit der Ergebnisse von Studentenbefragungen

5.4.4 Weitere Anforderungen an Studentenbefragungen als Leistungsindikator

5.4.5 Fazit zur Eignung von Studentenbefragungen

#### 5.5 Die Eignung von Absolventenbefragungen zur Messung der Wirkungen der Hochschulausbildung

- 5.5.1 Wie können Wirkungen der Hochschulausbildung mittels Absolventenstudien „gemessen“ werden?
- 5.5.2 Die Datenqualität der Ergebnisse von Absolventenbefragungen
- 5.5.3 Vergleichbarkeit der Ergebnisse von Absolventenbefragungen
- 5.5.4 Weitere Anforderungen
- 5.5.5 Fazit zur Eignung von Absolventenbefragungen
- 5.6 Lehrendenbefragungen als Bestandteil der Qualitätsorientierung
  - 5.6.1 Die Datenqualität der Ergebnisse von Lehrendenbefragungen
  - 5.6.2 Vergleichbarkeit
  - 5.6.3 Weitere Anforderungen
  - 5.6.4 Fazit zur Eignung von Lehrendenbefragungen

## 5.1 Kriterien, Kennziffern und Indikatoren

### 5.1.1 Unterscheidung zwischen Kriterien, Kennziffern und Indikatoren

Zunächst soll auch hier eine genauere Klärung von Begriffen vorgenommen werden, da sie häufig synonym verwendet werden und dies z.T. zu Verwirrungen führen kann. Während im Abschnitt zu Leistungsanreizmodellen in Anlehnung an die entsprechende Literatur noch relativ allgemein von „Parametern“ die Rede war, soll es nun darum gehen, diese zu konkretisieren, denn letztlich müssen die Indikatoren eindeutig operationalisiert bzw. operationalisierbar sein.

Als *Kriterien* werden hier – wie es der alltagssprachlichen Diskussion an den Hochschulen wohl am ehesten entspricht – wie beim Begriff Parameter ebenfalls noch relativ allgemein verstanden, welche Aspekte als Meßlatte für Leistung bzw. Qualität dienen sollen.

Zur Unterscheidung von Kennziffern und Indikatoren weist Pasternack (2004: 111) auf die Differenz zwischen beiden Begriffen hin, „die im Kontext von Qualitätsorientierung zentrale Bedeutung hat (...): *Kennziffern* drücken schon vorhandene Quantitäten aus und führen zu absoluten Aussagen. (...) *Indikatoren* dagegen werden über ein theoretisches Konzept definiert, (wobei aus Kennziffern durch ein solches Konzept Indikatoren werden können).“ Die zugrunde liegenden Konzepte hierbei könnten dabei z.B. Effizienz, Effektivität oder Qualität sein (vgl. Pasternack 2004: 111). Sie sind damit keine neutralen Realitätsbeschreibungen, sondern selektieren, transformieren und kombinieren Daten, basierend auf normativen oder definatorischen Vorgaben. Sie sind „Maßgrößen, die die Realität lediglich ausschnittsweise bzw. stellvertretend abbilden wollen“ (vgl. Weber 1996: 81). Außerdem sollen Indikatoren auch Schlüsse auf nicht unmittelbar wahrnehmbare

Phänomene zulassen, indem sie als Hilfsgröße direkt wahrnehmbare Phänomene heranziehen (vgl. Schmidberger 1994: 297).

Beispielsweise wäre die Anzahl der Absolventen eines Studienganges in einem Studienjahr, da sie als absolute Aussage einzuordnen wäre, zunächst nur eine Kennziffer. Kombiniert man diese Anzahl der Absolventen jedoch mit der Anzahl der Studienanfänger in diesem Studiengang, indem die Relation Absolventen je Anfänger berechnet wird, so ist dies keine absolute Aussage mehr. Wird diese Relation dann z.B. in Prozentwerten ausgedrückt und zudem mit der wertenden Aussage verknüpft, daß eine sehr geringe „Produktivität“ eines Studienganges von z.B. um die 10% über mehrere Studienjahre hinweg gesehen eine Verschwendung von Steuergeldern und Lebenszeit von Studierenden bedeute, so bezeichnet diese Relation einen Indikator.

Abschließend soll erwähnt werden, daß noch zwischen verschiedenen Arten von Indikatoren unterschieden werden kann, „bei Leistungsindikatoren für die Universität mindestens zwischen qualitativen, quantitativen, absoluten und relativen Indikatoren“ (vgl. Sendldorfer 1997: 64). Andere unterscheiden (zusätzlich) in objektive und subjektive Indikatoren (vgl. Sendldorfer 1997: 64, Seidenschwarz 1992: 156ff.). In diesem Kapitel soll lediglich eine Gliederung in objektive und subjektive Indikatoren erfolgen (Studenten- und Studienanfängerzahlen werden hier von vornherein nur als Kennziffern aufgefasst). Hierbei stellen als objektiv bezeichnete Indikatoren fast immer gleichzeitig quantitative Indikatoren dar (Ausnahme: Noten – vgl. Pasternack 2004: 120, Seidenschwarz 1992: 156ff.). Außerdem sind fast alle der diskutierten objektiven Indikatoren auch relative Indikatoren (Ausnahme: Anzahl abgelehnter Bewerber, solange nicht als Relation Anzahl der Bewerber pro Studienplatz verwendet).

Da für die Einbeziehung in Leistungsbewertungen und Leistungsanreizmodelle eine Einordnung in entsprechende theoretische Konzepte und wertende Aussagen unerlässlich sind, soll in folgendem Abschnitt zunächst beschrieben werden, welche Anforderungen an Kriterien bzw. Indikatoren gestellt werden.

### **5.1.2 Anforderungen an Kriterien, Kennziffern und Indikatoren**

Zur Wichtigkeit der Kriterien bzw. Indikatoren wird einerseits betont, daß mit der *Auswahl* der Kriterien bzw. Indikatoren bereits eine Steuerungswirkung (oder auch Fehlsteuerungswirkung) impliziert ist. Deshalb müssen sie zentrale Aspekte - insbesondere Wirkungen,

aber auch Nebenwirkungen – messen (vgl. Buschor 2002: 74, Schenker-Wicki 2002: 34 sowie Abschnitt zu Leistungsanreizmodellen in dieser Arbeit). Das Meßsystem solle zudem vollständig sein. Wenngleich diese Anforderung meistens eine Illusion sei, so Buschor, gelte sie doch als Meßlatte für Indikatorensysteme.

Andererseits wird auf die Wichtigkeit der *Qualität* der Daten bzw. deren Validität hingewiesen (vgl. ebenfalls Buschor 2002: 74, Sendldorfer 1997: 65 u.a.). Bülow-Schramm (vgl. HRK 2003a: 32) betont dies auch noch einmal explizit für Varianten der Verknüpfung von Leistungsbewertung und -anreizen: „Das immer wieder beschworene Verhältnis von Evaluation und Mittelverteilung hat ja nur dann eine realistische Basis, wenn tatsächlich valide Daten im Evaluationsverfahren gesammelt werden konnten“.

Einen Überblick über weitere wesentliche Anforderungen an Indikatoren in Evaluation und „New Public Management“, wie sie in der Schweiz und im deutschsprachigen Raum Anwendung finden, gibt Buschor (2002: 74): So wird der Aspekt der Zeitnähe als zentrale Anforderung an Indikatoren genannt. Außerdem müssten Indikatoren auch Vertraulichkeitsaspekte berücksichtigen (indem sie die Persönlichkeitsschutzvorschriften beachten). Bei allen genannten Anforderungen müsse aber auch der Erhebungsaufwand beachtet werden. Dieser solle nicht größer sein als der Beitrag zur Lösung von Führungsaufgaben, so Buschor abschließend. Was Buschor hier nicht noch einmal explizit erwähnte (dieser Anspruch steckt jedoch in der eingangs erwähnten Erfassung von Wirkungen und Nebenwirkungen), ist, daß in die Betrachtung des Erhebungsaufwandes im Verhältnis zum Nutzen auch mögliche Folgekosten von Fehlanreizwirkungen enthalten sein sollten.

Ziegele (2002: 1) nennt als Anforderungen, die teilweise denen von Buschor entsprechen:

- „Indikatoren müssen praktikabel sein. Hierzu gehört (...) die schnelle und problemlose Verfügbarkeit, Vergleichbarkeit und die Eindeutigkeit der Indikatoren.“
- „Gleichzeitig müssen die Indikatoren Anreize schaffen, die Freiräume bei der Mittelverausgabung sinnvoll zu nutzen.“
- „Indikatoren dürfen keine unerwünschten Anpassungseffekte von Hochschulen hervorrufen.“

Von Buschor werden zudem noch als Grenzen von Leistungsindikatoren formuliert, daß sie nicht auf allen Gebieten gleichwertig und zielführend quantifiziert werden können. Außerdem würden Indikatoren und selbst auch Evaluationen allein nicht genügen. Es sei auch Führung wichtig.

Weiterhin nennt Buschor (2002: 76) neben Fehlanreizwirkungen auch „Kurzichtigkeit“ (Vernachlässigung der Betrachtung langfristiger Wirkungen) und die Gefahr der „bewußten Irreführung“ als typische Gefahren der Anwendung von Leistungsindikatoren. Letzlich wird als möglicher Ausweg vorgeschlagen, Defizite der „reinen Leistungsorientierung“ durch die Integration von Ansätzen des Qualitätsmanagements zu ergänzen (vgl. Buschor 2002: 80). Das entspricht der in dieser Arbeit vertretenen Auffassung, daß neben quantitativen Aspekten der Leistung auch deren Qualität berücksichtigt werden sollte.

Während Buschor die Diskussion um die Wichtigkeit und die Anforderungen an Leistungsindikatoren in der öffentlichen Verwaltung anhand seiner Erfahrungen im Bildungs- und Gesundheitswesen zusammenfasste, führte Grüning (2002) für den Bereich der Privatwirtschaft in Deutschland eine Studie durch, in der er die Sicht der Akteure in den Unternehmen empirisch untersuchte. Angesichts der im Abschnitt zu Leistungsanreizen dargestellten Entwicklungen, daß ursprünglich aus der Privatwirtschaft stammende Modelle z.T. für den Hochschulbereich adaptiert werden, erscheinen die Ergebnisse dieser Studie als Hintergrundinformationen für das Ausmaß der Wichtigkeit von Indikatoren interessant:

Grüning (2002: 206) beschreibt in seiner Studie, welche Voraussetzungen in Unternehmen als erfolgskritisch angesehen werden, die bereits Performance Measurement Systeme eingeführt haben: Auf Rang zwei wird dort (bei 64 von 108 Unternehmen fast gleichauf mit der Mitarbeiterbeteiligung auf Rang eins) die Art und Ausführung der Meßkonstruktion gesehen, „wobei vorwiegend die Güte der Indikatoren (insbesondere Validität) und deren Zeitnähe als bedeutend angesehen werden“ (vgl. Grüning 2002: 206).

Bisher wurde an den Hochschulen als Anforderung an Kriterien bzw. Indikatoren vor allem deren grundsätzliche Angemessenheit sowie deren Validität und Zuverlässigkeit diskutiert. Deren Zeitnähe spielte bislang noch keine große Rolle. Es erscheint jedoch schwer vorstellbar, daß ein Arbeiten mit günstigstenfalls zwei bis drei Jahre alten Daten, wie es z.T. an einigen Hochschulen anzutreffen ist, für die zukunftsorientierte Steuerung einer Hochschule effektiv sein kann. Daher ist zu erwarten, daß dieser Aspekt um so mehr an Bedeutung gewinnt, je stärker Indikatoren tatsächlich zur Steuerung herangezogen werden und deshalb wird hier auch die Zeitnähe als wichtige Anforderung an Indikatoren gesehen.

Kurz zusammengefasst werden daher insgesamt als Anforderungen an Indikatoren formuliert: Ausreichende Datenqualität, Praktikabilität und dabei insbesondere Vergleichbarkeit, hohe Zeitnähe, sinnvolle Anreizwirkung, Vermeidung unerwünschter Anpassungseffekte und ein vertretbares Verhältnis von Aufwand und Nutzen.

## 5.2 Die Eignung „objektiver“ Kennziffern und Indikatoren

Können sogenannte "objektive" Kennziffern und Indikatoren die subjektiven Eindrücke der Studierenden objektivierend untermauern? Lässt sich die Zuverlässigkeit der studentischen Einschätzung ihrer Lehr- und Studiensituation gar mittels solcher Kennziffern und Indikatoren überprüfen? Welche Aussagekraft haben z.B. Betreuungsrelation, Studiendauer, Notendurchschnitt, Erfolgs- und Absolventenquote für die Lehrqualität? Nachfolgend soll die Aussagekraft der am häufigsten verwendeten Kennziffern und Indikatoren und ihre Brauchbarkeit für die Einschätzung der Lehrqualität anhand ihrer methodischen Grundlagen sowie an konkreten Beispielen aus einem seit dem Jahr 2000 laufenden Projekt zur Erfassung und Aufbereitung von Daten aus den sächsischen Lehrberichten (vgl. Krempkow/König 2004) diskutiert werden. Eine Diskussion aller in Deutschland potentiell verwendbaren Kennziffern und Indikatoren kann an dieser Stelle nicht erfolgen. Hierzu sei z.B. auf die Publikation von Webler u.a. (1993) verwiesen, die mögliche Indikatoren für die Darstellung der Situation der Lehre in Lehrberichten und ihre empirischen Grundlagen in größerer Breite darstellt. Weitere mögliche Indikatoren, die in neuerer Zeit angeführt wurden und hier aufgrund bislang nur weniger vorliegender Erfahrungen mit ihrem Einsatz nicht diskutiert werden, sind auch Gleichstellungskriterien wie z.B. eine Frauenquote (vgl. HIS 2004: 30ff., Minssen/ Wilkesmann 2003: 109, HRK 2001) oder Kriterien für Internationalität wie die Anzahl internationaler Studierendenaustausche (vgl. Leszcensky u.a. 2004: 195, Ziegele 2003: 18).<sup>153</sup>

### 5.2.1 Objektive Kennziffern und Indikatoren - Ersatz subjektiver Bewertung?

Oft wird darauf hingewiesen, daß die Eindrücke z.B. von der Überfüllung an einem Fachbereich oder in einem Kurs im Vergleich verschiedener Hochschulen und innerhalb einzelner Hochschulen und Fachbereiche sehr unterschiedlich ausgeprägt sein können und nicht zuletzt auch von der Erwartungshaltung bzw. vom "gewohnten Leidensdruck" der Studierenden abhängen. Deshalb wäre es denkbar, den (inter)subjektiven Eindrücken der Studierenden von der Überfüllung an ihrem Fachbereich „objektive“ Kennziffern und Indikatoren gegenüberzustellen. Eine Möglichkeit, die Hinweise auf die "tatsächliche"

Überfüllung geben können, wären beispielsweise eine Erfassung der Teilnehmerzahl zur Bestimmung der Kursgrößen, das zahlenmäßige Betreuungsverhältnis Dozenten zu Studierenden und die Anzahl der Studierenden an ihrem Fachbereich. Dies, so hoffen die Verfechter solcher Kennziffern und Indikatoren, würde die subjektiven Eindrücke der Studierenden und auch der Professoren erst in ein richtiges Licht rücken und dabei entweder "objektivierend unterstützen" oder relativieren. Versucht wurde dies sowohl von Hornbostel & Daniel im Spiegel-Ranking (1993) als auch im Focus-Ranking (1997) und in den Studienführern des CHE (seit 1998 jährlich). Auch die Kultusministerkonferenz (KMK) legte 1995 einen Katalog mit als objektiv bezeichneten Kennziffern und Kriterien vor.

Jedoch stießen alle Autoren, die sich empirisch mit solchen Kennziffern auseinandersetzten, auf enorme Schwierigkeiten, die ihnen vorliegenden Daten vergleichbar zu machen. Hierauf verweisen beispielsweise Webler u.a. (1993) in ihren Empfehlungen zur Erstellung von Lehrberichten. Darin setzten sie sich intensiv mit der Verwendung von Leistungskennziffern und Indikatoren zur Messung von Lehrqualität auseinander und kamen zu dem Schluss: "Selbst bei einwandfrei erhobenem Datenmaterial sollte der Versuch von Leistungsvergleichen zwischen Hochschulen bzw. Fachbereichen nur von geeigneten Personengruppen mit eingehender Methodenkenntnis vorgenommen werden, um unzutreffende Schlußfolgerungen aus fehlerhaft interpretiertem Material zu vermeiden." (vgl. Webler u.a. 1993: 38) Doch die Schwierigkeiten beginnen häufig bereits bei der Qualität, insbesondere der Validität, der zur Verfügung stehenden Daten, die erst die Grundlage für Schlussfolgerungen und Interpretationen wäre. "Es muss versucht werden, die methodischen Grundlagen z.B. der genannten Indikatoren<sup>154</sup> (Studenten- und Absolventenzahlen, Studienabbrecher, Fachstudiendauer, Examensnoten) zu revidieren, um sie aussagefähiger für lehrbezogene Evaluationen zu machen..." (vgl. Webler u.a. 1993: 37). Dies mag das Beispiel der Gegenüberstellung von Studienanfängern und Absolventen zur Bildung einer *Abbruchquote* zeigen, obwohl hierbei lediglich die "Schwundquote" bzw. die "Absolventenquote" erfasst wird. Will man so den Abbrecheranteil ermitteln, ist dies solange nicht valide, wie nicht auch die Wanderungsbewegungen von Studierenden zwischen den einzelnen Hochschulen und Fächern berücksichtigt werden. Will man dagegen

---

<sup>153</sup> Weitere mögliche Indikatoren wurden auch in den Experteninterviews genannt, z.B. für Internationalität der Anteil ausländischer Studenten und Dozenten (vgl. Abschnitt 4.4.2 in dieser Arbeit).

<sup>154</sup> Webler u.a. verwendeten hier den Begriff Indikatoren. Das Zitat wurde im Original belassen, obwohl die genannten Kriterien hier auch als Kennziffern angesehen werden könnten, soweit keine normative Festlegung zu ihrer Bewertung erfolgte.

den „Schwund“ insgesamt erfassen, lassen sich u. U. durchaus zutreffende Schlußfolgerungen aus einer solchen Gegenüberstellung ableiten. Aber dieses Beispiel sollte nur der Verdeutlichung der möglichen Probleme dienen. Zunächst soll genauer auf Studenten- und Studienanfängerzahlen und das Betreuungsverhältnis als Kennziffern bzw. Indikatoren eingegangen werden.

### **5.2.2 Studenten- und Studienanfängerzahlen als Orientierungshilfe, aber nicht als Leistungsindikator**

Vielleicht mag es denjenigen, welche die einschlägigen empirischen Studien zu den Gründen für die Wahl einer Hochschule in Deutschland kennen, abwegig erscheinen, die Anzahl von Studenten als *Leistungsindikator* aufzufassen. Sie werden dennoch häufig als solcher diskutiert (vgl. z.B. Minssen/ Wilkesmann 2003: 108, Rammelt 2003, Hartwig/ Küpper 2002: 12, aber auch bereits Seidenschwarz 1992: 156ff).

In anderen Ländern (wie Großbritannien und Australien) mit stärker ausgeprägten Stipendiensystemen (als ökonomische Grundlage für eine tatsächliche Wahlmöglichkeit der Hochschule für Studierende) und einem größeren Maß an Transparenz der Hochschulleistungen (als Grundlage für ein wettbewerbsähnliches Verhalten) erscheint dies eher als angemessen, zumal Studentenzahlen in gewisser Weise auch die Funktion eines in diesen Ländern nicht existierenden Sockelbetrages mit übernehmen müssen (vgl. Leszcensky u.a. 2004 bzw. die Übersicht zur Gliederung leistungsorientierter Mittelvergabe ausgewählter OECD-Länder in dieser Arbeit).

In der deutschen Diskussion wird zudem häufig nicht zwischen Leistung und Auslastung bzw. „Lehrbelastungen“ differenziert (vgl. Pasternack 2004: 117). Denn sowohl nach den weit verbreiteten als auch nach dem in dieser Arbeit zugrundegelegten (stärker auch den Qualitätsaspekt einbeziehenden) Leistungsbegriff kann die reine Anzahl immatrikulierter Studierenden Leistung und erst recht Qualität nicht valide erfassen und daher kein Leistungsindikator, sondern lediglich eine Kennziffer sein. Ziegele (2002: 1) weist in diesem Zusammenhang darauf hin: „Probleme treten immer dann auf, wenn Output-Indikatoren gemessen werden, die eigentlich Inputs im Produktionsprozeß Hochschule darstellen (...)“. Als unerwünschte Anpassungseffekte würde zudem generell hervorgerufen, daß eine Senkung der Qualität der von den Hochschulen erbrachten Leistungen stattfinden könne, wenn mengenbezogene Indikatoren gewählt werden. Dies gilt aber nicht nur für die

Verwendung der Studierendenzahl als Leistungsindikator, sondern z.B. auch für die Absolventenzahl. Bei der Absolventenzahl besteht allerdings ein Anreiz, die Studierenden zum Abschluß zu bringen (vgl. Leszcensky u.a. 2004: 194).

Einmal davon abgesehen, daß die Bezeichnung „Produktionsprozeß Hochschule“ strittig ist, beschreibt Ziegele Effekte, die so bereits in anderen Ländern im Hochschulsystem auftraten, so z.B. in Großbritannien (vgl. Leszcensky u.a. 2004, Orr 2001).

Daneben gibt es jedoch weitere Gründe, die nicht nur gegen eine Verwendung von Studentenzahlen als Leistungsindikator in Mittelverteilungsmodellen sprechen, sondern auch gegen ihre Verwendung als Kennziffer bei der Leistungsbewertung an Hochschulen: Studien zur Hochschulwahl in Deutschland zeigen, daß hier meist die „Nähe zum heimatischen Herd“ (vgl. Müller-Böling u.a. 2001: 7), oder die Attraktivität des Studienortes (Studentenkneipen, Kulturszene usw.) immer noch die maßgeblichen Gründe für die Wahl einer Hochschule sind (vgl. z.B. HIS 1999, Lenz/ Wolter/Vogel 2002, Lenz/Wolter/Winter 2000). Die Qualität der Hochschule/ des Studienganges bzw. die Studienbedingungen spielen oft nur eine untergeordnete Rolle. Dies ist aber vor allem darauf zurückzuführen, daß sich die Studienanfänger hierzu häufig (noch) nicht ausreichend informiert fühlen. Zwar zeigen einige Studien, daß die Qualität der Hochschule mit zunehmendem Informationsgrad wichtigerer Auswahlgrund wird (vgl. Müller-Böling 2001). Aber noch sind für die Mehrheit andere Gründe ausschlaggebend.

In der Praxis wirken Studentenzahlen dennoch indirekt als Leistungskriterium, denn über die Kapazitätsverordnung wirken sie auf die Finanzierung von Lehreinheiten sehr wohl ähnlich einer expliziten Leistungskennziffer (wie z.B. in Großbritannien – vgl. Orr 2001). Und auch im universitären Alltag messen sich Hochschulen sehr wohl an ihren Studentenzahlen und konkurrieren z.B. um den inoffiziellen Titel „größte Hochschule“ in Sachsens.<sup>155</sup>

---

<sup>155</sup> In Jahr 2004 ist an dieser Stelle hierzu ein weiteres Problem zu nennen: Die Universität Leipzig lieferte leider bis Redaktionsschluß für die Aufbereitung der sächsischen Lehrberichtsdaten keine Studierendenzahlen. Demzufolge konnten nicht wie bei allen anderen Hochschulen Zahlen direkt aus der Universität verwendet, sondern diese mußten mit Angaben des Statistischen Landesamtes ergänzt werden. Die Verwendung der Angaben des Statistischen Landesamtes erfolgte für das aktuelle Jahr und das Vorjahr, sodaß der Trend auf einer einheitlichen Datenbasis berechnet wurde und dies hierauf keine Auswirkungen hatte. Durch die andere Datenquelle und die etwas andere Systematik kann es aber zu Abweichungen gegenüber hochschuleigenen Zahlen kommen. Eine Durchsicht der Studierendenzahlen am Beispiel der nächstgrößeren Hochschule TU Dresden nach beiden Quellen zeigte jedoch, daß es selbst bei einigen Fächern, in denen sich etwas größere Abweichungen ergeben (u.a. Psychologie, Geschichte, Germanistik – insgesamt zehn von rund 150 untersuchten Fächern), dadurch keine Änderung in der Grundaussage gibt, an welcher Universität die größere Studentenzahl immatrikuliert ist. Die Zahlen sind zwar leider, aus durch die Verfasser nicht zu vertretenden

Ebenso spielen die Studentenzahlen bei hochschulinternen Verteilungskämpfen eine wichtige Rolle, wie Ende der 1990er Jahre z.B. der starke Anstieg der Studienanfängerzahlen in der Informatik illustrierte. Wie viele von den Studienanfängern dann auch tatsächlich das Studium erfolgreich abschließen, spielt hierbei zunächst oft keine Rolle. Dieser Aspekt soll aber erst unter dem Stichwort Absolventenquote ausführlicher diskutiert werden.

Die Folgen dessen, daß die Studentenzahlen in der Praxis bereits indirekt als Leistungsindikator für die Mittelverteilung wirken, sind im universitären Alltag bereits jetzt spürbar: So betreiben immer wieder einige Fakultäten die Strategie, Studentenzahlen als Verhandlungsmasse zu nutzen. Wenn dann letztlich doch nicht die erhofften finanziellen Mittel oder nur ein Teil zu erzielen war, setzt oft ein Prozess der künstlichen "Auslese" von Studenten ein, um die Zahl der zu betreuenden Studenten wieder auf ein als erträglich empfundenenes Maß zu reduzieren. Da die „Auslese“ häufig über sogenannte Multiple-Choice-Prüfungen erfolgt (vgl. hierzu auch die Diskussion zur Zuverlässigkeit und Validität des Kriteriums Noten), oder über ein Zurückfahren der Betreuung auf das unbedingt Nötige (z.B. Sprechstunden nur zu für Studierende ungünstigen Zeiten montags früh oder freitags nachmittag), werden wohl letztlich tendenziell nicht die – gemessen an den späteren beruflichen Anforderungen – tatsächlich besten Studenten (mit guten Noten) das Studium abschließen. Vielmehr sind dies tendenziell eher diejenigen, die – z.B. weil sie nicht so stark Nebentätigkeiten nachgehen (müssen) – die Zeit haben, Prüfungsinhalte auswendig zu lernen und weniger Wert auf die Anwendbarkeit des Wissens legen bzw. neben dem Studium nicht arbeiten müssen. Die Ergebnisse der Analysen zum Einfluss der Noten auf den späteren beruflichen Erfolg der Absolventen der TU Dresden im Rahmen der Dresdner Absolventenstudien 2000-2003 legen diesen Schluss jedenfalls nahe (vgl. Krempkow/ Popp 2003a).

Während die hier beschriebenen Nebenwirkungen lediglich an einigen Hochschulen bzw. Fachbereichen in Deutschland bereits jetzt spürbar sind, haben andere Länder, die die

---

Gründen, nicht so genau vergleichbar wie angestrebt. Eine aus Gründen der Vergleichbarkeit grundsätzlich mögliche einheitliche Verwendung der Daten des Statistischen Landesamtes für alle Hochschulen war aus Zeit- und Ressourcengründen nicht mehr möglich. Eine zutreffende Orientierung für Studien- und Wechselwillige, an welcher Hochschule tendenziell eine größere Anonymität herrschen könnte, sind die verwendeten Zahlen jedoch allemal. Für zukünftige Veröffentlichungen wird aber zu überlegen sein, ob in solchen Fällen nicht auf die größere Detailliertheit und Aktualität der aus den Hochschulen stammenden Studierendenzahlen aus methodischen Gründen bzw. zur besseren Vergleichbarkeit wieder verzichtet werden muß.

Anzahl der Studierenden direkt als Leistungsindikator verwendeten, unerwünschte Nebenwirkungen im gesamten Hochschulsystem zu beachten. So wird in beinahe allen von Leszcensky u.a. (2004: 191 ff.) untersuchten OECD-Ländern die Anzahl der Studierenden als Indikator im Rahmen der Formelfinanzierung eingesetzt (Ausnahme ist Irland, dort gibt es allerdings ein System staatlicher Bildungsgutscheine, das ähnliche Wirkungen erzielt). Häufig werden in den Ländern, wo der Hochschulzugang nicht ohnehin staatlich geregelt ist, der Anreizwirkung jedoch bewusst dadurch Grenzen gesetzt, daß zwischen Staat und einzelnen Hochschulen Vereinbarungen über die maximale Anzahl der aufzunehmenden Studierenden geschlossen werden. Auf diese Weise sollen (im Zusammenwirken mit anderen Maßnahmen) auch unerwünschte Nebenwirkungen wie das beschriebene Absenken der Qualität vermieden werden. Andere Maßnahmen, die die Nebenwirkung vermeiden sollen, daß Hochschulen bei der Finanzierung nach Studierendenzahlen kein Interesse daran haben, daß Studierende zügig ihren Abschluß machen können, sind die Begrenzung der Finanzierung auf Studierende in der Regelstudienzeit und/ oder die Nutzung weiterer Indikatoren, wie z.B. Anzahl der Absolventen (vgl. Hartwig/ Küpper 2002: 12). Einige OECD-Länder versuchen die Nebenwirkung des geringen Interesses an einer guten Betreuung der Studierenden oder Konflikte mit anderen Zielen (z.B. Stärkung der Internationalität) auch zu vermeiden, indem sie als Indikator nicht Studierendenzahlen, sondern Credit Points verwenden, so Schweden und Norwegen. Auf diese Weise wird auch bei einem (gewünschten) Hochschulwechsel derjenige Anteil honoriert, den die jeweilige Hochschule vor einem Abschluß erbracht hat.

So kritisch man die Aussagekraft der Studenten- und Studienanfängerzahlen als Indikatoren für die "Leistung" einer Hochschule aber auch sehen muß, in der Praxis wirken sie aus den genannten Gründen mehr oder weniger als solche und sollten schon allein deshalb bei Gegenüberstellungen statistischer Kennziffern von Hochschulen mit ausgewiesen werden (vgl. auch Webler u. a. 1993). Hinzu kommt, daß die Studenten- und Studienanfängerzahlen für Studieninteressierte neben der Betreuungsrelation eine wichtige Orientierungsgröße darstellen, denn die empfundene Anonymität oder auch das manchmal erwünschte "aufgehoben sein in der Masse" an einer Hochschule hängen nicht nur von der Betreuungsrelation ab, sondern auch davon, wie viele Studenten in einem Studienanfängerjahrgang bzw. im Studiengang insgesamt anzutreffen sind.

### **5.2.3 Anzahl (abgelehnter) Bewerber – Indiz für Attraktivität des Studienganges?**

Von verschiedenen Seiten werden immer wieder auch die Bewerberzahl oder – vor allem von kleineren Hochschulen – die Anzahl der Bewerber pro Studienplatz bzw. der Anteil der abgelehnten Bewerber als Indikator für die Attraktivität eines Studienganges betrachtet, der die Förderung eines nachfragegerechten Studienangebotes bewirken soll (vgl. Ziegele 2002: 10). Auch das Sächsische Ministerium für Wissenschaft und Kunst fordert Informationen über die Anzahl abgelehnter Bewerber als Bestandteil der Lehrberichte, ohne jedoch festzulegen, welchem Zweck diese dienen sollen (vgl. SächsLehrberVO § 4 (1.) a). Von einigen Hochschulen wird diese zu Marketing-zwecken verwendet. So warben z.B. die Wirtschaftsfakultät der TU Dresden und das Institut für Kommunikationswissenschaft der TU Dresden, aber auch die HTW Dresden und andere Hochschulen mit einer hohen Zahl abgelehnter Bewerber und benutzten diese zugleich als Argumentation für die Schaffung neuer Stellen bzw. gegen Stellenkürzungen.

Wie bereits bei der Diskussion der Studienanfängerzahlen erläutert, sind in Deutschland Gründe für die Wahl einer Hochschule auch nach neueren Studien meist die Nähe zum Herkunftsort oder die Attraktivität des Studienortes (HIS 1999, Lenz/ Wolter/ Winter 2000, Müller-Böling 2001). Da also die Qualität der Hochschule/ des Studienganges bzw. die Studienbedingungen oft nur von untergeordneter Wichtigkeit sind, spiegeln die Bewerberzahlen meist eher die Anzahl der Studienberechtigten in der Umgebung des Hochschulortes als die tatsächliche Attraktivität des Studienganges im Verhältnis zu demselben Studiengang an anderen Hochschulen wider. Als Attraktivität eines Studienganges können Bewerberzahlen bzw. Bewerberrelationen also nur dann gelten, wenn die Studienanfänger sich auch tatsächlich bewußt wegen (der Qualität) des Studienangebotes für den jeweiligen Studiengang entschieden haben (vgl. auch Ziegele 2002: 11), was durch Studentenforschungen (vgl. z.B. Lenz/ Wolter/ Winter 2000) oder auch Abiturientenforschungen (vgl. Lenz/ Wolter/ Vogel 2002) überprüfbar wäre.

Als methodisches Problem bei der Berechnung z.B. der Bewerber pro Studienplatz kommt hinzu, daß in Sachsen laut LehrberVO nur die Anzahl abgelehnter Bewerber ausgewiesen werden soll, z.B. an der TU Dresden in den Lehrberichten jedoch alle Bewerber ausgewiesen wurden – unabhängig davon, ob sie aus freier Entscheidung von ihrer Bewerbung zurücktraten (z.B. weil sie sich letztlich doch für eine andere Hochschule/ Studiengang entschieden) oder ob sie tatsächlich abgelehnt wurden. Durch dieses

methodische Problem würde die TU Dresden bei Verwendung der vorhandenen Zahlen als Indiz für Attraktivität tendenziell attraktiver erscheinen als andere Hochschulen, was jedoch evtl. allein auf eine unterschiedliche Berechnungsgrundlage zurückzuführen wäre.

Aufgrund der auch nach neueren Studien kaum veränderten Motive der Hochschulwahl, wegen der fehlenden Möglichkeit der Überprüfung der Motive bei den jeweiligen Studiengängen an allen (sächsischen bzw. deutschen) Hochschulen und schließlich auch wegen z.T. unterschiedlicher, zur Verzerrung der Ergebnisse führender Berechnungsgrundlagen wird eine Verwendung der Anzahl abgelehnter Bewerber als (Leistungs-)Indikator eher kritisch beurteilt.<sup>156</sup>

#### **5.2.4 Betreuungsrelation als ein Indikator für Überfüllung**

Die Betreuungsrelation (oder: das Betreuungsverhältnis) ist für Studienanfänger, Hochschulen, aber auch für Personalentscheider in den Ministerien eine als äußerst wichtig eingeschätzte, aber dennoch mit vielen Problemen behaftete Kennziffer. "Das zahlenmäßige Verhältnis zwischen Lehrpersonal und Studierenden lässt sich als Indikator für die Betreuungskapazität eines Fachbereiches nutzen", führen beispielsweise Hornbostel & Daniel dazu aus. Auch die Kultusministerkonferenz (1996) sah dies als eine der wichtigsten Kennziffern an und schließlich wurde auch im "Focus"-Ranking (1997) und in weiteren nachfolgenden Rankings das Betreuungsverhältnis einbezogen.

Je nach Perspektive bzw. je nach für die Qualitätsorientierung zugrundegelegte Zielgruppe kann eine relativ „idyllische“ Betreuungsrelation von beispielsweise 15 Studierenden je Professor als Indikator für das Potential zu einer guten Betreuung angesehen werden (Zielgruppe Studierende bzw. Studienanfänger) oder als Indikator für eine Unterauslastung und damit mangelnde Effizienz des Studienganges an einer Hochschule (Zielgruppe Ministerium). Bei letzterer Perspektive ist allerdings als mögliche unerwünschte Nebenwirkung eine Qualitätssenkung zu befürchten (vgl. auch Ziegele 2002: 12). Hier ist demzufolge – wie am Anfang dieses Kapitels zur Klärung des Begriffes Indikator ausgeführt, eine normative Entscheidung zur Bewertung entsprechend der Einordnung in das theoretische Konzept zu treffen (vgl. auch Pasternack 2004: 118). Da in dieser Arbeit

---

<sup>156</sup> Daher wurde konsequenterweise (im Unterschied zur Erstveröffentlichung des Studienführers Sachsen für die Ingenieurwissenschaften – vgl. Sächsische Zeitung vom 13. /14. Mai 2000) seit der Veröffentlichung des Studienführers Sachsen für die Geistes- und Sozialwissenschaften – auf die Verwendung des Anteils der abgelehnten Bewerber pro Studienplatz als Indikator für die Attraktivität eines Studienganges verzichtet (vgl. Krempkow/ König/ Winter 2001 und Folgejahre).

der Aspekt der Qualität gegenüber dem der Effizienz im Vordergrund steht, wäre hier die Entscheidung folgerichtig zugunsten der Qualität zu treffen.

Aber auch wenn diese Grundentscheidung getroffen ist, gibt es noch einige andere Probleme: Denn je nachdem, ob man dabei nur die Anzahl der Professoren oder auch die des übrigen wissenschaftlichen Personals berücksichtigt, ergeben sich sehr unterschiedliche Ranglisten", gaben schon Hornbostel & Daniel (1996) zu bedenken.<sup>157</sup> So blieben auch beim Focus-Ranking (1997) erboste Leserbriefe nicht aus, die den Vergleich von unvergleichbaren Statistiken rügten. Dabei hatte sich "Focus" in diesem Falle nicht auf eigene Erhebungen gestützt, sondern auf die von den offiziellen Behörden, also den Hochschulen selbst bzw. die von den statistischen Landesämtern veröffentlichten Zahlenwerke.

Allerdings wurden diese Statistiken offenbar von vielen offiziellen Stellen ein wenig anders berechnet. Hier sollte also zunächst eine einheitliche Regelung oder zumindest Klarheit über die Berechnungsgrundlagen hergestellt werden, damit diese Zahlen vergleichbar werden (vgl. auch Anforderungen an Indikatoren nach Ziegele 2002). Aber auch wenn diese Voraussetzung erfüllt würde, wäre die Kennziffer wohl noch umstritten. Denn während im Grundstudium ein Grossteil der Lehrleistungen durch das wissenschaftliche Personal abgedeckt wird und dann die Einbeziehung der wissenschaftlichen Mitarbeiter durchaus berechtigt oder sogar notwendig erschiene, werden von den Studierenden im Hauptstudium, insbesondere in der Prüfungsphase, vom wissenschaftlichen Personal deutlich weniger Lehrleistungen nachgefragt. Zur besseren Charakterisierung der tatsächlichen Betreuungsrelationen empfehlen Hornbostel & Daniel deshalb für das Hauptstudium die Kennziffer Absolventen je Professor anstelle von Studierenden je Dozent. Dies wiederum ist aber nicht in jedem Studienfach in gleichem Maße der Fall, da beim Vergleich verschiedener Studienfächer hierbei erhebliche Probleme bei der abzusichernden Betreuung der Studierenden auftreten könnten, wie z.B. bei Personalentscheidungen (bzw. -einsparungen) oft argumentiert wird. Auch Webler u.a. (1993) weisen darauf hin, daß aufgrund unterschiedlich hoher Freiheitsgrade der Abwesenheit, Intensität u.a. Umständen die tatsächliche Nachfrage nach Lehrleistungen nur schwer quantifizierbar ist. Wegen der Unterschiedlichkeit der Fächerkulturen sei daher nur im Vergleich derselben Fächer die

---

<sup>157</sup> Diese Probleme entschärfen sich etwas, wenn statt der häufig noch verwendeten Soll-Stellenzahlen (lt. Haushaltsplänen) einheitlich die Ist-Stellenzahlen verwendet würden und v.a. eine getrennte Ausweisung der Stellen nach Finanzierungsquelle erfolgt, so daß z.B. die meist nicht oder nicht im selben Umfang für die Betreuung von Studierenden zur Verfügung stehenden Drittmittelstellen herausgerechnet werden könnten.

Annahme einigermaßen richtig, daß die Studiengewohnheiten sich zwischen Fachbereichen hinreichend ähneln (vgl. Webler u.a. 1993: 39).

Einige weiterführende Informationen ließen sich noch durch die Auswertung des Veranstaltungsangebotes der einzelnen Institute und eine Gegenüberstellung der zur Verfügung stehenden Lehrkapazität (gemessen z.B. am Lehrdeputat) gewinnen. "Art und Anzahl der angebotenen Lehrveranstaltungen geben sehr viel genauer die zur Verfügung stehende Lehrkapazität wieder als die Angaben der Personalstatistik", führen Hornbostel & Daniel (1996) hierzu aus.

Allerdings sagt auch diese nur etwas über die Betreuungsrelation in Lehrveranstaltungen. Über die außerhalb der Lehrveranstaltungen zur Verfügung stehende Betreuungskapazität (z.B. Sprechstunden/ Beratungsangebote allgemein) sagt dies wenig aus. Es gibt zwar einige Untersuchungen, die dies z.T. ansatzweise einbezogen, z.B. zum Lehrangebot in der Soziologie (vgl. Klima 1975, Heitbrede 1985, Hartmann 1989), bisher wurde das Lehrangebot jedoch v.a. auf Einheitlichkeit und die Anzahl der angebotenen Lehrveranstaltungen untersucht. Der Zusammenhang mit der Zahl der Studierenden, denen dieses Lehrangebot zur Verfügung steht, wurde bisher noch in keinem Fall hergestellt, so Hornbostel & Daniel weiter. Genau dies würde aber vermutlich weit näher an der Realität liegen als die bisher verwendeten Kennziffern und Indikatoren und wohl auch mit der (inter)subjektiv empfundenen Güte der Betreuungssituation stärker in Zusammenhang stehen. Beachtet werden müsste bei der Analyse von Vorlesungsverzeichnissen jedoch außerdem, daß es sich lediglich um Veranstaltungsankündigungen handelt, die oft mit der Art und Anzahl der tatsächlich durchgeführten Veranstaltungen nicht übereinstimmen. Dieses Problem ließe sich evtl. dadurch lösen, daß man zur Analyse nicht die offiziellen Vorlesungsverzeichnisse verwendet, sondern beispielsweise die an einem Stichtag in der Mitte eines Semesters gültigen und dann wahrscheinlich eher der Realität entsprechenden Veranstaltungspläne.

Dennoch ist auch hierbei ein weiteres zu all den genannten Problemen hinzukommendes Dilemma noch nicht gelöst: Daß sowohl die für die Erfassung der Studierendenzahl als auch für das Betreuungsverhältnis wichtigen Dienstleistungsverflechtungen (z.B. der Mathematischen Institute für die Statistik-Vorlesungen in der Soziologie) und die Lehnachfrage der Nebenfachstudierenden oft nicht berücksichtigt werden. Auch Webler u.a. (1993: 39) weisen auf solche "Zurechenbarkeitsprobleme" von Lehrleistungen hin. Diese Dienstleistungsverflechtungen wären also ebenfalls in die Überlegungen einzubeziehen,

wenn mit der Kennziffer Betreuungsverhältnis operiert wird – wobei jedoch nur die Hochschulen selbst diese Informationen bereitstellen könnten. Bislang nicht diskutiert wurde, ob auch die Kapazitätsberechnungen der Hochschulen hilfreich sein könnten, um die zahlenmäßige Betreuungsrelation zu ergänzen – sofern sie auf einheitlicher Berechnungsgrundlage verfügbar wären (vgl. z.B. Rammelt 2003).<sup>158</sup> Diese könnte ggf. auch das Problem der mangelnden Erfassbarkeit von Dienstleistungsverflechtungen lösen helfen. Auch die räumliche Kapazitätsberechnung könnte evtl. ergänzende Aussagen über die räumliche Auslastung bzw. Überfüllung ermöglichen. Ob dies jedoch umsetzbar und der Aufwand vertretbar ist, müssten erst praktische Erfahrungen mit einer Analyse von Daten z.B. eines Bundeslandes zeigen.

Aber auch wenn all dies berücksichtigt wird, können solche Relationen dennoch niemals eine vollständige Entsprechung für die von Studierenden teilweise beklagte mangelnde Beratung und Betreuung und für die empfundene Anonymität der Hochschule sein. Denn hierfür spielt auch die Anzahl der Studierenden am Fachbereich insgesamt, in dem entsprechenden Studiengang, sowie die Kommunikation der Studierenden untereinander eine Rolle (vgl. HIS 1992 u. 1995). Die Anzahl der Studierenden wiederum kann sich in Grund- und Hauptstudium und in den einzelnen Semestern erheblich unterscheiden.

Deshalb kann das Fazit dieser Diskussion der Betreuungsverhältnis nur lauten, daß diese in Form isoliert betrachteter Kennziffern über die tatsächliche Betreuungssituation nur wenig aussagt. Sie kann lediglich als Anhaltspunkt für das zur Verfügung stehende Potential dienen. Für eine der Realität nahekommende Einschätzung ist die subjektive Bewertung durch die Studierenden unerlässlich, wobei auch Aspekte zugänglich würden, die sich einer „objektiven“ Erfassung entziehen, wie z.B. die Ansprechbarkeit der Dozenten oder die Qualität von Beratungsangeboten. Die „objektive“ Kennziffer und die subjektive Bewertung dabei müssen keineswegs als sich in ihrem "Wahrheitsgehalt" widersprechende Informationen aufgefasst werden, wie dies mit einer unterstellten Ersetzbarkeit suggeriert würde. Vielmehr können die auf so verschiedene Weise gewonnenen Informationen als sich gegenseitig sinnvoll ergänzende Kriterien für ein der Realität möglichst nahe kommendes Bild der (Betreuungs-)Situation an einem Fachbereich betrachtet werden.

---

<sup>158</sup> Für das Land Berlin hat HIS (2004: 17ff.) eine kapazitätsbezogene „Auslastungsquote“ als Anzahl der Studierenden in der Regelstudienzeit in Relation zur Anzahl der Studienplätze definiert, die einen Anreiz zur besseren Auslastung unterausgelasteter Studiengänge schaffen soll und mit 10% Gewicht in die leistungsorientierte Finanzierung eingeht.

Als Indikator für Leistungsanreize, z.B. bei der leistungsorientierten Mittelverteilung, eignet sich die Betreuungsrelation aufgrund der beschriebenen Probleme nicht. Sie hat jedoch ggf. über die Veröffentlichung von auffallend positiven oder auffallend negativen Relationen dennoch indirekt einen steuernden Effekt, indem als idyllisch beschriebene Betreuungsrelationen potentiell Studenten anziehen könnten, während sehr ungünstige Betreuungsrelationen bei einer solchen Wirkung weniger Studierende anziehen sollten. Inwieweit ein solcher Effekt jedoch tatsächlich stattfindet, wurde bislang nach Kenntnis des Verfassers nicht untersucht. Aber selbst wenn der beschriebene Effekt nicht oder nur in geringem Ausmaß stattfindet, sollte eine zahlenmäßige Betreuungsrelation als Orientierungshilfe für Studierende auch künftig bei vergleichenden Veröffentlichungen zur Studiensituation berücksichtigt werden, da sie – wie bereits weiter oben beschrieben – als ein die subjektive Sicht der Studierenden ergänzendes Kriterium für eine umfassende Darstellung der (Betreuungs-)Situation in einem Fach wichtig anzusehen ist.

#### **5.2.5 Durchschnittliche Fachstudiendauer - ein Indikator für Lehrqualität?**

Eine höhere Studiendauer der Hochschulabsolventen im Vergleich zu vielen anderen Industrienationen (vgl. HIS 2003: 27) wird oft als Indiz für schlechte Lehrqualität genannt. Im Umkehrschluss wird eine kurze Studiendauer von Politik und Öffentlichkeit oft als Indiz und Ziel für "gute Lehre" angesehen. Welches Bildungsverständnis auch immer dahinterstehen mag, in jedem Fall scheint eine kurze Studiendauer auch seitens der Wirtschaft ein Ziel "guter" Lehre zu sein und unter der gegebenen sozialen Situation (da bei Überschreitung der Regelstudienzeit in einigen Bundesländern Sanktionen drohen und BAföG-Empfänger verzinste Darlehen aufnehmen bzw. bei weiterer Überschreitung komplett selbst ihren Lebensunterhalt finanzieren müssen) auch für die Studierenden angeraten. Angesichts knapper öffentlicher Kassen erscheint dies verständlich, zumal durch die geringere zur Verfügung stehende Lebensarbeitszeit auch tendenziell weniger Beiträge in das Steuer- und Sozialsystem fließen und der Volkswirtschaft mögliches Innovationspotential entzogen wird (vgl. HIS 2003: 24). Andere argumentieren, daß das Studium ein Recht auf Bildung darstelle und die Dauer des Studiums als persönliche Entwicklungsmöglichkeit jedem Individuum freistehen sollte.

Bei allen politischen Kontroversen um die Verkürzung der Studiendauer besteht jedoch Konsens, daß denjenigen, die es möchten, ein Abschluss des Studiums innerhalb der Regel-

studienzeit ermöglicht werden soll. Mit der Zielstellung, einen „Anreiz zur Schaffung der Bedingungen zur Einhaltung der Regelstudienzeit durch gute Lehre und Betreuung“ zu schaffen, wird die durchschnittliche Fachstudiendauer (evtl. in Relation zur Regelstudienzeit) auch als Leistungsindikator für Mittelverteilungsmodelle diskutiert (vgl. Ziegele 2002: 13).

In mehreren Bundesländern, wie z.B. in Sachsen, ist diese Formulierung sogar explizit ins Hochschulgesetz aufgenommen worden. Deshalb soll die durchschnittliche Studiendauer der Absolventen trotz einiger Zweifel an ihrer Aussagekraft für die Qualität der Lehre als in Rankinglisten oft verwendete Kennziffer diskutiert werden. Die Auseinandersetzung um die Ziele der Studienzeitverkürzung führt allerdings mehr zu einer politischen als zu einer methodischen Diskussion. Lediglich letztere soll an dieser Stelle geführt werden. Methodisch muss die Studiendauer zunächst unterschieden werden nach Fachstudiendauer (Studiendauer gemessen in Fachsemestern) und der Verweildauer an Hochschulen (gemessen in Hochschulsemestern). Als Hochschulsemester werden alle Semester bezeichnet, in denen jemand an einer Hochschule eingeschrieben war, z.B. einschließlich anderer Fachrichtungen vor einem Fachrichtungswechsel. Als Fachsemester werden nur solche Semester bezeichnet und für die Fachstudiendauer angerechnet, in denen jemand in dem jeweiligen Fach eingeschrieben war.

Trotz dieser Einschränkung in der Betrachtung der Studiendauer nur als Fachstudiendauer ist es immer noch ein Zusammenspiel von verschiedensten Faktoren, die zu den in der Tat vorhandenen gravierenden Unterschieden in der Studiendauer gleicher Fachrichtungen an verschiedenen Hochschulen führen. Dies zeigen die großen Schwankungen der durchschnittlichen (Fach-)Studiendauer innerhalb kürzester Frist an einigen Studienorten, obwohl es keine gravierenden Veränderungen der Rahmenbedingungen des Studiums gab. Um wiederum das Beispiel der Soziologie aufzugreifen: In Saarbrücken schwankte die durchschnittliche Studiendauer binnen Jahresfrist um 4,5(!) Semester (1989-1990; Datengrundlage: offizielle Statistik des Wissenschaftsrates). Auch an anderen Hochschulen kamen Schwankungen um 2-3 Semester in einem Jahr vor (vgl. Abb. 1 in Hornbostel & Daniel 1996). Als Ursachen für diese Schwankungen wurden starke Schwankungen der Anfängerjahrgänge, aber auch relativ geringe Absolventenzahlen ausgemacht, bei denen "Ausreißer" die Mittelwerte stark beeinflussen. Hinzu kommt als weiteres Problem, daß die Fachstudiendauer an den Hochschulen teilweise nicht einheitlich erfasst wird. So wird bei einigen Hochschulen das Datum der Zeugnisübergabe, bei anderen das Datum der letzten

Prüfung und wieder anderen die Abgabe der Abschlußarbeit als Studienabschlußzeitpunkt gewertet. Zudem ist nicht immer klar, was als mittlere Fachstudiendauer angegeben wird, da diese z.T. als arithmetisches Mittel und z.T. als Median ausgewiesen wird. Dadurch ist sie teilweise weder eindeutig noch vergleichbar.

Die durchschnittliche Fachstudiendauer allein ohne Berücksichtigung des Kontextes und der Fallzahl hat also kaum einen Aussagewert über die Studienbedingungen bzw. die Qualität der Lehre an einer Hochschule.<sup>159</sup> Erst wenn die Studiendauer über einen längeren Zeitraum und mit ausreichender Datengrundlage konstant hoch ist, könnte dies ein Indiz für studienorganisatorische Probleme sein, die dann z.B. durch Befragung von Studierenden genauer lokalisiert werden könnten. Anhaltspunkte für Faktoren, welche die durchschnittliche Fachstudiendauer maßgeblich beeinflussen können, bietet eine Studie von Daniel (1996b), in der über 1000 Betriebswirtschaft-Absolventen an der Universität Mannheim befragt wurden. Die Ergebnisse dieser Studie zeigen, daß vor allem der Bildungsweg, die Abiturnote und die Form der Studienfinanzierung einen "statistisch und praktisch bedeutsamen Einfluss auf die Fachstudiendauer" haben. Als einziger Faktor, der im Bereich der Verantwortung der Hochschulen liegt und dem mit besserer Studienorganisation abzuhelfen wäre, stellte sich die Anzahl der Wiederholungsprüfungen heraus.

Durch eine zeitlich enger gestaffelte Folge von Wiederholungsprüfungen (am Ende der Vorlesungszeit und kurz vor Beginn eines neuen Semesters) könnte das Studium in diesem Fall um jeweils ein Semester verkürzt werden, denn bisher fanden Prüfungen immer nur jeweils einmal im Semester statt. Weitere wichtige Ansatzpunkte zur Verkürzung der Studienzeit sieht Daniel bei der Berufs- und Studienberatung, beim Hochschulzugang und der Studienfinanzierung. So interessante Ansatzpunkte die Studie bieten mag, von einer Generalisierbarkeit der Ergebnisse kann noch nicht gesprochen werden, da in der Studie nur Betriebswirtschaftler untersucht wurden und das Studierverhalten in verschiedenen Fächern von den jeweiligen Fächerkulturen geprägt ist.

Jedoch ergab eine seit 1999 in den Absolventenstudien der TU Dresden verwendete Fragestellung ähnliche Resultate. Hier wurde in Anlehnung an Studentenbefragungen der HIS GmbH gefragt, ob die Absolventen in ihrem Studium Zeit verloren haben, weil sie aufgrund organisatorischer Regelungen (Lehrangebot) Leistungsnachweise nicht erwerben oder nicht

---

<sup>159</sup> In diesem Sinne äußern sich auch Interviewpartner bei den im Rahmen dieser Arbeit durchgeführten Experteninterviews. So wurde beispielsweise auch der Anteil von Studierenden mit Kindern angeführt, die tendenziell eine längere Studiendauer benötigen und der daher berücksichtigt werden sollte.

bestandene Klausuren, Prüfungen o. ä. erst später wiederholen konnten. Mit Hilfe dieser Fragestellung sollte untersucht werden, inwieweit Zeitverluste neben Erwerbstätigkeit und anderen studienzeitverlängernden Aktivitäten auch auf studienorganisatorische Probleme zurückgeführt werden können. Die Frage erklärte einen beachtlichen Teil der Zeit, die die Absolventen über die Regelstudienzeit hinaus studierten (vgl. Krempkow 1999, 2000, 2001, Krempkow u.a. 2002, 2003, 2004).

Insgesamt ist die Eignung der durchschnittlichen Studiendauer als Kriterium für die Einschätzung der Studienbedingungen und die Qualität des Lehrbetriebes nach den vorliegenden Ergebnissen als fragwürdig einzustufen, da sie zu sehr auch von Faktoren abhängt, die außerhalb des Einflussbereiches der Hochschulen liegen. Zur genaueren Klärung dieser Frage wären spezielle Auswertungen bundesweiter Absolventenbefragungen (z.B. von Teichler/ Schomburg u. a. 2001) notwendig, die m. E. noch nicht vorliegen. Eine große Hilfe hierfür wären mittels Studenten- oder Absolventenbefragungen erhobene Antworten z.B. auf die Frage, ob und wieviel im Studium aufgrund ungünstiger studienorganisatorischer Regelungen Zeit verloren wurde. Solche Informationen liegen bislang jedoch nur von wenigen Hochschulen vor (vgl. Krempkow/ Popp 2003b). Solange kann die durchschnittliche Fachstudiendauer lediglich eine grobe Orientierung sein, wie lange die Studenten in etwa für ein Studium in diesem Studiengang einplanen müssen, nicht jedoch für sich genommen ein „objektives“ Qualitätskriterium für einen Studiengang. Für Studierende, die auf eine Förderung nach dem BAFÖG angewiesen sind, ist trotz aller methodischen Probleme des Kriteriums die Studiendauer aber eine wesentlich bessere Orientierungshilfe als die in den Studienordnungen ausgewiesenen Regelstudienzeiten.

Voraussetzung ist aber, daß ausreichende Fallzahlen zugrundeliegen. Außerdem sollte dabei die Entwicklungstendenz der aktuellen Studiendauer im Vergleich zu vorhergehenden Semestern als zusätzliche Information zur Verfügung gestellt werden, um ein evtl. Vorhandensein größerer Schwankungen zumindest kenntlich zu machen. Eine grobe Prognose der künftigen Studiendauer eines Faches als die Studiendauer ergänzende Information erlaubt unter bestimmten Voraussetzungen (u.a. Abschluß des Studiums in diesem Studiengang) der Anteil der Studierenden in der Regelstudienzeit, der hier als Studierbarkeit bezeichnet und nachfolgend diskutiert wird.

### 5.2.6 Studierbarkeit – besseres Kriterium als die Studiendauer?

Unter Studierbarkeit wird alltagssprachlich meist die Möglichkeit verstanden, das Studium "zu schaffen". Hier sind Verwechslungen mit der Erfolgs- bzw. Durchfallquote in Prüfungen oder auch mit der Abbruch-, Schwund- bzw. Absolventenquote und natürlich mit der Studiendauer vorprogrammiert.

In Abgrenzung zu diesen Kennziffern und Indikatoren muss das Kriterium Studierbarkeit also klarer definiert werden. Für eine evtl. Verwendung als Leistungsindikator sollte dabei auch der Zeitaspekt eine Rolle spielen, es muß bei der Studierbarkeit also darum gehen, inwiefern das Studium in der vorgesehenen Zeit „zu schaffen“ ist.<sup>160</sup> Um die genannten Kennziffern und Indikatoren möglichst gut zu ergänzen, wird die Studierbarkeit hier definiert als Anteil der Studierenden innerhalb der Regelstudienzeit (Universitäten meist 9, Fachhochschulen 8 Semester) im Verhältnis zur Gesamtanzahl der Studenten im jeweiligen Studiengang (vgl. auch Ziegele 2002: 13). Sie unterscheidet sich damit deutlich von der Erfolgsquote in Abschlussprüfungen und auch von der Absolventenquote.

Die Studierbarkeit ist bei dieser Definition durchaus im Zusammenhang mit der Studiendauer zu sehen, da ein sinkender Anteil von Studenten innerhalb der Regelstudienzeit bedeutet, daß die über die Regelstudienzeit hinaus Studierenden letztlich u.U. zu einer höheren durchschnittlichen Studiendauer führen würden. Genau durch diesen Zusammenhang wird die Studierbarkeit zu einer Art "Frühwarnsystem" für Probleme, sei es nun in der Studienorganisation, in der Motivation der Studenten oder in der Aufnahmefähigkeit des Arbeitsmarktes, die letztlich alle zu einer längeren Studiendauer führen. Eine ungünstige Studierbarkeit führt nur in wenigen Fällen nicht zu einer höheren Studiendauer, z.B. wenn durch unterschiedliche hochschulgesetzliche Regelungen in den einzelnen Bundesländern Abwanderungstendenzen aus diesen in andere Bundesländer bestehen (etwa bei Studiengebühr für sogen. Langzeitstudenten, Verlust des Prüfungsanspruches oder Zwangsexmatrikulation bei Überschreitung der Regelstudienzeit um mehr als 4 Semester wie z.B. in Sachsen). Letztlich führt dies lediglich zu einer Verlagerung des Problems in andere Bundesländer, die oft mit neuen studienorganisatorischen Problemen z.B. bei der Anerkennung von Studienleistungen und zu noch längeren individuellen Studienzeiten führen. Gesellschaftlich gesehen wäre deshalb eine bessere Betreuung bei Studienproblemen sowie die offizielle Möglichkeit des Teilzeit-

---

<sup>160</sup> In einigen Publikationen wird daher auch eine sogen. Regelstudienzeitquote diskutiert (vgl. HIS 2004: 17).

studiums bei ständiger Erwerbstätigkeit mit entsprechend verlängerter Regelstudienzeit wahrscheinlich eine bessere Lösung (wobei dann die Studierbarkeit für Voll- und Teilzeitstudenten getrennt zu berücksichtigen wäre). Rein methodisch gesehen würde natürlich auch eine bundeseinheitliche restriktive Regelung zu einer höheren Zuverlässigkeit der Kriterien Studiendauer und Studierbarkeit führen. Allerdings würde dies wohl in vielen Fällen bedeuten, daß Studierende mit Problemen in der Studienabschlussphase keinen Abschluss mehr erhalten, was sich dann in der Absolventenquote niederschlagen würde. Damit würde sowohl die individuelle Investition an Lebenszeit als auch die gesellschaftliche Bildungsinvestition entwertet.

Vergleicht man die Aussagekraft von durchschnittlicher Studiendauer und Studierbarkeit, so läßt sich formulieren, daß die Studierbarkeit durchaus eine höhere Aussagekraft hat als die Studiendauer, da sie durch ihre größere Zeitnähe (die von mehreren Autoren als Anforderung an Indikatoren genannt wird) schneller Probleme in der Studienorganisation anzeigen kann. Zudem wird sie nicht so stark durch Abwanderungstendenzen beeinflusst wie die Studiendauer.<sup>161</sup> Für den speziellen Fall relativ junger Studiengänge mit bisher erst wenigen Absolventen kommt hinzu, daß die Studierbarkeit auch deshalb zuverlässiger ist, weil sie auf höheren Fallzahlen beruht. Allerdings ist hierbei zu beachten, daß die Studierbarkeit nur dann ausgewiesen werden kann, wenn der Studiengang lange genug existiert, um überhaupt Studierende außerhalb der Regelstudienzeit haben zu können. Dies gilt jedoch in noch stärkerem Maße für die Studiendauer. Falls es größere Schwankungen in der Anzahl der Studienanfänger gibt (wenn z.B. eine Zulassungsbeschränkung eingeführt oder abgeschafft wurde), sollte auch dies berücksichtigt werden. Dies ist jedoch erfahrungsgemäß nur bei relativ wenigen Studiengängen der Fall.

Dennoch bleibt auch hier anzumerken, daß ein geringer Anteil von Studenten in der Regelstudienzeit mit vielen anderen, nicht in der Lehrqualität zu suchenden Gründen zusammenhängen kann, wie z.B. mit der Attraktivität des Hochschulortes („Ablenkung“ vom Studium), der Arbeitsmarktlage („lieber Student als arbeitslos“) u. v. a. m. Allerdings könnten hier, wie auch zu den Ursachen einer hohen durchschnittlichen Studiendauer, Studenten- bzw. Absolventenbefragungen zu konkreteren Informationen führen, die mögliche Ursachen aufklären könnten. Unter Berücksichtigung der Einschränkungen zur

---

<sup>161</sup> Verlassen Studierende erst im 14. Semester kurz vor der Zwangsexmatrikulation noch den Studiengang, weil es z.B. in einem anderen Bundesland keine solche Regelung gibt, zeigt sich dies in der Studiendauer überhaupt nicht.

Aussagekraft und des Kontextes lässt sich zur Eignung der Studierbarkeit als Indikator für Leistungsbewertungen formulieren, daß diese zwar ebenfalls nicht unproblematisch ist, aber u. a. aufgrund der größeren zugrundeliegenden Fallzahlen, ihrer nicht so starken Beeinflussung z.B. durch Abwanderungen und ihrer größeren Zeitnähe besser geeignet ist als die Fachstudiendauer. Sie hat zudem neben der schnellen und problemlosen Verfügbarkeit den Vorteil der besseren Vergleichbarkeit und Eindeutigkeit, da solche Probleme wie unterschiedliche Definitionen des Abschlußdatums und unterschiedliche Berechnungen der mittleren Fachstudiendauer nicht auftreten.

Betrachtet man die Studierbarkeit als potentiellen Leistungsindikator für die Vergabe finanzieller Mittel, so sind zusätzlich weitere Aspekte zu beachten. So existiert als mögliche unerwünschte Nebenwirkung der Anreiz zur vollen Ausschöpfung der Regelstudienzeit (vgl. Ziegele 2002: 14). Dieser Anreiz kann mit anderen Maßnahmen konfliktieren, die das Ziel der Studienzeitverkürzung verfolgen, wie z.B. sogenannten „Freischußregelungen“.

### **5.2.7 Abschlußnoten: Kein Kriterium für Studienerfolg, aber Orientierungshilfe**

Die erwarteten oder erteilten Noten werden nicht nur als potentieller Einflussfaktor auf die Bewertung der Lehre durch die Studierenden gesehen, sie werden selbst auch als Instrument zur Messung der Leistung der Studierenden diskutiert (vgl. Ziegele 2002: 10ff.) und – da sie das Ziel verfolgen, Qualität zu erfassen – gelegentlich auch als qualitativer Indikator bezeichnet (vgl. Pasternack 2004: 117). Im Zusammenhang mit der Diskussion von Noten als potentieller Einflußfaktor auf die Bewertung der Lehre wies jedoch bereits Rindermann (1996) darauf hin, daß Noten jedoch bei weitem nicht so valide und zuverlässig sind, wie man es von studentischen Lehrbewertungen als anderem qualitativen Indikator verlangt.

Zunächst ist hier auf die enormen Unterschiede in den Durchschnittsnoten der gleichen Fächer an verschiedenen Universitäten hinzuweisen. Schon "Der Spiegel" (1993) und auch "Focus" (1997) hatten die Durchschnittsnoten in ihre Fachbereichsvergleiche einbezogen. Die Unterschiede waren groß genug, so daß sie nicht allein auf Stichprobeneffekte zurückgeführt werden können. Hornbostel & Daniel (1996) hatten die Durchschnittsnoten am Beispiel verschiedener Fachbereiche der Soziologie genauer untersucht und festgestellt, daß an einigen Hochschulen (z.B. der FU Berlin) ein extrem guter Notenschnitt erwartet wird, obwohl es sonst keine Hinweise für eine besondere Motivation gibt. (eher umgekehrt: vgl. Kromrey 1996) Auch zu der Vermutung, daß die Studierenden an der FU Berlin sich

durch besondere Fähigkeiten auszeichnen, (etwa durch eine höhere Durchschnittsnote im Abitur, Aufnahmeprüfungen o.ä.) gab es keinen Anlass. Deshalb wurde der hohe Notenschnitt auf eine unterschiedliche Praxis der Notenvergabe durch die Lehrenden zurückgeführt. Ähnliche enorme Unterschiede in der Praxis der Notenvergabe sind auch immer wieder in anderen Fächern zu beobachten (vgl. auch Krempkow/ König 2004). Dies bedeutet, daß ein Absolvent mit demselben Leistungsstand an der einen Hochschule eine gute Note und an einer anderen Hochschule eine deutlich schlechtere Note erhalten kann. Daher erfüllen Noten eine wesentliche Anforderung an Indikatoren, deren Vergleichbarkeit über dieselben Fächer an verschiedenen Hochschulen, nicht.

Mit der Validität und Zuverlässigkeit von Noten als Kriterium des Lernerfolgs am Beispiel des Studiums der Medizin hat sich Westhoff (vgl. Westhoff in Koebke 1996: 379-383) noch detaillierter auseinandergesetzt. Er unterschied dabei die Notenvergabe grundsätzlich in zwei Verfahren, die für die Einschätzung ihrer Zuverlässigkeit von ganz entscheidender Bedeutung sind: Dies ist zum einen die Vergabe der Noten nach zuvor festgelegtem Maßstab, wie dies z.B. eine zu erreichende Punktzahl für eine bestimmte Note darstellt. Das zweite, sich davon deutlich unterscheidende Verfahren ist die Notenvergabe nach Normalverteilung (bzw. zusätzlich mit einem Faktor versehen). Hierbei ist die Durchschnittsnote der Studierenden und im Normalfall auch die größte Notenhäufigkeit in der Nähe der Note 3 zu finden. Bei diesem Verfahren kann es vorkommen, daß in einem sehr guten Jahrgang ein Studierender mit exzellenten Kenntnissen eine schlechtere Note erhält als ein Studierender aus einem schlechten Jahrgang mit nur mittelmäßigen Kenntnissen. Deshalb hält Westhoff die klare Festlegung eines einheitlichen Maßstabes vor der Notenvergabe für die Verwendung von Noten als zuverlässiges Instrument zur Messung von Lern- bzw. Studienerfolg für notwendig. Für eine Vergleichbarkeit von Noten verschiedener Fachbereiche wäre, dieser Logik folgend, ein einheitlicher Maßstab notwendig. Dies erscheint allerdings aufgrund verschiedener Schwerpunktsetzungen in der Ausbildung oder gar sehr verschiedener Ausbildungsinhalte innerhalb des gleichen Fachgebietes, wie es zwischen den einzelnen Hochschulen häufig der Fall ist, schwer realisierbar (vgl. auch Ziegele 2002: 11). Wenn es aber unmöglich ist, eine einheitliche Benotungsgrundlage herzustellen, so sollten zumindest die Ursachen für die Unterschiede im Notenniveau transparent gemacht werden, damit Absolventen von besonders streng benotenden Einrichtungen keine Nachteile entstehen, denn für einstellende Unternehmen sind Noten nach wie vor ein Kriterium.

Westhoff u.a. hatten deshalb im Zusammenhang mit der Evaluation der sogenannten "Ergebnisqualität" der Lehre verschiedene Prüfungsverfahren, insbesondere die in der Medizinausbildung weit verbreiteten Multiple-Choice-Prüfungen (MCP) kritisch hinterfragt und sie unter Bezugnahme auf verschiedene Testtheorien hinsichtlich ihrer Validität und Reliabilität untersucht. Hierbei stellte Westhoff fest, daß MCP zwar sehr reliabel und leicht auswertbar sind, jedoch oft nicht das "messen", was sie vorgeben zu messen (Validität). Mit MCP würden eher Persönlichkeitsmerkmale erfasst als die Berufsfähigkeit, also die Tatsache, ob jemand als Arzt geeignet sei oder nicht, so Westhoff. "Und das sei das schlimmste, was mit einer Ausbildung passieren könnte" schreibt er in seinem Beitrag: "Daß trotz langjähriger Ausbildung und vorgeblich guten Ergebnissen am Ende Leute als Ärzte auf die Menschen losgelassen werden, die als Ärzte nicht taugen." Für die Messung von Lernerfolg hält er deshalb die gründliche Untersuchung anderer Prüfungsverfahren auf ihre Validität hin für angebracht. Seiner Meinung nach sei es durchaus "prinzipiell machbar, die Objektivität von mündlichen Prüfungen sicherzustellen".

Auch wäre vorstellbar, daß ein Credit-point-system, wie es in den USA, Kanada und einigen anderen vom angelsächsischen Bildungssystem dominierten Staaten üblich ist und im Zuge des sogenannten Bologna-Prozesses auch für Europa zunehmend relevant wird, eher den Anforderungen an Validität und Reliabilität genügt (vgl. Dalichow/ Fritz in: BMBF 1997). Das System hätte den Vorteil, daß kontinuierlich während des gesamten Studiums die Leistung der Studierenden "gemessen" würde. Das rein prüfungsmotivierte Lernen, wie es bei der Medizinausbildung, aber auch in anderen Fächern in Deutschland oft als Problem wahrgenommen wird, würde als Problem in diesem Ausmaß gar nicht erst entstehen. "International sind die zentralen testtheoretischen Bedingungen für gültige Prüfungen in der Medizin z.B. in Kanada oder den USA erfüllt, in Deutschland ist der internationale Standard bei weitem nicht erreicht", so Westhoff u. a.. Diese Einschätzung deckt sich auch mit der anderer Autoren für weitere Fächer (vgl. z.B. Meer 1999: 478).

Aber auch wenn die Noten allen testtheoretischen Ansprüchen genügen, könnten sie ohne Überprüfung ihrer Relevanz für den späteren Berufsweg nicht als Kriterium für den Erfolg des Studiums im Sinne des Studienangebotes der Hochschule gelten. Hierfür müsste anhand von Absolventenstudien zunächst tatsächlich nachgewiesen werden, daß Absolventen mit besseren Noten auch erfolgreicher im Beruf sind. Daß dies nicht unbedingt der Fall sein muss, zeigen beispielsweise die Dresdner Absolventenstudien (vgl. Krempkow 2001,

Krempkow u.a. 2002, 2003, 2004), aber auch die der Münchner Universität (vgl. Brüderl 1996).

Zusammenfassend kann daher zur Eignung von Noten als Kriterium für den Studienerfolg und erst recht zur Eignung von Noten als Indikator für Leistungsanreizmodelle formuliert werden, daß sie wegen ihrer fraglichen Validität und ihrer mangelnden Vergleichbarkeit aufgrund der heterogenen Praxis der Notenvergabe an deutschen Hochschulen kaum geeignet sind. Hinzu kommt, daß bei Verwendung als Leistungsindikator ein „starker Anreiz zur Senkung der Anforderungen und damit der Ausbildungsqualität“ entstehe (vgl. Ziegele 2002: 11). Gerade aufgrund der z.T. sehr unterschiedlichen Praxis der Notenvergabe sollte diese aber sowohl für Studienanfänger und Studierende, als auch für potentielle Arbeitgeber transparent gemacht werden, damit sie als Orientierungshilfe und zur Einordnung der erreichten Noten in die jeweilige Notenvergabepraxis dienen kann.

### **5.2.8 Erfolgsquote bei Prüfungen: Qualitätskriterium eines Studienganges?**

Wenn Prüfungsnoten als Kriterium diskutiert werden, in diese aber nur die bestandenenen Prüfungen einfließen, dann sollten auch die nicht bestandenenen Prüfungen bzw. die Erfolgsquote bei (Abschluß-)Prüfungen in die Diskussion einbezogen werden.

Die Erfolgsquote bei Prüfungen wird hier definiert als Anteil der endgültig bestandenenen Abschlussprüfungen im Verhältnis zur Gesamtzahl der im selben Zeitraum abgelegten Prüfungen. Die Erfolgsquote ist also direkt umgekehrt proportional dem alltagssprachlich als "Durchfallquote" bezeichneten Anteil der (endgültig) nicht bestandenenen Prüfungen. Die Interpretation bzw. Bewertung erfolgt je nach Perspektive unterschiedlich. Während für Studienanfänger und Studierende eine möglichst hohe Erfolgsquote (bzw. gute Note) als anstrebenswert und als positiver Aspekt eines Studienganges gilt, argumentieren einige Professoren bzw. Fachbereichs- und Fakultätsleitungen z.T. umgekehrt mit einer möglichst niedrigen Erfolgsquote (bzw. strengen Notenvergabe) als „Qualitätskriterium“ für die Absolventen eines Studienganges nach dem Prinzip „Nur die besten bestehen die (Abschluss-) Prüfungen“. Auch diese Frage der Bewertung müßte bei der Diskussion von Erfolgsquoten (wie auch von Noten) letztlich normativ entschieden werden. Wichtig erscheint hierbei auch die Berücksichtigung des Aspektes, daß Noten eine sinnvolle Differenzierung des Leistungsstandes ermöglichen sollten (vgl. Wissenschaftsrat 2003).

Dies wäre weder bei sehr niedrigen noch bei sehr hohen Erfolgsquoten bzw. überwiegend sehr guten Noten möglich.

Wie diese Argumentation zeigt, ist die Erfolgsquote in Zusammenhang zu sehen mit den Durchschnittsnoten, da ein strenger(er) Notenmaßstab meist mit niedrig(er)en Erfolgsquoten einhergeht. Die Erfolgsquote stellt trotz ihres Zusammenhanges mit den Noten aber eine zusätzliche Information dar, weil in den Lehrberichten oft nur die Durchschnittsnoten oder nur die Noten der bestandenen Prüfungen enthalten sind. Außerdem sind die Folgen einer endgültig nicht bestandenen Abschlussprüfung für die (Nicht-)Absolventen ungleich gravierender.

Sehr niedrige Erfolgsquoten (wie z.B. bei den Abschlußprüfungen der Informatikstudiengänge an der HTWK Leipzig und der Hochschule Mittweida - vgl. z.B. Krempkow/ König/ Winter 2000 und Folgejahre) bedeuten – und dies wurde vielen Informatikern ohne Abschluss wahrscheinlich erst nach „Zerplatzen“ der „Spekulationsblase New Economy“ klar – letztlich eine „Fehlinvestition“, da diese de-facto-Studienabbrecher nach dem Verlust ihres ersten gut dotierten Arbeitsplatzes nun zumindest teilweise mit weniger gut dotierten Jobs vorliebnehmen müssen als ihre ehemaligen Kommilitonen, die das Studium bis zum Abschluss absolvierten. Bei anderen Fachrichtungen mit Arbeitskräfteüberschuß zeigt sich das Problem von Einkommensverlusten oft noch deutlicher als bei der Informatik (vgl. CHE/ Focus Nr. 16/2000).

Problematisch in seiner Aussagekraft wird das Kriterium Erfolgsquote, wenn diese wie bei den Geistes- und Sozialwissenschaften fast überall bei 100% liegt, gleichzeitig aber bekannt ist, daß nur ein Bruchteil der Studienanfänger letztlich auch einen Abschluss erwirbt. Grundsätzlich aussagekräftiger wäre deshalb anstelle der Erfolgsquote bei Abschlussprüfungen die Erfolgsquote bei den Zwischenprüfungen, die in der einschlägigen Literatur bereits früh als unverzichtbar selbst für Minimalversionen von Lehrberichten vorgeschlagen wurde (vgl. Webler u. a. 1993). Außerdem ließen sich bei Beobachtung der Erfolgsquoten von Zwischenprüfungen bereits zu einem früheren Zeitpunkt im Studienverlauf „Fehlinvestitionen“ und mögliche Probleme beim voraussichtlichen Studienerfolg oder bei der Studienmotivation bzw.- information erkennen und ggf. gegensteuern. An vielen Hochschulen erfolgt eine solche Erfassung bei Zwischenprüfungen jedoch nicht.

Aber auch wenn die Erfolgsquote für die Zwischenprüfungen von allen Hochschulen vorläge und Differenzen zeigte, hätte diese nur so lange Aussagekraft, wie die Zwischen- bzw. Vordiplomprüfungen tatsächlich Hauptursache dafür sind, ein einmal begonnenes

Studium nicht weiterzuführen. Ist das Nichtbestehen von Prüfungen jedoch nur ein (minder wichtiger) Grund von vielen für den großen "Schwund" von Studierenden mit fortschreitender Semesterzahl, so gibt auch die Erfolgsquote bei den Zwischenprüfungen nicht einmal ein grob zutreffendes Bild der durchschnittlichen "Erfolgschancen" eines Studienanfängers wieder, letztlich in diesem Studiengang an dieser Hochschule auch einen Abschluss zu erwerben. Hierfür wäre zumindest eine Gegenüberstellung der Relation von Studienanfängerzahlen und Absolventenzahlen notwendig, die auch als Absolventenquote bezeichnet werden kann und im nächsten Abschnitt diskutiert wird.

Die Erfolgsquote kann also aus methodischen Gründen ebenfalls nur als Orientierung und ergänzende Information zu den Durchschnittsnoten dienen und stellt kein "objektives" Leistungskriterium dar. Hinzuzufügen wäre, daß bei einer Verwendung von Erfolgsquoten bei Prüfungen als Leistungskriterium (und der Definition hoher Erfolgsquoten als positiv) die Folgen für den Ruf einer Institution und evtl. für die Qualität eines Hochschulabschlusses auf dem Arbeitsmarkt unabsehbar sind. Eine Verwendung als Leistungsindikator könnte weiterhin ähnlich wie bei der Verwendung von Noten dazu führen, daß generell alle Hochschulen allen Studenten zu einem Abschluss zu verhelfen trachten, was sich bei Personalentscheidern in Unternehmen aber bald herumsprechen dürfte. Umgekehrt könnten sich Hochschulen mit sehr niedrigen Erfolgsquoten aufgrund ihrer vorgeblich sehr hohen Leistungsanforderungen auch als "Elite"-Hochschulen zu profilieren versuchen, obwohl es keine weiteren Hinweise auf eine hohe Ausbildungsqualität gibt.

### **5.2.9 Absolventenquote als Indikator für „Produktivität“?**

Aus der Studienanfänger- und der Absolventenzahl wird häufig eine beliebte Kennziffer berechnet, die zum Teil fälschlicherweise als Abbruchquote bezeichnet wird. Eine solcherart berechnete Abbruchquote ist jedoch als nicht zutreffend einzuschätzen, wenn damit die Anzahl bzw. die relative Häufigkeit der Studienabbrecher dargestellt werden soll (vgl. z.B. Tinto 1993, HIS 1995, Ziegele 1997, HIS 2002). Das liegt daran, daß in einigen Fällen eine erhebliche Anzahl von Studierenden während ihres Studiums die Hochschule wechselt. So kommt es, daß einige Hochschulen nach dem Grundstudium einen sehr starken Schwund haben, während andere Hochschulen einen sehr starken Zulauf erfahren. So hatten von den Soziologiestudierenden im Hauptstudium an der Universität Bielefeld 39% zuvor bereits an

einer anderen Universität studiert (vgl. Hornbostel & Daniel 1996). Zwar war hier die Fragestellung nicht ganz eindeutig, da Studierende, die ein Zweitstudium an der Universität Bielefeld aufnahmen und zuvor an einer anderen Universität studierten, ebenfalls mit erfasst wurden, jedoch ist deren Zahl erfahrungsgemäß nicht so groß, daß das Ergebnis wesentlich beeinflusst würde. Deshalb wäre es berechtigt, hier nur noch von "Schwundquote" zu sprechen, solange nicht Wanderungsbewegungen der Studierenden berücksichtigt werden (vgl. Webler u.a. 1993). Ziegele (2002: 11) bezeichnet dies auch als „Drop-Outs“.

Hiermit ist jedoch nicht die Schwundberechnung nach dem sogenannten Hamburger Verfahren gemeint, die für die nach Kapazitätsverordnung zu errechnende Kapazität eines Studienganges verwendet wird, da hier ein Durchschnittswert über alle Semester anhand eines Beobachtungszeitraumes von lediglich 4 Semestern errechnet wird. Diese Art der Berechnung wird zwar von den Landesministerien, beispielsweise dem Hessischen Ministerium für Wissenschaft und Kunst, für ausreichend gehalten (vgl. MWK/ Arnold: undatierte Arbeitsanweisung). Das Hamburger Verfahren ist jedoch aufgrund der "Hochrechnung" anhand eines nur kurzen Beobachtungszeitraumes deutlich ungenauer als andere, anhand von Lehrberichten mögliche Verfahren – obwohl auch diese durchaus Wünsche offenlassen. So können mangels Zugriff auf Immatrikulationsdatenbanken meist keine echten Kohortenanalysen (Analysen nach Immatrikulationsjahrgängen) vorgenommen werden. Es kann somit nicht genau nachvollzogen werden, welcher Student/ welche Studentin welchen Jahrganges sich in welchem Semester befindet und ob genau diese/r dann nach Ablauf von 5 oder 6 Jahren einen Abschluß erwarb oder ein anderer Student/ eine andere Studentin. Ebenfalls unberücksichtigt bleiben (müssen) bei der Ermittlung der sogenannten Schwundquote oft die Fachrichtungswechsler, da die Hochschulen hierzu oft keine Informationen oder keine Informationen auf Studiengangsebene liefern. Gleiches gilt natürlich für eine Absolventenquote, die sich (als Differenz zu 100%) aus der Abbruchquote ergäbe (also als Anzahl der Absolventen in einem bestimmten Zeitraum im Verhältnis zur Anzahl der Studienanfänger, die mit ihnen zusammen das Studium begonnen hatten). In anderen Veröffentlichungen wird für die Absolventenquote synonym auch die Bezeichnung "Erfolgsquote" verwendet (vgl. z.B. HIS 2004: 15). Um Verwechslungen mit der Erfolgsquote bei Abschlussprüfungen auszuschließen, soll nachfolgend nur noch von

Absolventenquote gesprochen werden.<sup>162</sup> Die Fachrichtungswechsler sind zwar für das jeweilige Institut genauso "verloren", aber aus gesamtgesellschaftlicher Perspektive anders zu bewerten als die "echten" Studienabbrecher (vgl. Ziegele 1997): Während Fachrichtungswechsler meist wegen enttäuschter Erwartungen (z.T. auch aufgrund mangelnder Vorinformation) ihr zunächst begonnenes Studienfach aufgeben, um nach gründlicherer Information dann oft sehr zielstrebig in einem anderen Studienfach weiterzustudieren, halten die "echten" Studienabbrecher offenbar einen Studienabschluss generell für nicht notwendig, sinnvoll oder machbar. Dies zeigt bereits, daß die Gründe für das Handeln der Studierenden und daraus folgend die Ansätze, wie man etwas für eine höhere Absolventenquote tun könnte, sehr unterschiedlich sein können. Hierzu bedarf es detaillierterer Untersuchungen zu den Gründen des Studienabbruchs bzw. des Fachrichtungswechsels, die an dieser Stelle nicht diskutiert werden können. Hinweise hierzu geben z.B. die von HIS (1995) durchgeführten Studierendenbefragungen, von denen einzelne Fragestellungen ggf. auf eine Übernahme in Anträge auf Fachrichtungswechsel bzw. auf Exmatrikulation geprüft werden könnten. Unter Hinzuziehung von Studierenden- und Absolventenbefragungen erarbeitete HIS (2002) inzwischen noch eine genauere Studienabbruchstatistik. Im Wesentlichen gingen sie bei der Berechnung des „Schwundes“ jedoch analog der Berechnung der Absolventenquote vor, wie sie weiter oben beschrieben wurde.

In den amtlichen Statistiken werden die Gründe für Abbruch oder Wechsel bisher oft nicht oder nicht vollständig erfasst. Hinzu kommt, daß es aus datenschutzrechtlichen Gründen in der Bundesrepublik so gut wie unmöglich ist festzustellen, wie viele der Studienanfänger an welcher Hochschule auch tatsächlich ihr Studium abgeschlossen haben, da keine Studienverlaufsstatistiken erstellt werden dürfen. Durch ihre eigene, auf anonymer Befragung basierende Untersuchung konnten Hornbostel & Daniel deshalb zwar zeigen, daß die Absolventenquote in der Soziologie an der oben bereits genannten Universität Bielefeld mit 82% fast doppelt so hoch ist wie beispielsweise an der Uni Wuppertal. Aber dies könnte zu einem Großteil auf die nach dem Grundstudium zur Universität Bielefeld wechselnden Studierenden zurückgeführt werden, da – wie bereits erwähnt – 39% der Bielefelder Soziologiestudierenden zuvor an einer anderen Universität immatrikuliert waren und Bielefeld bei Rankings regelmäßig weit vorn lag. Bundesweit lagen dazu bis vor einiger Zeit

---

<sup>162</sup> Alternativ könnte auch von Absolventen-Anfänger-Relation gesprochen werden. Obwohl dieser Begriff den hier betrachteten Aspekt wohl am besten trifft, scheint zu unhandlich, so daß er nicht weiter Verwendung findet.

keine Daten vor, so daß es bei der Vermutung bleiben muss, daß viele auf diese Weise ermittelte Absolventenquoten, nähme man sie als "harte" Leistungskriterien für die "Güte" der Lehre an einer Universität, "verzerrt" sein können und daher für Schlussfolgerungen erst weitere Untersuchungen nötig sind.

Welchen Begriff man jedoch auch immer für dieses in den Geistes- und Sozialwissenschaften stark ausgeprägte Phänomen verwendet, die Bedeutung bleibt für alle hiervon betroffenen Fächer bzw. Fachbereiche dieselbe: Da die Finanzierung der Hochschulen sich mehr oder weniger nach Studentenzahlen (in der Regelstudienzeit) bemisst, muss ein hoher Schwund bzw. eine niedrige Absolventenquote negativ bewertet werden (vgl. Langer u. a. 2001: 49). Sehr niedrige Absolventenquoten haben beim derzeitigen Finanzierungssystem (und wohl auch in Zukunft) nicht nur ungünstige Betreuungsrelationen im Grundstudium zur Folge, sondern müssen auch als Ausdruck enttäuschter Erwartungen jedes einzelnen betroffenen Studienanfängers und letztlich als Verlust an Lebenszeit und -qualität gesehen werden. Studentenbefragungen zeigen, daß viele Studenten der Geistes- und Sozialwissenschaften mit Erwartungen in das Studium gehen, die nicht erfüllt werden können und die sich zu einem großen Teil auf verfehlte oder schlicht fehlende Beratung zurückführen lassen. Solange aber ein großer Teil der Studenten zumindest das Grundstudium noch absolviert, bestehen für Hochschulen kaum Anreize, hier etwas zu ändern. Offenbar kommt man in der Politik jedoch mehr und mehr zu der Ansicht, daß hier etwas zu ändern ist. So wurde die Absolventenquote in den letzten Jahren nicht nur in die Lehrberichtsverordnungen mehrerer Bundesländer aufgenommen, sondern soll (z.B. in Berlin) auch als eine von mehreren Kennziffern und Indikatoren mit zur Finanzierung der Hochschulen herangezogen werden. Angesichts dieser Entwicklungen erscheint es in jedem Falle sinnvoll, über Absolventenquoten zu informieren, um mögliche Entscheidungen und ihre Grundlagen möglichst schon im Vorfeld transparent zu machen. Allerdings wäre es für eine bessere Interpretierbarkeit und die Suche nach möglichen Veränderungsansätzen hilfreich, wenn künftig auch auf Studiengangsebene zumindest der Anteil der Fachwechsler, besser auch der Hochschulwechsler ausgewiesen wird, um besser Schlussfolgerungen aus Absolventenquoten ziehen zu können. Dies ist leider derzeit meist (noch) nicht geschehen. Insbesondere bei Erfolgsquoten um die 100 Prozent und dennoch relativ geringen Absolventenzahlen sind – wie bereits am Schluß der Ausführungen zu Erfolgsquoten bei Prüfungen erwähnt – Absolventenquoten eine notwendige Ergänzung der bereits beschriebenen Kennziffern und Indikatoren.

Um die Absolventenquoten aussagekräftiger zu machen, sollten sie aber über mehrere Jahre gemittelt werden. Hierdurch können relativ große Schwankungen zwischen einzelnen Abschlussjahren, wie sie im Projekt zur Erfassung und Aufarbeitung von Lehrberichtsdaten in Sachsen (im Unterschied zu den anderen in den vergangenen Jahren dort verwendeten Indikatoren) festgestellt wurden, ausgeglichen werden (vgl. Krempkow/ König 2003, 2004).<sup>163</sup>

Zweitens sollte der Zeitraum, der zwischen der erhobenen Studienanfängerzahl und der hierzu in Beziehung gesetzten Absolventenzahl vergangen ist, der durchschnittlichen Studiendauer entsprechen, aber dennoch für eine bessere Vergleichbarkeit und Nachvollziehbarkeit einheitlich sein. So wurde entsprechend der Entwicklung der durchschnittlichen Fachstudiendauer der für das sächsische Projekt verwendete Zeitraum seit dem Jahr 2003 von fünf auf sechs Jahre erhöht. Diese Erhöhung ergibt sich aus der insgesamt in ganz Sachsen gestiegenen Studiendauer, denn inzwischen wurde in einer relativen Mehrheit der Studiengänge in Sachsen von den Absolventen eine Studiendauer von rund 11 und 12 Fachsemestern benötigt. Während nur noch in 39% der Studiengänge 9 und 10 Fachsemester benötigt werden, beträgt in 46% der Studiengänge die Studiendauer 11 und 12 Fachsemester. Inzwischen wird zudem in rund 10% der Studiengänge sogar rund 13 Fachsemester studiert.

Da die Daten (unter Hinzuziehung des jeweiligen Vorjahres) über zwei Jahre gemittelt wurden, gleichen sich „vorzeitige Abgänge“ ebenso wie „Überhänge“ von Studierenden aus vorhergehenden Jahrgängen zumindest zu einem relativ großen Teil aus (vgl. auch HIS 2004: 16)<sup>164</sup>.

Neben dem Ausgleich von Schwankungen zwischen einzelnen Abschlussjahren kann durch die Mittelung die Zuverlässigkeit der Daten auch noch in anderer Hinsicht erhöht werden. Das Problem niedriger Fallzahlen (kleiner oder gleich 5), welches sich bei den ersten

---

<sup>163</sup> Dies war vor der Veröffentlichung des Studienführer Sachsen 2003 noch nicht möglich, da die Absolventen- und v.a. die Studienanfängerzahlen noch nicht für ausreichend viele Jahrgänge zurückgehend nach einheitlichem Raster vorlagen (vgl. Krempkow/ König 2003).

<sup>164</sup> Eine andere Möglichkeit stellt die Schaffung einer ‚synthetischen‘ Jahrgangsstärke dar, indem die Gesamtzahl der Studierenden durch die Regelstudienzeit dividiert wird (vgl. HIS 2004: 15). Allerdings ist diese insbesondere bei geringen Absolventenzahlen immer noch anfälliger gegenüber zufälligen Schwankungen, da sie sich nur auf einen Absolventenjahrgang bezieht. Außerdem vernachlässigt sie Absolventen, die in einem vertretbar erscheinendem Zeitraum von etwa 1 Jahr nach Ablauf der Regelstudienzeit ihren Abschluß erwerben, so daß es kaum Anreize gibt, Studierende auch in der Abschlußphase noch bis zum erfolgreichen Abschluß zu betreuen. Bei HIS (2004: 17) wird dies in der „Regelstudienzeitquote“ berücksichtigt, die der hier als Studierbarkeit bezeichneten Quote entspricht, bei der allerdings 2 Semester zur Regelstudienzeit addiert werden.

Analysen der sächsischen Lehrberichte v.a. bei neu eingerichteten Studiengängen mit noch wenigen Studienanfängern und/ oder wenigen Absolventen ergab (aber auch bei Studiengängen mit generell nur wenigen Absolventen), tritt bei Betrachtung mehrerer Jahrgänge fast überhaupt nicht mehr auf. Als nachteilig bei der Mittelung über mehrere Jahre erwies sich, daß Veränderungen der Ergebnisse erst später als Tendenzen deutlich werden. Da speziell bei der Absolventenquote jedoch relativ viele Veränderungen auftraten, wurde dieser Nachteil aus Sicht der Verfasser durch die Vorteile mehr als aufgewogen.

Zur Aussagekraft der Absolventenquote ist also festzuhalten, daß diese keine Aussagen über Fach- und Hochschulwechsler ermöglicht. Da diese nach den gegebenen Rahmenbedingungen aber in jedem Fall einen Verlust für die Hochschule bedeuten, scheint eine Berücksichtigung solcher Wechsler derzeit nicht vorrangig. Um künftig stärker zu erwartende Konflikte mit anderen diskutierten Zielen wie einer stärkeren Internationalisierung zu vermeiden, sollten diese jedoch möglichst bald mit erfasst und berücksichtigt werden. Alternativ wären evtl. zusätzliche Indikatoren zur Internationalisierung (wie z.B. der Anteil von Absolventen mit Auslandsaufenthalten) als Ausgleich unerwünschter Nebenwirkungen heranzuziehen.

Ebenfalls keine Aussagen möglich sind über Studierende, die außerhalb des betrachteten Zeitraumes von z.B. 6 Jahren noch einen Abschluß erwerben. Da die Regelstudienzeiten und die Förderungshöchstdauer nach dem BAföG aber deutlich unter 6 Jahren liegen und auch die Möglichkeit einer Studienabschlußförderung nach dem BAföG in den allermeisten Fällen spätestens nach zwölf Semestern endet und die Möglichkeit eines Abschlusses in jedem Fall unter 6 Jahren und möglichst in der Regelstudienzeit einhelliges politisches Ziel ist, erscheint dies vertretbar. (Wenn die Absolventenquote mit der Studierbarkeit gemeinsam als weiteres Indikator verwendet wird, gibt es auch einen Anreiz zur Einhaltung der Regelstudienzeit).

Werden außerdem die Absolventenquoten über mehrere Jahre gemittelt und orientieren sich die Berechnungszeiträume an der (überwiegenden) durchschnittlichen Fachstudiendauer, so lassen sich zwar aufgrund der in Deutschland nicht zulässigen Erhebung von Studienverlaufsdaten noch immer keine exakten Aussagen treffen. Dies hängt aber zugleich damit zusammen, daß der Einfluß der Hochschulen auf den Studienerfolg begrenzt ist, weil auch hochschulunabhängige Faktoren diesen beeinflussen (wie z.B. Studienfinanzierung, Arbeitsmarktprobleme, Krankheit u.ä., wobei dies meist in ähnlichem Ausmaß auf die Hochschulen einer Region zutrifft). Trotz dieser auch bei der Analyse anderer Länder

vorgebrachten Bedenken wird dort die Annahme als plausibel eingeschätzt, daß die Hochschulen jedenfalls mittel- bis langfristig den Studienerfolg gestalten können. Lediglich „eine zu starke Berücksichtigung dieses Lehrindikators scheint nicht ganz unproblematisch“ (Leszcensky u.a. 2004: 194). Daher scheint zumindest eine grobe Aussage über die Wahrscheinlichkeit eines Abschlusses in einem bestimmten Studiengang einer Hochschule, die z.T. auch als „Produktivität“ bezeichnet wird, mit Hilfe der Absolventenquote möglich. Einschränkend ist zur Absolventenquote dennoch hinzuzufügen, daß diese lediglich ein Indikator für die quantitative Seite der „Produktivität“ einer Hochschule in einem bestimmten Fach sein kann. Über die Qualität der ausgebildeten Absolventen kann hiermit keine Aussage gemacht werden.<sup>165</sup> Damit könnte bei Verwendung von Absolventenquoten ein Anreiz entstehen, die Anforderungen zu senken (vgl. Ziegele 2002: 12). Dies wurde auch bereits von anderen OECD-Ländern berichtet. Als weniger ausgeprägt wurde dies beschrieben, wenn wie in Schweden und Norwegen Kreditpunkte oder wie in Dänemark die Anzahl der durchgeführten Prüfungen (anstelle nur der bestandenenen) verwendet werden (vgl. Leszcensky u.a. 2004: 194). In der Analyse von OECD-Ländern zeigte sich zudem, daß durch den Einsatz entsprechender Qualitätssicherungssysteme diesen negativen Anreizwirkungen Grenzen gesetzt werden können. Belastbarere Aussagen zur Qualität nicht nur des Ausbildungsprozesses, sondern auch der Absolventen sind jedoch nur mit relativ aufwendigen Absolventenverbleibsstudien möglich, die es bislang in Deutschland erst an wenigen Hochschulen gibt (vgl. in Sachsen z.B. Pawlowski 1999 und 2002, Krempkow u.a. 1999 und Folgejahre).

Da es ausdrücklich das politische Ziel und die gesellschaftliche Aufgabe einer Hochschule ist, Studierende auszubilden und ihnen (in angemessener Zeit) einen Abschluß zu ermöglichen, sollte die Absolventenquote (oder ähnliche Indikatoren<sup>166</sup>) trotz der genannten Einschränkungen ihrer Aussagekraft als Indikator in Leistungsbewertungen und Leistungsanreizen Berücksichtigung finden. So wird im Rahmen des Berichtssystems der Bundes-

---

<sup>165</sup> So könnte es durchaus sein, daß eine Hochschule trotz einer geringeren Absolventenquote letztlich gesamtgesellschaftlich gesehen einen größeren Beitrag liefert, wenn deren Absolventen z.B. aufgrund ihrer deutlich besseren Kenntnisse und Kompetenzen einen schnelleren/ besseren Übergang ins Berufsleben erfahren.

<sup>166</sup> So diskutieren Ziegele (2002: 10) die Zahl der Absolventen pro Lehrkraft als Indikator mit ähnlichen Zielen. Andere Autoren wie z.B. Pasternack (2004: 117), Rammelt (2003) nennen die absolute Anzahl der Absolventen, die auch in den Niederlanden, Finnland und Belgien Verwendung findet (vgl. Leszcensky u.a. 2004: 189ff.). Die Absolventenquote „mißt jedoch den Studienerfolg direkter“ als die absolute Anzahl der Absolventen und ist daher angemessener für das Ziel, „die Qualität der Lehre und der Lehrorganisation so

regierung zur technologischen Leistungsfähigkeit Deutschlands formuliert: „Die eigentliche Outputgröße des Bildungssystems bilden die Hochschulabsolventen (...). Die bisherige und künftige Entwicklung der Absolventenzahlen determiniert damit in hohem Maße die Entwicklungsmöglichkeiten einer wissensbasierten Gesellschaft“ (vgl. HIS 2003: 28). Wie relevant Unterschiede in der Absolventenquote sein können, zeigt auch ein von Fohrbeck (2000: 178) vorgenommener Vergleich der USA und Großbritannien (mit Daten von 1995): Da in Großbritannien ca. 80% der Studierenden ihr Studium erfolgreich abschließen und in den USA nur 60%, „produziert“ Großbritannien im Verhältnis zur Bevölkerungszahl genauso viele Graduierte wie die USA (obwohl die Studierquote in den USA bei 52% und in Großbritannien bei 43% lag). Auch in neueren Untersuchungen zu *Studienabbruchquoten* im internationalen Vergleich (mit Daten von 1999) zeigen sich enorme Unterschiede zwischen den USA (34%) und Großbritannien (15%). Die höchste Abbruchquote hat Italien (58%), die niedrigste Japan (6%). Deutschland (30%) liegt im unteren Mittelfeld (vgl. HIS 2003: 24).

Letztlich ist die Verwendung der Absolventenquote auch im Zusammenhang mit möglichen Folgen einer alleinigen oder stark gewichteten Berücksichtigung von Studierendenzahlen in Leistungsanreizmodellen zu sehen, da dies wenig Anreiz für die Verwirklichung des o.g. Zieles einer Hochschule bietet, sondern im Gegenteil als Fehlanreiz einen möglichst langen Verbleib honorieren würde, ohne daß ein Anreiz zur Förderung des Studienabschlusses bestände.

#### **5.2.10 Fazit zur Aussagekraft „objektiver“ Kennziffern und Indikatoren**

Mit so genannten "objektiven" Kennziffern und Indikatoren wie Studenten- und Studienanfängerzahlen, Betreuungsrelation, Studiendauer und Durchschnittsnote, Erfolgsquote, Absolventenquote ist jeweils für sich genommen "nur ein sehr unvollständiges Bild der Studiensituation" zu zeichnen (vgl. Hornbostel & Daniel 1996). Das liegt zum größten Teil daran, daß bisher den offiziellen Stellen vorliegende hochschulstatistische Daten in ihrer Qualität oft nicht befriedigen können (z.B. zur Abbrecherquote). Eine Erhöhung der Aussagekraft durch Weiterverarbeitung vorhandener und durch Ergänzung mit weiteren Daten z.B. zum Hochschul- und Fachrichtungswechsel ist zwar grundsätzlich möglich, wie das Beispiel der Absolventenquote zeigen sollte. Dies ist allerdings auch entsprechend

---

verbessern zu wollen, daß möglichst viele Studierende eines Jahrgangs ihren Studienabschluß erreichen“ (vgl.

aufwendig. Selbst dann erlaubt dies jedoch keine exakte, sondern nur eine grobe Aussage über die Wahrscheinlichkeit eines Abschlusses in einem bestimmten Studiengang einer Hochschule und sagt noch nichts über die Qualität eines solchen Abschlusses und darüber, unter welchen Bedingungen er erworben wurde. Ersteres wäre nur mit relativ aufwendigen Absolventenverbleibsstudien möglich, letzteres wäre auch mit ohnehin durchzuführenden Studentenforschungen zur Qualität der Lehre erfaßbar. Andere, mit vergleichsweise geringem Aufwand verfügbare Indikatoren wie die Betreuungsrelation können selbst bei klarerer Definition der Aussagekraft nur potentielle Möglichkeiten für eine gute Betreuungssituation aufzeigen, nicht aber die tatsächlich realisierten. Dies wiederum könnte durch eine subjektive Einschätzung der Studierenden (oder auch retrospektiv von Absolventen) ergänzt werden.

Die Relation Bewerber pro Studienplatz ist aus inhaltlichen und methodischen Gründen als Indikator für die Attraktivität eines Studienganges ungeeignet, obwohl sie ebenfalls meist leicht zu berechnen wäre. Diese Gründe können auch nicht durch Hinzuziehen anderer Indikatoren ausgeglichen werden.. Nach einschlägigen Studien bildet die Relation Bewerber pro Studienplatz eher die Anzahl verfügbarer Studienberechtigter in der Umgebung eines Studienortes bzw. die Attraktivität des Studienortes ab (nicht aber die eines bestimmten Studienganges einer Hochschule). Da hierbei Lebensbedingungen wie niedrige Mieten, kulturelles Angebot usw. eine größere Rolle spielen als die Studienqualität, ist sie kein inhaltlich valides Kriterium für die hier verfolgten Zwecke. Weitere Gründe können darin liegen, daß die Datengrundlage z.T. uneinheitlich ist.

Bei der Fachstudiendauer als einem potentiellen Indikator für die Ergebnisqualität des Studiums sind ebenfalls Probleme der Vergleichbarkeit z.B. aufgrund unterschiedlicher Definitionen des Abschlußzeitpunktes oder der Mittelwertberechnung zu konstatieren. Zudem können Ursachen für Schwankungen auch außerhalb der Verantwortung der Hochschulen liegen, so daß eine Verwendung als alleinige oder zentrale Leistungskennziffer fragwürdig scheint. Zur genaueren Klärung von Ursachen für Schwankungen könnten ggf. auch hier Absolventenforschungen beitragen. Dennoch kennzeichnet die Studiendauer als wesentlicher Nachteil gegenüber der Studierbarkeit (neben der meist geringeren zugrunde liegenden Fallzahl) die mangelnde Zeitnähe. Wenn es Probleme oder auch Verbesserungen bezüglich der Möglichkeit gibt, das Studium in der vorgesehenen Zeit abzuschließen, zeigt

---

HIS 2004: 15).

sich dies erst, wenn Studierende als Absolventen die Hochschule verlassen. Verlassen Studierende erst kurz vor der Zwangsexmatrikulation z.B. im 14. Semester noch den Studiengang, weil es in einem anderen Bundesland keine solche Regelung gibt, zeigt sich dies in der Studiendauer überhaupt nicht. Die Studierbarkeit kann daher insgesamt als besser geeigneter Indikator angesehen werden als die Fachstudiendauer.

Auch die Noten, nach wie vor ein Einstellungskriterium für Arbeitgeber, sind nur unter bestimmten Bedingungen valide und zuverlässig, die in Deutschland meist nicht erfüllt sind. Zudem korrelieren Noten zumindest an einzelnen Hochschulen kaum oder nicht in der gewünschten Richtung mit dem beruflichen Erfolg. Noten können also am ehesten noch ein Maß für die Notenvergabepraxis im jeweiligen Studienfach sein, jedoch keinesfalls ein absolut zu sehendes zuverlässiges Maß für die Leistungsfähigkeit der Absolventen oder des Studienfaches. Dasselbe gilt in ähnlicher Weise auch für die Erfolgsquoten bei den Abschlußprüfungen. Diese sind als Ergänzung zu den Noten wichtig, da in die Durchschnittsnoten der Absolventen nur die bestandenen Prüfungen einfließen. Die Erfolgsquoten wiederum sollten – insbesondere in Fächern mit hohen Erfolgsquoten und dennoch geringen Absolventenzahlen – durch Absolventenquoten ergänzt werden. Umgekehrt ermöglichen die Erfolgsquoten ergänzend zu den Absolventenquoten Aussagen, welcher Anteil des „Schwundes“ auf endgültiges Nichtbestehen von Abschlußprüfungen zurückzuführen ist.

Vergleiche und Interpretationen der sogenannten „objektiven“ Kennziffern und Indikatoren bedürfen also neber einer sorgfältigen und sachkundigen Analyse der gegenseitigen Ergänzung, aber auch der Ergänzung durch „subjektive“ Bewertungen von Absolventen, Studierenden oder auch Lehrenden. Die genannten Kennziffern und Indikatoren können zwar durchaus als Orientierungshilfe für Studieninteressierte oder für Arbeitgeber zur Einordnung von Stellenbewerbern im Vergleich zu anderen Absolventen desselben Studienganges dienen. Sie sollten daher regelmäßig für einen breiten Leserkreis veröffentlicht werden. Für die Zukunft sollte jedoch geprüft werden, inwieweit zu diesen Kennziffern und Indikatoren stärker auch subjektive Bewertungen verfügbar gemacht werden können. Für die Auswertungen von Studiengängen kann zwar, wenngleich nicht vollständig für alle Hochschulen und Fächer vorhanden, eine Auswahl subjektiver Bewertungen aus bundesweiten Studienführern wie dem des CHE herangezogen werden. Allerdings sind diese z.T. bereits mehrere Jahre alt. Eigene Erhebungen wären daher aus Gründen einer zu verbessernden Vollständigkeit und Aktualität und nicht zuletzt auch aufgrund eines größeren

Einflusses auf die Fragestellungen und einer zeitnahen Verfügbarkeit von Daten zur Ergänzung der bereits vorhandenen „objektiven“ Kriterien und Indikatoren sehr wünschenswert.

Für eine ausgewogene Erfassung der Studienqualität im Sinne einer Leistungsbewertung und mit dem Ziel, Anreize für Leistung *und* Qualität zu setzen, wäre die Einbeziehung subjektiver Bewertungen darüber hinaus eine wichtige Voraussetzung, da die bisher überwiegend diskutierten „objektiven“ Kennziffern und Indikatoren Aspekte der Qualität nicht erfassen können. Hierdurch könnte es leicht zu einer Fehlsteuerung in Richtung höherer „Stückzahlen“ von Absolventen kommen, die nicht in der gewünschten und erforderlichen Qualität ausgebildet werden. Außerdem wären einige der hier diskutierten Kennziffern und Indikatoren, wenn sie nicht vorwiegend der Information und Orientierung, sondern zur Verwendung in Anreizsystemen dienen sollen, ggf. anders zu berechnen und z.B. über weitere Jahre zu mitteln, um sie – insbesondere bei kleinen Studiengängen – weniger abhängig von evtl. zufälligen Schwankungen zu machen (dann allerdings auf Kosten der Aktualität bzw. Zeitnähe). Diese Diskussion kann an dieser Stelle nicht abschließend geführt werden, da hierzu bislang zu wenig Erfahrungen in Deutschland vorliegen. Dennoch wird sie vermutlich an Bedeutung gewinnen. Daher soll auch darauf hingewiesen werden, dass hierzu zusätzliche Analysen und Auswertungen sowie Weiterentwicklungen notwendig sind, um den gewünschten Zielen möglichst nahe zu kommen. Hierbei wären Erfahrungen aus anderen Ländern teilweise hilfreich, da eventuelle Fehlanreizwirkungen von vornherein bedacht werden können. Eine direkte Übertragbarkeit von teilweise bereits langjährigen Erfahrungen wie in Großbritannien oder Australien ist jedoch aufgrund der anderen Struktur und ggf. andersartigen Wirkungsweisen nicht gegeben. So würde beispielsweise die Verwendung von stark gewichteten Studierendenzahlen als Indikator für Leistungsanreize ohne entsprechendes Korrektiv vermutlich zu unerwünschten Nebenwirkungen wie der Absenkung der Qualität führen. In Ländern, die dies bereits anwenden, kann das z.B. durch ein (im Vergleich zu Deutschland) stärkeres Qualitätsmanagement und eine hohe Transparenz über die Qualität von Studium und Lehre an den Hochschulen vermieden oder jedenfalls vermindert werden.

Zusammenfassend bleibt festzuhalten, daß von den diskutierten Kriterien und Indikatoren für die meisten zwar ein relativ geringer Erhebungsaufwand anfällt, da sie häufig bereits in dieser oder ähnlicher Form erfasst werden. Jeweils für sich genommen kann aus den verschiedensten Gründen keines allein als wirklich „objektiv“ gelten. Als am meisten

problematisch erscheinen für das Ziel, Leistung einschließlich Qualität zu fördern, die Verwendung von Studenten-, Studienanfänger- und Bewerberzahlen, aber auch von durchschnittlichen Abschlußnoten. Auch die Betreuungsrelation kann nur das Potential für eine gute Betreuungssituation aufzeigen, nicht aber die tatsächlich realisierte. Ebenfalls nicht unproblematisch, aber etwas besser geeignet erscheinen (im Vergleich zur häufig diskutierten durchschnittlichen Fachstudiendauer) die Studierbarkeit sowie die Absolventenquote, wobei diese unter Einbeziehung zusätzlicher Informationen weiterentwickelt werden sollte. Insgesamt ermöglicht erst ein „Datenkranz“ aus mehreren, sich gegenseitig ergänzenden Kennziffern und Indikatoren unter Einbeziehung subjektiver Indikatoren ein realistisches (i.S.v. der Realität möglichst nahe kommendes) Bild der Studiensituation eines Faches an einer bestimmten Hochschule.

Als subjektive Indikatoren ergänzt werden könnten beispielsweise Einschätzungen von Studierenden und/ oder Absolventen zur Betreuungssituation, aber auch Einschätzungen, ob es aus Sicht der Studierenden/ Absolventen in der Studienorganisation Gründe gab, die eine Einhaltung der Regelstudienzeit erschwerten (und ggf. um welchen Zeitraum). Daneben können zusätzlich zu solchen sich direkt aus der Ergänzung „objektiver“ Indikatoren ergebenden weitere subjektive Indikatoren nützlich sein. Dies soll in nachfolgendem Abschnitt diskutiert werden.

### **5.3 Subjektive Bewertungen als Leistungskriterien?**

In den vorhergehenden Abschnitten wurde dargestellt, welche Probleme die sogenannten objektiven Kriterien bzw. Indikatoren bei der Verwendung als Leistungsindikatoren noch in sich bergen. Doch selbst wenn deren methodische Probleme sich im Laufe der Zeit mindern und sie sich damit zuverlässiger machen ließen, so können die meisten sogenannten „objektiven“ Kriterien kaum den Aspekt der Qualität erbrachter Leistungen erfassen - und schon gar nicht das bei angestrebter stärkerer Zielgruppenorientierung wichtiger werdende subjektive Qualitätsempfinden der Zielgruppen. Nachfolgend sollen die Möglichkeiten und Grenzen der Einbeziehung subjektiver Bewertungen durch verschiedene Gruppen von Hochschulmitgliedern diskutiert werden. Diese Gruppen sind Absolventen und Studierende, aber natürlich auch Lehrende. Für diese wird jeweils anhand bereits erprobter Befragungsinstrumente und eigener Untersuchungen die Eignung von Ergebnissen solcher Befragungen

als Kriterien für Leistungsbewertungen und Leistungsanreize diskutiert. Diese Gruppen entsprechen im Wesentlichen auch den bereits am Anfang dieser Arbeit im Abschnitt „Zielgruppen für die Qualität der Hochschullehre“ diskutierten, wobei die Befragungen von Absolventen Informationen über die Perspektive der einstellenden Unternehmen bzw. der Wirtschaft ermöglichen.<sup>167</sup>

### 5.3.1 Die Bedeutung subjektiver Bewertungen für die Qualitätsorientierung

Folgt man oben genannter Argumentation, so können subjektive Bewertungen eine immense Bedeutung für Qualitätsorientierung erlangen. Deutlich wird dies insbesondere, wenn Hochschulabsolventen als ein „Produkt“<sup>168</sup> der Hochschulen aufgefaßt werden und die Qualität der Hochschulbildung anhand der Anzahl der Absolventen (evtl. im Verhältnis zur Anzahl der immatrikulierten Studienanfänger des jeweiligen Absolventenjahrganges), anhand deren Studiendauer, deren Ausbildungskosten usw. bewertet werden soll. Denn diese Kennziffern sagen zwar etwas über Quantitäten bzw. die Effizienz der Hochschulbildung aus, jedoch nichts darüber, ob die Absolventen tatsächlich *gut* ausgebildet wurden. Über die Qualität sind mittels objektiver Kennziffern und Indikatoren nur eingeschränkt Aussagen möglich. Eine Einbeziehung dieser Aspekte und damit Aussagen darüber sind jedoch notwendig, wenn die Gefahr von Fehlanreizwirkungen vermindert werden soll, wie auch bereits im Abschnitt zu Leistungsanreizmodellen erläutert wurde. In diesem Zusammenhang betonten mehrere Autoren die Bedeutung qualitativer bzw. subjektiver Indikatoren in Modellen des Qualitätsmanagements und Leistungsanreizmodellen, (wobei dies leider häufig nicht konkreter ausgeführt wurde). Dagegen nannten einige Autoren, die sich vor allem mit Indikatoren beschäftigten, konkrete Möglichkeiten zur Einbeziehung subjektiver bzw. qualitativer Indikatoren, z.B. durch die Befragung von Absolventen und Studierenden (vgl. Pasternack 2004: 120, Ziegele 2002: 11ff., Tropp 2002: 118, HRK 2000b: 17, Sendldorfer 1997: 64). Solche Möglichkeiten wurden auch in den im Rahmen dieser Arbeit durchgeführten Experteninterviews angesprochen (vgl. entspr. Abschnitt). Die *Qualität* der Ausbildung entscheidet schließlich darüber, ob Studierende ggf. später z.B. in Unternehmen als Absolventen eingestellt werden (oder ob sie als

---

<sup>167</sup> Eine Einbeziehung von weiteren subjektiven Einschätzungen aus „der Gesellschaft“ als Zielgruppe kann nicht erfolgen, da „die Gesellschaft“ nicht unmittelbar, sondern überwiegend durch vermittelnde Institutionen auf die Qualitätsorientierung und die Strukturen an Hochschulen einwirkt und deren Untersuchung den Rahmen der hier vorgelegten Arbeit sprengen würde.

Selbständige erfolgreich sind). Folgt man diesen Überlegungen, dann ließe dementsprechend der Verbleib der Absolventen Rückschlüsse auf die Qualität der Ausbildung zu, wobei allerdings eine ganze Reihe intervenierender Variablen zu berücksichtigen wäre (wie z.B. Arbeitsmarktlage, Struktur der Absolventen bezüglich Geschlecht usw.).

### **5.3.2 Möglichkeiten zur Erhebung subjektiver Bewertungen**

Will man nachhaltige Qualitätssicherung betreiben und dabei den Prozeß auch zeitnah steuern, sollte nicht nur die Ergebnisqualität (und die Absolventen als "Endprodukt") betrachtet werden, wie bereits bei der Diskussion der Konzepte und Modelle ausgeführt wurde. Vielmehr erscheint daneben auch eine Betrachtung des Prozesses bzw. der Prozeßqualität notwendig. Diese kann durch dreierlei Personengruppen eingeschätzt werden, die an diesem Prozeß beteiligt sind bzw. waren, wobei jede ihre eigene Perspektive hat und die Hochschule bzw. der Fachbereich entscheiden muß, in welchem Rahmen die jeweilige Perspektive leitend für ihre Qualitätsorientierung sein soll. Die drei Personengruppen sind:

1. Studierende
2. Absolventen
3. Lehrende

*Studierende* können aus ihrer Perspektive insbesondere die aktuelle Situation von Lehre und Studium beurteilen. Während Lehrveranstaltungsbewertungen inzwischen in den meisten Bundesländern gesetzlich verankert sind und in vielen Fächern regelmäßig durchgeführt werden, findet eine als genauso wichtig einzuschätzende Bewertung der Studienbedingungen bisher meist nur unregelmäßig statt. Eine Erfassung aller wesentlichen Dimensionen der Situation der Lehre und des Studiums ist jedoch nur einschließlich ihrer Rahmenbedingungen möglich. Zudem können die Rahmenbedingungen wichtige Korrekturfaktoren (auch im Sinne von potentiell die Ergebnisse verzerrenden "bias"-Variablen) zur Einordnung und Interpretation der Ergebnisse studentischer Bewertungen von Lehre und Studium darstellen. Wie die besondere Perspektive der Studenten umfassend und dennoch mit vertretbarem Aufwand erfaßt werden kann, soll im Abschnitt zu Studentenerbefragungen dargestellt werden. Hierbei kann auf die umfassenden Analysen und Auswertungen zum

---

<sup>168</sup> Zur kritischen Diskussion über die Angemessenheit von Begriffen aus der Betriebswirtschaft vgl. die

Heidelberger Inventar der Lehrveranstaltungsevaluation (HILVE) sowie auf eigene Analysen und Erfahrungen mit einer Adaption dieses Instrumentes bei der Durchführung und Auswertung von Studentenbefragungen an der TU Dresden (und an 9 weiteren Hochschulen in Sachsen) verwiesen werden, die im Zuge der Etablierung des Projektes "Evaluation der Lehre" an der TU Dresden und des Studentischen Evaluationsbüro Sachsen (SES) gesammelt wurden.<sup>169</sup> Außerdem werden weitere Studentenbefragungen für einen Vergleich von studentischen Bewertungen der Situation von Lehre und Studium und der Bewertung durch Absolventen herangezogen. Dies geschieht beispielhaft für die Wirtschaftswissenschaften anhand der Ergebnisse einer 1998 am Institut für Soziologie durchgeführten Studentenbefragung und der Ergebnisse der Studentenbefragung für die Imagestudie der TU Dresden (Lenz u.a. 2000) sowie der Ergebnisse der Dresdner Absolventenstudien (vgl. Krempkow 2000, Krempkow/ Popp 2003b).

*Absolventen* können das Studium (retrospektiv) insbesondere hinsichtlich der auf dem Arbeitsmarkt geforderten Qualifikationen und Kenntnisse beurteilen. Außerdem kann die Einschätzung der Lehre und der Studienbedingungen durch Absolventen – wie oben erwähnt – für einen Vergleich ihrer Einschätzungen mit denen von Studenten sein, da sie bereits einigen Abstand zum Studium gewonnen haben. Vor allem wird im Abschnitt zu Absolventenbefragungen jedoch dargestellt, wie der berufliche Verbleib bzw. der Erfolg der Absolventen auf dem Arbeitsmarkt erfaßt werden kann, da dieser als Indikator für die langfristigen Wirkungen des Studiums gilt. Zusätzlich sind mittels Zusammenhangsanalysen auch unabhängig von individuellen Einschätzungen der Absolventen Rückschlüsse auf Einflussfaktoren bzw. Bestimmungsgründe für den beruflichen Erfolg möglich, die den einzelnen Absolventen evtl. nicht (oder nicht in entsprechendem Ausmaß) bewusst sind.<sup>170</sup> Inzwischen kann auf mehrjährige Erfahrungen mit umfangreichen Absolventenstudien in allen Fakultäten der TU Dresden mit insgesamt über 2000 Befragten zurückgegriffen werden (vgl. Krempkow/ Popp 2003b).

*Lehrende* können sowohl als Peers (externe bzw. nicht derselben Hochschule angehörenden Experten), aber auch als hochschulinterne Experten und Verantwortungsträger fungieren. Außerdem lassen sich die Lehrenden nach dem Modell der Gruppenuniversität in

---

Anmerkungen zum Kundenbegriff im Abschnitt zu Modellen des Qualitätsmanagements.

<sup>169</sup> Die Leitung des TU-Projektes hatte der Verfasser von Mitte 1997 bis Ende 2000 inne. Mitarbeiter des Studentischen Evaluationsbüro Sachsen war er von Anfang 1999 bis Ende 2000 (vgl. Krempkow/ Heldt 1999, 2000 sowie Krempkow/ Winter 2000).

Professoren und (wissenschaftliche) Mitarbeiter einteilen, wobei auch letztere oft große Teile der Organisation und Durchführung der Lehre übernehmen und deshalb hier ausdrücklich einbezogen werden sollen. Außerdem kann die Perspektive der Lehrenden wichtig sein, auch ohne daß sie eine besondere Verantwortung übernehmen; nämlich, indem sie als Mitarbeiter und damit wichtiger "Produktivitätsfaktor" des „Unternehmens“ Hochschule aufgefaßt werden. Durch Lehrendenbefragungen ähnlich wie bei Mitarbeiterbefragungen in Unternehmen ließe sich dann nicht nur deren besondere Perspektive erfassen, sondern zugleich Arbeitsbedingungen, Arbeitszufriedenheit, Akzeptanz von geplanten Veränderungen usw.. Wie dies geschehen kann, wird im Abschnitt zu Lehrendenbefragungen vorgestellt. Dabei wird die umfangreiche Literatur zu Mitarbeiterbefragungen in Unternehmen nur in spezifischen Zusammenhängen herangezogen, da Lehrende an Hochschulen einen deutlich anderen Status innehaben und die Bedingungen an Hochschulen nicht ohne weiteres mit den Bedingungen in Unternehmen vergleichbar sind. Es kann jedoch auf die Erfahrung mit eigenen Lehrendenbefragungen an mehreren Hochschulen und auf bereits Anfang/ Mitte der 1990er Jahre durchgeführte bundesweite Lehrendenbefragungen mit vergleichbaren Fragestellungen verwiesen werden, so daß dennoch auf bereits erprobte Befragungsinstrumente zurückgegriffen werden konnte.

Befragungen von Studenten werden im Vergleich zu Absolventen- und Lehrendenbefragungen kritischer diskutiert. So wurden in den letzten Jahren im Zusammenhang mit studentischen Lehrveranstaltungsbewertungen, aber auch im Zusammenhang mit Hochschulrankings eine Reihe von wissenschaftlichen Beiträgen publiziert, die sich skeptisch zur Eignung von Studentenbefragungen für Qualitäts- und erst recht für Leistungsbewertungen äußern. Auch in den für diese Arbeit durchgeführten Experteninterviews äußerten sich die Interviewten zu Studentenbefragungen kritischer als zu Absolventenbefragungen. Andererseits gibt es umfassende Untersuchungen, die für viele von den Skeptikern vorgetragenen Bedenken keine Belege finden. Da solche wissenschaftlichen Kontroversen weder in den vergangenen Jahren noch aktuell etwas an der festen Absicht von Hochschulpolitikern änderten, das Urteil von Studierenden in Leistungsbewertungen einzubeziehen (vgl. Zitate der Bundesbildungsministerin am Anfang dieser Arbeit), soll nachfolgender Abschnitt mit einer ausführlichen Diskussion der Eignung

---

<sup>170</sup> So erwiesen sich praktische Erfahrungen vor oder während des Studiums für viele Fächer als wichtiger als eine kurze Fachstudiendauer oder (sehr) gute Noten.

von Studentenforschungen begonnen werden. Hierbei werden die von den Skeptikern vorgetragenen Bedenken aufgegriffen und anhand empirischer Untersuchungen überprüft.

## **5.4 Die Eignung von Studentenforschungen zur Qualität der Lehre**

Studentische Lehrbewertungen gehören an vielen Hochschulen als Instrument zur Rückmeldung über die Qualität von Lehrveranstaltungen bereits zum Alltag. Einige Hochschulen und Fachbereiche vergeben auf der Grundlage von Studentenforschungen auch Preise „für gute Lehre“. Unter vielen Hochschullehrern bestehen jedoch nach wie vor Vorbehalte gegenüber Studentenforschungen zur Qualität der Lehre. Ein Teil dieser Vorbehalte beruht oft auf unklaren Zielvorstellungen, ein anderer darauf, daß die Ergebnisse empirischer Untersuchungen zur Validität und Zuverlässigkeit von Studentenforschungen in Deutschland immer noch zu wenig bekannt sind. Da Validität auch eine zentrale Anforderung an Indikatoren ist, sollen in folgendem Abschnitt deshalb zunächst zentrale Ergebnisse solcher Studien vorgestellt und diskutiert werden. Daneben wird die Vergleichbarkeit als wichtige Anforderung an Indikatoren genannt, so daß sie hier für die Studentenforschungen mittels der vorhandenen Studien ebenfalls zu diskutieren ist. Da hierbei auch die Berücksichtigung eines potentiellen Noteneinflusses erfolgt, kann unter dem Stichwort Vergleichbarkeit zugleich angesprochen werden, ob die Nutzung der Ergebnisse von Studentenforschungen als möglicher Leistungsindikator unerwünschte Anpassungseffekte wie z.B. eine mildere Notenvergabe bzw. eine Absenkung der Leistungsanforderungen zur Folge haben können. Schließlich erfolgt die Diskussion weiterer Anforderungen an Studentenforschungen, wie Zeitnähe und das Verhältnis von Nutzen und Erhebungsaufwand.

### **5.4.1 Wie kann Lehrqualität mittels Studentenforschungen „gemessen“ werden?**

Wenn man Lehre bewerten (lassen) will, muß man sich zunächst darüber klar werden, was „gute Lehre“ ist (vgl. Abschnitte zur Definition von Lehrqualität in Kapitel 2). Von der Festlegung der Lehrziele hängt ab, welche Methoden, Instrumente und letztlich, welche einzelnen Fragestellungen (Items) verwendet werden, mittels derer „gemessen“ werden soll, ob die Lehre den festgelegten Zielen entspricht.

Nachfolgend werden als Instrument nur standardisierte Fragebögen vorgestellt, auf die sich die meisten empirischen Untersuchungen beziehen. Diese sind im allgemeinen für Vorlesungen und Seminare mit einer größeren Teilnehmerzahl (über 30) das geeignete Instrument. Für kleinere Lehrveranstaltungen mit stark diskursivem Charakter oder überwiegend praktischen Anteilen sind oft andere Methoden und Instrumente besser geeignet (z.B. Gruppendiskussions- oder Rezensionsverfahren<sup>171</sup> (vgl. Lohnert/ Rolfes 1997).

Es gibt jedoch einige "Meß"-instrumente in Form von standardisierten Fragebögen zur Evaluation von Lehrveranstaltungen, die den verschiedenen Lehrzielen mit einer gewissen Variabilität Rechnung zu tragen versuchen. Rindermann (1996) fügte seinem Fragebogen einen Pool von Ergänzungsfragen (auch für Seminare/ Übungen) bei, den er zusammen mit dem Fragebogen-Hauptteil auf seine Eignung, Validität und Reliabilität hin überprüfte. Unter Validität wird hierbei verstanden, ob durch das Meßinstrument - also in diesem Falle die Fragebögen zur Qualität der Lehre - tatsächlich das gemessen wird, was gemessen werden soll (vgl. z.B. Schnell u. a. 1999, Diekmann 2002: 223).

Als reliabel wird eine Messung dann bezeichnet, wenn wiederholte Messungen am gleichen Objekt zu den gleichen Ergebnissen kommen. Andere Autoren sprechen auch von "Reproduzierbarkeit" der Ergebnisse. Die Frage nach der Reliabilität ist hierbei zwar eine notwendige, aber keine hinreichende Voraussetzung für Validität. Eine ausführliche Beschreibung der Methoden, mittels derer die Reliabilität und Validität eines Meßinstrumentes ermittelt werden kann, findet sich z.B. bei Diekmann (2002: 224). Speziell für Studentenbefragungen diskutiert dies ausführlich Rindermann (2001). Genau durch diese Methoden, die nach wie vor außerhalb der empirischen Sozialwissenschaften oft nicht sonderlich bekannt sind, unterscheiden sich jedoch die hier vorgestellten Instrumente von vielen anderen, die Fakultäten verschiedener Hochschulen meist ohne sozialwissenschaftliches Know-How erstellten. Vorbehalte bezüglich der Zuverlässigkeit von Studentenbefragungen rühren wohl auch daher.

---

<sup>171</sup> Bei sogenannten Rezensionen werden Lehrveranstaltungen schriftlich, jedoch mittels offener Fragestellungen ohne Antwortvorgaben bewertet. Hierzu eignen sich je nach Interessenschwerpunkt und Umfang der Fragestellungen anstelle eines Fragebogens auch handelsübliche Karteikarten im Format DIN A6 oder DIN A5. Fragestellungen könnten sein: Was hat Ihnen/ Dir an der Lehrveranstaltung besonders gefallen/ überhaupt nicht gefallen? Was hätten/st Sie/Du sich/Dir anders gewünscht? Die Karteikarten werden dann durch Studenten eingesammelt, durchmischt wieder ausgeteilt und mit den Teilnehmern besprochen, so daß die schriftlichen Äußerungen nicht zwangsläufig einer Person zuzurechnen sind (geeignet ab 5 bis zu maximal 30 Teilnehmern).

Im folgenden wird der Schwerpunkt des Abschnittes deshalb auf Instrumente gelegt, deren Reliabilität und Validität geprüft wurde und die an verschiedenen Hochschulen in Deutschland bereits Anwendung fanden. An erster Stelle sind hier Fragebögen zu nennen, die sich an das Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE) anlehnen (vgl. Rindermann 1996),<sup>172</sup> wobei der Fragebogen an einigen Hochschulen (z.B. an der TU Dresden) um wichtige Rahmenbedingungen von Lehrveranstaltungen erweitert wurde (vgl. Krempkow/ Heldt 2000b). Der Schwerpunkt der Fragestellungen dieses Instrumentes liegt auf der Bewertung didaktischer Aspekte der Lehrveranstaltung in Interaktion mit Motivation und Leistungsfähigkeit der Studierenden.

Studentenbefragungen können im Gegensatz zu Lehrveranstaltungsbewertungen aber auch das Ziel einer Bewertung von Studienbedingungen verfolgen. Da die Studienbedingungen hier aber als Bestandteil guter Lehre aufgefasst werden und Diskussion der Eignung von Studentenbefragungen zur Bewertung von Studienbedingungen und der Lehrveranstaltungsqualität sich relativ wenig voneinander unterscheiden (vgl. z.B. Multrus 1995, Hornbostel/ Daniel 1996), sollen Studentenbefragungen zu Studienbedingungen hier nicht separat diskutiert werden. Die Zuverlässigkeit und Vergleichbarkeit von Befragungen zu Studienbedingungen wurden bereits an anderer Stelle exemplarisch für die TU Dresden diskutiert (vgl. Krempkow 2003b).

#### **5.4.2 Die Datenqualität von Ergebnissen studentischer Bewertungen der Lehre - oder: Welche Argumente gegen Befragungen halten empirischer Prüfung stand?**

Entgegen immer noch weit verbreiteter Ansicht sind die in den vorgestellten Studien verwendeten Fragebögen dem Konzept nach also keineswegs nur "einfache" Erhebungen der Teilnehmerzufriedenheit. Vielmehr wird versucht, von den Studierenden auf mehreren

---

<sup>172</sup> Rindermann hatte ausgehend von der Tatsache, daß zu Beginn seiner Untersuchungen im deutschen Sprachraum noch keine allgemein verwendbaren und auf Testgütekriterien untersuchten Fragebögen zur Veranstaltungsbewertung zur Verfügung standen, den Schwerpunkt seiner Arbeiten auf die Untersuchung und Entwicklung solcher Fragebögen für einzelne Lehrveranstaltungen gelegt (Vgl. Rindermann 1994: 11). Der von Rindermann verwendete standardisierte Bogen sollte die Lehre des Dozenten, Veranstaltungscharakteristika und das studentische Verhalten valide und zuverlässig erfassen. Rindermann wandte zur Überprüfung seines Fragebogens ein flexibles, dimensionenorientiertes Konzept an. In seinem ersten Schritt versuchte er, möglichst breit möglichst viele Aspekte von Lehrqualität zu erfassen. Die Items für einen ersten Fragebogen wurden dann durch Befragung von Dozenten und Studierenden sowie die Durchsicht in- und ausländischer Fragebögen und hochschuldidaktischer Literatur ausgewählt. In mehreren weiteren Schritten wurden dann die Items überarbeitet und auf ihre Relevanz geprüft. Nach diesem Prozeß der Itemselektion wurde der überarbeitete Fragebogen anhand von drei zeitlich aufeinanderfolgenden Untersuchungsstichproben (Sommersemester 1992 bis Sommersemester 1994) auf seine Brauchbarkeit

Dimensionen eine umfassende und detaillierte Einschätzung von Lehrveranstaltungscharakteristika, Dozenten- und Studentenverhalten (und z.T. der Rahmenbedingungen von Lehrveranstaltungen) vornehmen zu lassen (vgl. Rindermann 2003). Hiermit wird bereits ein Aspekt von Validität angesprochen, nämlich die Inhaltsvalidität.

Da es wenig Sinn machen würde, Ausführungen über Reliabilität zu machen, ohne daß Validität gegeben ist (s.o.), soll zunächst die Validität diskutiert werden. Grundsätzlich ist bei der Bestimmung der Validität nach Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität zu unterscheiden.

*Inhaltsvalidität* meint, ob alle Aspekte des zu messenden Sachverhaltes erfaßt werden. Das bedeutet in diesem Fall, daß möglichst viele verschiedene Aspekte „guter Lehre“ erfaßt werden sollten. Nach Diekmann (2002: 224) ist Inhaltsvalidität gegeben, wenn eine Auswahl von Items die zu messende Eigenschaft in hohem Grade repräsentiert. Dies entspricht in etwa auch der am Anfang dieses Kapitels genannten Anforderung an Indikatorensysteme, daß diese den zu bewertenden Gegenstand möglichst vollständig erfassen sollen (vgl. Buschor 2002: 74).<sup>173</sup>

Anders ist es bei der *Kriteriumsvalidität*: Hiermit ist gemeint, ob ein mit dem gerade zur Messung verwendeten Kriterium erzieltetes Ergebnis mit einem mittels einer anderen Methode erzielteten Meßergebnis übereinstimmt bzw. ob es einen Zusammenhang zwischen den beiden Kriterien gibt. Dies wird von einigen Autoren auch als externe Validierung bezeichnet. Kriteriumsvalidität würde also in diesem Fall bedeuten, ob die mittels Studentenumfragen erzielten Ergebnisse mit anders gewonnenen Aussagen über die Lehrqualität, z.B. mit Befragung von Absolventen oder den Einschätzungen externer Gutachter hinreichend übereinstimmen. Nach Diekmann (2002: 224) erfolgt die Überprüfung v.a. mittels Korrelationskoeffizienten.<sup>174</sup>

Die dritte Möglichkeit zur Einschätzung der Validität ist die *Konstruktvalidität*. Nach Schnell u.a. (1999) liegt Konstruktvalidität dann vor, wenn aus dem Konstrukt empirisch überprüfbare Aussagen über Zusammenhänge dieses Konstruktes mit anderen theoretisch hergeleitet werden können und sich diese Konstrukte empirisch nachweisen lassen. Im Gegensatz zur Kriteriumsvalidität sind bei der Konstruktvalidität explizite theoretische

---

untersucht, wobei noch kleinere Änderungen und Präzisierungen erfolgten. Eine grundsätzliche Überarbeitung erfolgte mit HILVE II (vgl. Rindermann 2001).

<sup>173</sup> Wengleich diese Anforderung meistens eine Illusion sei, so Buschor, gelte sie doch als Meßlatte für Indikatorensysteme.

Annahmen über die Zusammenhänge zwischen theoretischen Dimensionen Bestandteil des Validierungsprozesses. Ähnlich formuliert dies Diekmann (2002: 224): Demnach meint Konstruktvalidität, daß das erfasste Konstrukt mit möglichst vielen anderen Variablen in theoretisch begründbarem Zusammenhang steht.

### **Inhaltsvalidität**

Die Überprüfung der Inhaltsvalidität von Studentenbefragungen kann z.B. mittels Befragung von Experten und mittels (offener) Befragungen der Beteiligten geschehen, um zu erfahren, was Studierende und Lehrende als Kriterien guter Lehre auffassen. Zusätzlich können standardisierte Befragungen zur Wichtigkeit solcher Kriterien und/ oder Analysen zur Prognosekraft von Kriterien z.B. für den Lehrerfolg erfolgen. Einen Überblick über solche Untersuchungsansätze gibt Rindermann (2001: 36). Eine noch aktuellere Zusammenfassung unter Einbeziehung psychodiagnostischer Studien bieten Strobel/ Westhoff (2003: 373ff.).

Rindermann diskutierte mehrere empirische Studien aus dem amerikanischen und dem deutschen Hochschulbereich, darunter auch eigene offene und geschlossene Erhebungen bei Studierenden und Lehrenden zur Wichtigkeit von Kriterien. Diese Studien zeigten, daß die in seinem und ähnlichen Instrumenten verwendeten Items alle wichtigen und auch zahlreiche von den Befragten z.T. als weniger wichtig eingeschätzte, aber nach dem zugrunde liegenden Konzept (vgl. Rindermann 2001: 63ff.) bedeutsame Aspekte von Lehrqualität (wie z.B. Rahmenbedingungen der Lehre) enthalten. Didaktische Kompetenzen und das Engagement der Lehrenden sowie die Strukturierung der Lehre zählen in mehreren Studien zu den am häufigsten genannten bzw. am wichtigsten eingeschätzten Merkmalen. Hierbei gibt es häufig auch relativ große Übereinstimmung zwischen Lehrenden und Studierenden. Daneben werden von den Studierenden aber auch soziale und personale Kompetenzen der Lehrenden als relevant erachtet. Die Lehrenden nannten häufiger auch die fachliche Kompetenz als wichtiges Merkmal (die die Studierenden weniger einschätzen können als Lehrende – vgl. Rindermann 2001: 48). Zu ähnlichen Ergebnissen gelangten auch Spiel u.a. (2002: 27ff.) für österreichische Universitätslehrende. Demnach sind v.a. Strukturierung und Klarheit sowie verständliche Darstellungen und Erklärungen von Bedeutung.

---

<sup>174</sup> Rindermann (2001: 118) begründet am Beispiel von Studentenbefragungen ausführlicher, warum Korrelationskoeffizienten anderen Maßen der Übereinstimmung vorzuziehen sind.

Einen inhaltsanalytischen Ansatz zur Auswertung offener Fragen verwendete Bos (1995: 61ff.), den er zudem mit einer „latent-class-analysis“ verband. Ergebnis seiner Befragung von 180 Studierenden mit der Fragestellung „Was fällt Ihnen zur akademischen Lehre ein?“ war, das „die Bereiche, die in den ‘Rankings’, insbesondere in der Spiegelstudie, auch im Zentrum dessen stehen, was Studierende mit akademischer Lehre verbinden“ (vgl. Bos 1995: 76). Zwar wurde diese Befragung im Sommersemester 1993 etwa im Zeitraum der Veröffentlichung des zweiten Spiegel-Rankings durchgeführt. Es darf jedoch als unwahrscheinlich gelten, daß so viele befragte Studierende dieses gelesen haben, daß die Auswertung hierdurch beeinflußt wäre. Zudem weisen Analysen zur Faktorenstruktur und der Skalenhomogenität anhand des Konstanzer Studierendensurveys (der die im Spiegel-Ranking verwendeten Fragen enthält) darauf hin, daß in etwa dieselben Dimensionen als relevant erachtet werden können (vgl. Multrus 2001; 1995: 18ff.).<sup>175</sup>

Auch Rindermann (2001: 50) zog weitere Analyseverfahren wie Pfadanalysen und Regressionsanalysen zur Identifikation relevanter Merkmale der Lehre heran, da es seiner Einschätzung nach mittels solcher Analysen grundsätzlich möglich sei, Bedingungsfaktoren für Lehr- und Lernerfolg als relevante Merkmale für Lehrqualität zu finden. Voraussetzung dafür ist aber die Varianz der entsprechenden Merkmale. So ist eine Bedingung für Lernerfolg, daß Lehrveranstaltungen nicht häufig ausfallen. Wenn dies aber generell nicht der Fall ist, kann dieses Merkmal mit den hier gewählten Analysemethoden nicht identifiziert werden (vgl. Rindermann 2001: 50). Dennoch erwiesen sich auch in den Ergebnissen solcher Analysen viele der zuvor als relevant genannten Merkmale als wichtige Prediktoren. Außerdem zog Rindermann als mögliche Kriterien ergänzend die international und in den letzten Jahren auch an vielen Hochschulen in Deutschland entwickelten Lehrevaluationsfragebögen heran. Dieses Verfahren hat den Vorteil, daß es die impliziten und expliziten Theorien der Testkonstrukteure widerspiegelt und (dort, wo die Fragebögen breite Anwendung finden) als „offizielles Instrument“ auch den „Geist einer Institution“. Da die vorhandenen Fragebögen aber von vornherein nur die als „messbar“ eingeschätzten Aspekte beinhalten, stellt dies „ein konservatives, wenig innovatives“ Verfahren dar, wie Rindermann (2001: 51) selbst einschätzt. Auch dieses Verfahren führte jedoch im Wesentlichen zu denselben wie den bereits von Rindermann genannten Kriterien. Für

---

<sup>175</sup> An dieser Stelle ist darauf hinzuweisen, daß die im Spiegel-Ranking verwendeten Fragestellungen größtenteils nicht von den Autoren der Spiegelstudie selbst entwickelt wurden, sondern zuvor bereits mehrfach im Konstanzer Studierendensurvey (z.B. 1992/93 und 1994/95) verwendet wurden (vgl. Bargel 1995: 281ff.).

Deutschland wurden neben Rindermanns Untersuchungen auch im Rahmen einer Lehrveranstaltung der Universität Karlsruhe Vergleiche und Analysen der Lehrevaluation speziell an wirtschaftswissenschaftlichen Fakultäten durchgeführt (wobei häufig hochschulweit einheitliche Instrumente verwendet wurden), die bezüglich zentraler Merkmale guter Lehre ebenfalls eine relativ hohe Übereinstimmung ergaben (vgl. Faist 2000: 21ff.). Zu ähnlichen Ergebnissen kamen anhand einer fächerübergreifenden Analyse Souvignier/ Gold (2002: 269ff.), die allerdings nur auf vollständig in einschlägigen Publikationen veröffentlichte Fragebögen eingingen.

Zusammenfassend kann zur Inhaltsvalidität festgestellt werden, daß die Häufigkeit der Nennungen und die Einschätzung der Wichtigkeit von Merkmalen guter Lehre zwar je nach befragtem Personenkreis und verwendeter Erhebungs- bzw. Analysemethodik differieren. Die zentralen Merkmale sind jedoch weitgehend dieselben. Am häufigsten werden didaktische Kompetenzen und das Engagement der Lehrenden sowie die Strukturierung der Lehre genannt. Zusätzlich sind jedoch entsprechend den gewählten theoretischen Ansätzen weitere Merkmale hernazuziehen. Rindermann (2001: 53) beschreibt Lehrkompetenz in Anlehnung an diese Resultate als dreifaches Fähigkeitsbündel, zu dem neben guter Strukturierung und didaktischer Methodevielfalt und -sicherheit unter verschiedenen Unterrichtsbedingungen auch soziale Kompetenzen zählen. Zum HILVE-Fragebogen (und daran angelehnte Fragebögen) kann dem entsprechend eingeschätzt werden, daß er alle wesentlichen Merkmale guter Lehre erfasst.

### **Kriteriumsvalidität**

Hiermit war gemeint, ob ein mit dem gerade zur Messung verwendeten Kriterium erzieltetes Ergebnis mit dem eines „Außenkriteriums“ übereinstimmt. Ein solches Außenkriterium können einerseits die Einschätzungen desselben Sachverhaltes durch Lehrende oder externe Gutachter bzw. Fremdgutachter sein. Andererseits können als Außenkriterium auch die Einschätzungen von Absolventen herangezogen werden. Häufig wird unterstellt, daß erst diese die Lehrqualität einschätzen könnten, da sie genügend Abstand zur Hochschule hätten und ihre Einschätzungen an den Anforderungen der Praxis bzw. ihren (ersten) Berufserfahrungen messen könnten.<sup>176</sup> So sind häufig gegen eine „Messung“ von Lehr-

---

<sup>176</sup> Hierzu ist außerdem anzumerken, daß Einschätzungen zur Urteilskompetenz nach dem jeweils betrachteten Gegenstand von Qualität differenziert werden müssen (vgl. Abschnitt zur Definition von Lehrqualität am Anfang dieser Arbeit). So macht es einen Unterschied, ob Studierende den Inhalt oder die Verständlichkeit

qualität mittels Studierendenbefragungen vorgebrachte Argumente eine vermutete nicht vorhandene Urteilskompetenz bzw. Reife der Studierenden. Diese Argumentation führte dazu, daß in mehreren Hochschulrankings Studierenden im Hauptstudium eine entsprechende Urteilskompetenz zugestanden wurde, nicht jedoch Studierenden im Grundstudium. Die unterstellte unterschiedliche Urteilskompetenz lässt sich auch als Einflußfaktor auf die Bewertung der Lehrqualität bzw. intervenierende Variable auffassen. Diese Untersuchung intervenierender Variablen stellt eine zusätzliche Möglichkeit zur Überprüfung der Eignung von Studierendenbefragungen dar, der im Zusammenhang mit der Vergleichbarkeit ausführlicher diskutiert werden soll.

Zunächst soll das Argument der Sachkenntnis nur im engeren Sinne der Kriteriumsvalidität diskutiert werden. Bei differenzierten Analysen dieser Argumentation anhand von Studien zur Brauchbarkeit studentischer Lehrbewertungen mittels HILVE stellte sich heraus, daß Studierende, gemessen an den Mittelwerten, fast alle wesentlichen Aspekte von Lehrveranstaltungen (z.B. Didaktik und Struktur) nicht grundsätzlich anders beurteilen als Dozenten oder Fremdgutachter (vgl. Rindermann 1994: 28, 1996: 128ff.). Lediglich die Referate und die Beteiligung der Studierenden in Seminaren wurden von den Fremdgutachtern deutlich anders (mehr als einen Skalenpunkt positiver) eingeschätzt als von Studierenden. Bei den übrigen Items betragen die Abweichungen kaum mehr als einen halben Skalenpunkt (7er-Rating-Skala: „trifft nicht zu ... trifft zu“).

Grundlegende Voraussetzung dabei ist die Verwendung gleicher Kriterien. So kann beispielsweise aus unterschiedlichen Ergebnissen von Hochschulrankings nach Professorenbefragungen und Studentenforschungen nicht auf mangelnde Validität geschlossen werden, wenn völlig unterschiedliche Fragestellungen verwendet wurden.

Unter dieser Voraussetzung zeigte sich, daß Studierende und Lehrende eine relativ hohe Übereinstimmung der Mittelwerte fast über alle Items hinweg erreichten. Wo es Unterschiede zwischen Dozenten und Studierenden gab, waren diese vor allem dort zu beobachten, wo die Dozenten sich selbst einschätzen mußten. So haben die Dozenten im Vergleich zu den Studierenden und „Fremdgutachtern“ ihre „Lehrkompetenz“ negativer eingeschätzt, ihr eigenes Engagement jedoch positiver. Allerdings war die Übereinstimmung zwischen Studierenden und „Fremdgutachtern“ höher als zwischen Studierenden und

---

bzw. die Strukturierung einer Lehrveranstaltung beurteilen. Während ihnen für erstere Einschätzung die relevanten Informationen fehlen, sind letztere bei ihnen besser und zeitnäher verfügbar als bei Absolventen. Da die Zeitnähe eine wesentliche Anforderung an Indikatoren darstellt, sollte dies nicht vernachlässigt werden.

Lehrenden. Rindermann (2001: 164) weist aber zu Recht darauf hin, daß Studentenforschungen mittels Lehrendenforschungen grundsätzlich schwer zu validieren seien, da Selbstbeurteilungen durch soziale Erwünschtheit verzerrt und außerdem nur von einer Person stammen und deshalb auch weniger reliabel sind. Daher sei ein Vergleich mit Ergebnissen von Lehrendenforschungen von vornherein v.a. dort sinnvoll, wo diese sich nicht selbst einschätzen müssten (z.B. bei der Einschätzung von Rahmenbedingungen von Lehrveranstaltungen oder von studentischen Aktivitäten, wie z.B. Referaten. Besser geeignet sind Urteile von Fremdgutachtern, die neben Hochschuldidaktikern oder Berufspraktikern auch von anderen Lehrenden stammen könnten. Häufig gibt es jedoch Vorbehalte von Lehrenden gegenüber der Teilnahme anderer Lehrender (v.a. von derselben Institution) an ihren Lehrveranstaltungen, da dies in Deutschland sehr unüblich ist und wohl als Kontrolle aufgefasst wird.

Eine Metaanalyse von Studien zur Kriteriumsvalidität studentischer Lehrbewertungen (vgl. Rindermann 2001: 166) ergab, daß die Urteile von Studierenden und Fremdgutachtern durchschnittlich mit  $r=.54$  korrelieren.<sup>177</sup> Da bei Korrelationskoeffizienten von  $r=.4 - .6$  von mittlerer Validität ausgegangen wird, kann dies als Beleg für eine zumindest mittlere bzw. durchschnittlich gute Validität (im Vergleich zu anderen Untersuchungen) gelten. Die Korrelation der Urteile von Studierenden und Selbsteinschätzungen von Lehrenden fällt allerdings mit  $r=.24$  deutlich niedriger aus. Die Übereinstimmung der Urteile von Fremdgutachtern und Selbsteinschätzungen von Lehrenden ist insgesamt mit  $r=.06$  nur sehr gering.

Zwar wurden von Rindermann (2001: 169) in seinen eigenen Untersuchungen etwas höhere Korrelationen gefunden (Studierende und Selbsteinschätzungen von Lehrenden:  $r=.28$ , Fremdgutachter und Selbsteinschätzungen von Lehrenden:  $r=.29$ ). Er spricht aber mit Blick auf die Ergebnisse der Metaanalyse dennoch eher zurückhaltend davon, daß Selbsteinschätzungen von Lehrenden „nur mit Einschränkung als ein Kriterium unter anderen zur Beurteilung der Validität von studentischen Evaluationen“ herangezogen werden können.

Bisher wurden an dieser Stelle für Deutschland nur Studien zur Lehrveranstaltungsevaluation diskutiert, wobei die Selbsteinschätzungsproblematik auftritt. Hornbostel (2001: 8ff.) untersuchte den Zusammenhang zwischen Ergebnissen von Studierenden- und Professorenbefragungen auch für Studienbedingungen, wo die Problematik der Selbsteinschätzungen je nach betrachtetem Aspekt weniger stark auftreten dürfte. Tatsächlich finden sich z.B. für

---

<sup>177</sup> Sie ist eine Erweiterung der Metaanalyse von Feldmann (1989), die durchschnittlich  $r=.53$  ergab.

den Index „Räume“, der für Studierende und Lehrenden gleichermaßen eine wichtige Rahmenbedingung der Lehre darstellt und wohl am wenigsten im Verdacht einer Beeinflussung durch Selbsteinschätzungsprobleme steht, je nach Fach Korrelationskoeffizienten zwischen  $r=.8$  und  $.55$ . Im Durchschnitt der untersuchten 8 Fächer beträgt die Korrelation in diesem Aspekt  $r=.69$ . Bei anderen Indizes, die wiederum stärker auch Selbsteinschätzungsaspekte enthalten, wie „Kontakt Stud.-Prof.“ und „Gesamturteil Lehre“, fallen die Korrelationen erwartungsgemäß deutlich geringer aus. Aber auch hier finden sich keine negativen Korrelationen, was Scheuch (1990)<sup>178</sup> im Zusammenhang mit dem ersten Spiegel-Ranking einmal unterstellte. Vielmehr liegen die Korrelationen alle im positiven Bereich und betragen im Mittel beim „Kontakt Stud.-Prof.“  $r=.46$  und beim „Gesamturteil Lehre“  $r=.38$ , wobei zu letzterem Aspekt die Fragestellung zwischen Studierenden und Professoren nur annähernd identisch war.

Zur Validierung von studentischen Lehrveranstaltungsbewertungen wären (neben der Funktion einzelner Absolventen oder kleinerer Teams als Fremdgutachter) grundsätzlich auch Befragungen von Absolventen(jahrgängen) geeignet. Zwar haben sie den Nachteil, daß ihre Erfahrungen mit der Lehre schon eine ganze Weile zurückliegen und sie nicht die aktuellen Lehrveranstaltungen kennen. Wenn es jedoch um die Einschätzungen von langfristigen Wirkungen der Lehre geht und darum, daß Absolventen durch den zeitlichen Abstand unabhängiger urteilen können, wirkt dieser nicht als Nachteil.

Leider sind bislang jedoch nur wenige Vergleiche der Ergebnisse von Studenten- und Absolventenbefragungen bekannt (vgl. Daniel 1996, Krempkow/ Pastohr 2003). Wahrscheinlich hängt dies damit zusammen, daß Absolventenbefragungen relativ aufwendig sind und für sinnvolle Vergleiche möglichst dieselben Jahrgänge während ihres Studiums und dann als Absolventen befragt werden müssten, was idealerweise ein Paneldesign und damit eine langfristige Untersuchungsplanung voraussetzt. In den wenigen vorliegenden Studien konnte hierfür bisher in keinem Fall ein Paneldesign angewendet werden.<sup>179</sup>

---

<sup>178</sup> Vgl. dazu die intensive Diskussion von Scheuch, Lamnek, Endruweit, Gräf und Neidhardt in „Soziologie“ 2/90 und 1/91.

<sup>179</sup> Im Potsdamer Modell der Lehrevaluation wurde zwar ein als innovativ eingeschätztes Paneldesign eingesetzt (vgl. Engel/ Pohlenz 2001: 136). Allerdings wurden Panelvergleiche bisher lediglich zwischen Studierenden im Grund- und Hauptstudium bei relativ geringen Rücklaufquoten (zwischen 29 und 38%) und Fallzahlen (220 für die gesamte Hochschule) sowie mit teilweise veränderten Fragestellungen durchgeführt und publiziert. Hierbei ergaben sich in den untersuchten Aspekten Korrelationen von  $r=.3$  bis  $.45$  (vgl. Engel/ Pohlenz 2001: 140). Bei zu späteren Zeitpunkten durchgeführten Exmatrikuliertenbefragungen war offenbar kein Panelvergleich zwischen Studierenden und Absolventen möglich, da „keiner der im Rahmen früherer Studien befragten Studierenden auch in die vorliegende Stichprobe aufgenommen wurde“ (vgl. Pohlenz 2002:

Auch in einer eigenen Untersuchung war es lediglich möglich, annähernd die entsprechenden Jahrgänge zu befragen, die inzwischen (überwiegend) Absolventen sein müßten (vgl. Krempkow/ Pastohr 2003: 84ff.). Hierbei war es nicht möglich zu kontrollieren, inwieweit tatsächlich dieselben Personen erreicht werden konnten, die sich Jahre zuvor als Studierende an Befragungen zur Studiensituation Beteiligten. Da bis zum Zeitpunkt dieser Analysen keine andere Möglichkeit zu solchen Untersuchungen bestand, wurde trotz dieser Problematik der Versuch unternommen, der Frage nach grundsätzlichen Unterschieden zwischen der Perspektive der Studenten und Absolventen nachzugehen.

Nach diesen ersten Untersuchungsergebnissen finden sich bei aller gebotenen Vorsicht in der Interpretation keine Belege für eine grundsätzlich geringe Übereinstimmung und damit für das Argument, die Studierenden bräuchten zur Einschätzung der Studienbedingungen erst eine gewisse Distanz zur Lehranstalt. Die Studierenden beurteilen die Situation von Lehre und Studium ihres Faches auch nach Verlassen der Hochschule als Absolventen und entsprechender Berufserfahrung nicht vollkommen anders. Zwar gibt es, wenn man die Mittelwerte betrachtet, die Tendenz zu positiveren Urteilen gegenüber den Ergebnissen der Befragung als Studierende. Bringt man die Einschätzungen jedoch in eine Rangfolge bzw. unterteilt diese in Ranggruppen, so gibt es kaum Unterschiede zwischen den am besten und den am schlechtesten beurteilten Aspekten. Die Beurteilungen der Aspekte korrelieren mit  $r=.96$  (gesamte Hochschule) bzw.  $.98$  (Wirtschaftswissenschaften)<sup>180</sup> hoch. Allerdings kann man die Ergebnisse noch nicht verallgemeinern. Hierzu sollten weitere Untersuchungen auch an anderen Hochschulen und möglichst mittels Paneldesign erfolgen. Sollten diese jedoch zu ähnlichen Ergebnissen kommen, wäre dies – gemessen am „Außenkriterium“ Absolventenperspektive – ein Beleg für eine gute Kriteriumsvalidität.

Rindermann (2001: 175ff.) diskutiert außerdem noch den Zusammenhang mit dem Lernerfolg (z.B. gemessen am durchschnittlichen Prüfungserfolg) als potentielle Möglichkeit zur Überprüfung der Kriteriumsvalidität. Zwar sind auch hier die Zusammenhänge einer Metaanalyse mit  $r=.52$  im Vergleich zur Studenten-Fremdgutachter-Übereinstimmung zufriedenstellend hoch. Im Vergleich zu Korrelationen von

---

6 sowie Pohlenz 2003). Es dürften allerdings auch erst relativ wenige Absolventen das Studium beendet haben, da zwischen erstem Befragungszeitpunkt (Wintersemester 1997/98) und Exmatrikuliertenbefragung (Wintersemester 2001/02) ein Zeitraum lag, der deutlich geringer als die durchschnittlichen Studiendauern an deutschen Hochschulen ist. Hinzu kommt, daß mehr als die Hälfte der befragten Studierenden in der Exmatrikuliertenstudie als Studienabbrecher oder Hochschulwechsler eingestuft wurden, so daß die Zahl der Absolventen unter den Befragten ohnehin relativ gering war (vgl. Pohlenz 2002: 9).

<sup>180</sup> Für weitere einzelne Fächer war eine Überprüfung aufgrund zu geringer Fallzahlen nicht möglich.

Intelligenztests und Schulnoten, wo bei Korrelationskoeffizienten von  $r=.5$  von den höchsten Übereinstimmungen der Psychodiagnostik überhaupt gesprochen wird, sind dies sogar sehr hohe Werte. Dennoch ist bei Verwendung des Lernerfolgs zur Überprüfung der Kriteriumsvalidität grundsätzlich Skepsis angebracht. Einerseits werden bei diesem Zusammenhang nur partiell identische Aspekte erfaßt (vgl. Rindermann 2001: 177). Andererseits wäre, da auch unabhängig von der Lehrqualität z.B. durch Selbststudium oder praktische Erfahrungen Prüfungserfolg erzielbar ist, eine monokausale Interpretation nicht adäquat. Hinzu kommt, daß die Ergebnisse von Prüfungen an Hochschulen nicht mit einheitlichen Leistungstests vergleichbar sind, sondern vielmehr einer sehr unterschiedlichen Notenvergabepraxis unterliegen (vgl. Abschnitt zu Abschlußnoten als Kriterium für Studienerfolg in diesem Kapitel).

### **Konstruktvalidität**

Die Konstruktvalidität findet sich z.B. bei Rindermann in seinen explizit ausformulierten verschiedenen Dimensionen von „guter Lehre“ wieder (vgl. Rindermann 1996). Diese Dimensionalität kann mittels Faktorenanalysen überprüft werden, da diese Faktoren Konstrukte darstellen (vgl. Rindermann 2001: 77ff.). Wichtig ist hierbei noch die Unterscheidung zwischen Konvergenz und Diskriminanz. Konvergenz liegt vor, wenn unterschiedliche Operationalisierungen eines Konstruktes einander sehr ähnlich und damit im Prinzip austauschbar sind. Diskriminanz liegt dann vor, wenn ein Instrument andere Sachverhalte erfaßt als andere Instrumente. Die Überprüfung der Konvergenz oder Diskriminanz kann in der Praxis mit Multitrait-Multimethod-Analysen erfolgen (vgl. Marsh 1982b; sowie Rindermann 1996). Hierbei werden in einer Matrix verschiedene Methoden einander gegenübergestellt (vgl. z.B. Diekmann 1995: 227). Dann wird überprüft, ob die Zusammenhänge zwischen den Messungen desselben Konstruktes (Konvergenz) stärker sind als die Zusammenhänge zwischen den Messungen verschiedener Konstrukte. Ist dies der Fall, kann man von Konstruktvalidität sprechen. Zur Feststellung der Konstruktvalidität mittels Multitrait-Multimethod-Analysen bedarf es neben Kompetenzen zur Anwendung verschiedener Methoden jedoch einer genügend großen Datenmenge und dementsprechend umfangreicher finanzieller Mittel, um Messungen mit verschiedenen Methoden durchführen zu können. Deshalb konnte sie bisher in bezug auf Lehrqualität nur in wenigen Fällen angewendet und noch keine verallgemeinerbaren Aussagen getroffen werden.

Günstiger sieht die Lage bezüglich Faktorenanalysen aus, die ebenfalls eine Möglichkeit zur Überprüfung der Konstruktvalidität darstellen und mit bereits vorhandenen Daten ohne zusätzliche Erhebungen durchgeführt werden können. Rindermann (2001: 77ff.) legte inzwischen eine Vielzahl von faktorenanalytischen Ergebnissen anhand mehrerer unterschiedlicher Stichproben unter Einbeziehung von Studierenden, Fremdgutachtern und Dozenten an mehreren Hochschulen vor. Hierbei stellte er fest, daß sich die Dozentenitems z.B. innerhalb der Lehreffektivität zwar je nach Stichprobe stärker differenzieren lassen, die zugrundeliegende studentische Urteilsdimension ist für diese Items aber identisch. Zusammenfassend formuliert er: „In allen Varianten der faktorenanalytischen Untersuchungen mittels des HILVE konnte jedoch eine Differenzierung zwischen den einzelnen Hauptkomponenten des multifaktoriellen Modells der Lehrqualität beobachtet werden. (...) Die in verschiedenen Studien (Heidelberger vs. Bonner Daten) gewonnenen Faktorenlstrukturen ähneln sich weitgehend. Die extrahierten Faktoren können damit als stichprobenunabhängig gelten.“ (vgl. Rindermann 2001: 107). Kritisch wird festgehalten, daß die gewählte theoretische Dimensionierung des HILVE feiner als die empirische ist. D.h. Studierende differenzieren zwar zwischen verschiedenen Aspekten der Lehre, innerhalb einzelner Komponenten findet jedoch nicht zwischen allen Skalen eine empirisch nachweisbare Unterscheidung statt (vgl. Rindermann 2001: 112).

Dies kann auch für die Adaption des HILVE an der TU Dresden und mehreren weiteren sächsischen Hochschulen bestätigt werden, wo sich trotz Abwandlung, Erweiterung und z.T. Verkürzung des Instrumentes zwar nicht dieselben Faktoren, aber eine ähnliche Grundstruktur der Dimensionen finden ließ (vgl. Krempkow/ Winter 2000: 5):

## **Faktorenstruktur studentischer Lehrbewertung an der TU Dresden im WS 1999/2000**

Faktor 1a: Aufbau und Struktur der Lehrveranstaltung

- Die Ziele der Lehrveranstaltung sind transparent.
- Der Aufbau der Lehrveranstaltung ist gut nachvollziehbar.
- Die Relevanz der Themen für das Studium wird nahegelegt.

Faktor 1b: Interessenförderung und Lehr-/Lernerfolg

- Mein Interesse an den Inhalten der Lehrveranstaltung wurde gestärkt.
- Ich wurde zum Mitdenken motiviert.
- Durch die Lehrveranstaltung habe ich viel gelernt.

Faktor 2: Didaktik des/der Lehrenden

- Der/die Lehrende kann komplizierte Sachverhalte verständlich machen.
- Die verwendeten Beispiele sind hilfreich.
- Der Vortragsstil des/der Lehrenden ist anregend.

Faktor 3: Beratungsqualität

- Der/die Lehrende steht für Rückfragen zur Verfügung.
- Ich bin mit der Beratung und Betreuung insgesamt zufrieden.

Faktor 4: Dozentenengagement in Bezug auf die Lehrveranstaltung

- Der/die Lehrende wirkt immer gut vorbereitet.
- Der/die Lehrende nimmt die Lehre wichtig.

Faktor 5: Anforderungen in der Lehrveranstaltung<sup>181</sup>

- Die Stoffmenge der Lehrveranstaltung
- Die Anforderungen der Lehrveranstaltung
- Das Tempo in der Lehrveranstaltung

Faktor 6: Raumsituation

- Die Raumgröße ist der Teilnehmerzahl angemessen.
- Die Akustik in diesem Raum ist günstig.
- Die Sichtbedingungen in diesem Raum sind günstig.

Faktor 7: Verfügbarkeit von Arbeitsplätzen und Literatur

- Bei Bedarf bekomme ich in der Bibliothek einen Arbeitsplatz.
- Bei Bedarf bekomme ich im PC-Pool einen Arbeitsplatz.
- Die in der Lehrveranstaltung notwendige Literatur ist in der Bibliothek verfügbar.

International läßt sich eine Vielzahl unterschiedlicher Faktoren und Skalen in verschiedenen Lehrinventaren unterscheiden. Integrative Ansätze würden dagegen nach Rindermanns Einschätzung eher selten verfolgt. Trotzdem lasse sich ein zentrales Bündel an Dimensionen herauskristallisieren: Überwiegend wird in den Instrumenten das Lehrverhalten der Dozenten beschrieben (Strukturierung, Darstellungen u.ä.), Sozialverhalten (Dozenten-Studenten-Beziehung, Klima), Anforderungshöhe und Formen des Lehrerfolgs (Lernen, Allgemeinbeurteilung). In angloamerikanischen Verfahren werden zudem häufig die Benotungen/ Prüfungen hinzugezogen (vgl. Rindermann 2001: 109).

---

<sup>181</sup> Bei diesem Faktor ist zu beachten, daß die Skalierung im Vergleich zur sonst üblichen (1=„trifft zu“ bis 5=„trifft nicht zu“) anders erfolgte: Hier bedeuten die Werte der Skala: 1=„Viel zu hoch“, über 3=„Optimal“ bis 5=„Viel zu niedrig“.

Da sich die Dimensionen des HILVE zumindest teilweise auch in internationalen Studien wiederfinden, kann davon ausgegangen werden, daß es zwar international keine Einigkeit über die Faktoren- bzw. Dimensionenstruktur gibt, zumindest ein zentrales Bündel an Dimensionen jedoch überwiegend bestätigt wird.

### **Reliabilität**

Die Vorbehalte bezüglich der Zuverlässigkeit bzw. Reliabilität studentischer Befragungen lassen sich wie folgt zusammenfassen: Das Urteil wird von den Studierenden nicht mit dem notwendigen Ernst gefällt, es entspricht eher der jeweiligen Laune oder dem Zufall, ist leicht beeinflussbar und deshalb also insgesamt nicht zuverlässig. Wenn man dies unterstellt und als Hypothese auffaßt, dann müßte das Urteil der Studierenden unabhängig von den jeweiligen Lehrenden und unabhängig von der jeweiligen Befragtengruppe großen Schwankungen unterliegen, bzw. sich nicht replizieren lassen. Diese Diskussion um die Zuverlässigkeit von studentischen Urteilen ist neben der Frage der Validität ebenfalls bedeutsam für die Praxis der Lehrbewertungen. Denn gerade, wenn es um eventuelle Konsequenzen (leistungsorientierte Gehaltszulagen oder Mittelzuweisungen, Lehrpreise) geht, in die Lehrbewertungen mit einfließen sollen, ist die Frage wichtig, inwieweit diese Schwankungen unterliegen, die nichts mit Veränderungen der Lehrsituation zu tun haben.

Antwort auf solche Fragen zur Zuverlässigkeit bzw. Reliabilität könnten z.B. Analysen zur Übereinstimmung der Urteile bei wiederholter Befragung derselben Teilnehmer geben. Einige Autoren wie z.B. Diekmann (2002: 217) oder Sachs (1992: 132) sprachen daher auch allgemein von „Reproduzierbarkeit“ der Ergebnisse. Bortz (2004: 11) spricht von „Genauigkeit bzw. Zuverlässigkeit der erhobenen Daten“, während weitere Autoren wie z.B. Noelle-Neumann (1998: 445) darunter die „Verlässlichkeit“ verstehen bzw. „das Vermögen (...), bei sonst gleichen Umständen zu den gleichen Ergebnissen zu kommen“. Bortz/ Döring (2002: 689) sehen Reliabilität als „Gütekriterium eines Tests oder Fragebogens, das die Genauigkeit angibt bzw., wie stark Meßwerte durch Störeinflüsse und Fehler belastet sind“. Bortz/ Döring (2002: 689) unterscheiden ebenso wie Diekmann (2002: 217) vier Techniken zur Überprüfung der Reliabilität: Paralleltestmethode, Test-Retest-Methode, die Methode der Testhalbierung und die interne Konsistenz bzw. die Itemkonsistenzanalyse.

Zur Messung der Reliabilität kann man zunächst zeitlich parallele Tests mittels gleicher oder möglichst ähnlicher Methoden (*Paralleltestmethode*) und Messungen zur zeitlichen

Stabilität von Meßergebnissen über einen bestimmten Zeitraum hinweg (*Test-Retest-Methode*) unterscheiden. Die Paralleltestmethode eignet sich jedoch von vornherein weniger für studentische Lehrbewertungen, da es sowohl in Feldstudien, als auch in quasiexperimentellen Untersuchungsanordnungen schwer möglich ist, parallel identische Situationen herzustellen, also zum gleichen Zeitpunkt identisches Lehrverhalten bei identischen Studierendengruppen zu messen. Die Retest-Reliabilität und dabei insbesondere die Stabilität über verschiedene Zeiträume hinweg ist jedoch gut überprüfbar. Darauf wird im Zusammenhang mit verschiedenen Untersuchungen zur Qualität der Lehre, dabei schwerpunktmäßig jedoch mit den Untersuchungen Rindermanns (1994, 1996 und 2001) noch genauer eingegangen.

Bei der *Methode der Testhalbierung* wird ein Meßinstrument mit multiplen Indikatoren in zwei Hälften aufgeteilt. Aus der Korrelation der beiden Testhälften wird dann die sogenannte „Split-half-Reliabilität“ errechnet. Die *Itemkonsistenzanalyse* sieht Diekmann (2002: 218) als eine Erweiterung dieses Verfahrens. Mittels Itemkonsistenzanalyse wird eine Art Durchschnittsreliabilität für sämtliche Halbierungsvarianten (einer Itemliste) errechnet, um zu einer eindeutigeren und genaueren Reliabilitätsschätzung zu gelangen. Da sowohl für die Methode der Testhalbierung (für die individuelle Urteilerübereinstimmung), als auch für die Itemkonsistenzanalyse aber immer eine Messung mit multiplen Indikatoren erforderlich ist, kann sie nicht in jedem Falle angewendet werden.<sup>182</sup> Die Test-Retest-Methode kann dagegen auch dann verwendet werden, wenn die Messung nur durch einen Indikator erfolgt (vgl. Diekmann 2002: 219). Zunächst sollen Ergebnisse der Methode der Testhalbierung zur Bestimmung der Zuverlässigkeit der Urteile über mehrere Personen, anschließend Ergebnisse der Test-Retest-Methode berichtet und diskutiert werden.

### **Methode der Testhalbierung**

Bei der Methode der Testhalbierung (auch Split-half-Reliabilität genannt) als Möglichkeit zur Überprüfung der Reliabilität ist zunächst nach zwei unterschiedlichen Zielstellungen zu unterscheiden (vgl. Rindermann 2001: 119): Zum ersten ist dies die individuelle Urteilerübereinstimmung oder Einzelraterreliabilität. Zum zweiten ist dies die Übereinstimmung des gemittelten Urteils über mehrere Personen (Interraterreliabilität).

---

<sup>182</sup> Ein Beispiel für die Analyse der Itemkonsistenz findet sich bei Strobel/ Westhoff (2003: 375). Dort wurden anhand eines „Rückmeldebogen(s) zu einem freien wissenschaftlichen Vortrag“ mit 26 bzw. 24 Items bei vier Teilstichproben relativ hohe Werte zwischen 0,88 und 0,94 ermittelt.

Wie bereits oben erwähnt, werden für die Ermittlung der Einzelraterreliabilität immer mehrere Indikatoren bzw. Items (für denselben Sachverhalt) benötigt, um Aussagen über die Zuverlässigkeit der Urteile *einer* Person z.B. bezüglich mehrerer Items treffen zu können.

Bei der Interraterreliabilität werden die Urteile *mehrerer* Personen zu einem Item herangezogen, um die Zuverlässigkeit des Urteils einer Person bezüglich anderer Personen zu prüfen. Die Interraterreliabilität ist damit auch ein „Maß für die Repräsentativität des Veranstaltungsmittels für die Beurteilerpopulation“ (vgl. Rindermann 2001: 119).

Da die Berechnungsmodi für Einzel- und Interraterreliabilität einander ähneln und erstere nicht in allen Fällen möglich bzw. relevant ist,<sup>183</sup> sollen nachfolgend v.a. Untersuchungen der Interraterreliabilität vorgestellt werden.

Die Interraterreliabilitäten für Kursmittel bei Lehrveranstaltungsbewertungen durch Studierende liegen nach Rindermann (2001: 126, 2003: 238) generell über  $r=.80$  bis  $.90$  für Kurse mit 10 bis 20 oder mehr Teilnehmern. Ähnlich zeigt Rindermann (2001: 131) dies auch für die einzelnen Dimensionen des HILVE I und II. Bei einer geringeren Anzahl von Urteilern (z.B.  $N=2$ ) sind die Übereinstimmungen allerdings deutlich geringer ( $r=.18$  bis  $r=.34$ ), wobei die Übereinstimmung bei Betrachtung von Dimensionen anstelle von Items (aufgrund der Aggregation) höher ausfällt. Zieht man zum Vergleich die Urteilerübereinstimmungen von externen Gutachtern in Lehrveranstaltungen ( $r=.29$  bis  $r=.50$ ) oder Gutachtern für Manuskripte bei wissenschaftlichen Zeitschriften heran, fallen diese bei gleicher Anzahl von Urteilern aber ähnlich gering ( $r=.18$  bis  $r=.34$ ) aus (vgl. Rindermann 2001: 127). An anderer Stelle weist Rindermann (2001: 204) darauf hin, daß selbst zwischen hochaggregierten Leistungsdaten wie Abitur und (Hochschul-)Zwischenprüfungsnoten nur Korrelationen um  $r=.39$  bestehen.<sup>184</sup> In der Literatur wird häufig ein Wert von mindestens  $r=.80$  gefordert. In

---

<sup>183</sup> An dieser Stelle soll der Aspekt der Zuverlässigkeit der studentischen Urteile und nicht der Aspekt der Zuverlässigkeit von Items bzw. Skalen (vgl. Rindermann 2001: 136ff.) im Vordergrund stehen. Zudem wird davon ausgegangen, daß für Leistungsbewertungen und Leistungsanreize ohnehin v.a. bereits auf ihre Zuverlässigkeit geprüfte Items diskutiert werden (ungeprüfte dürften aufgrund kritischer Diskussionen von vornherein nicht zum Einsatz gelangen) und bei Neuentwicklungen von Items Experten zu Rate gezogen werden.

<sup>184</sup> Dieser Befund kann mit einer kleinen Zusatzauswertung der Dresdner Absolventenstudien (vgl. z.B. Krempkow/ Popp 2003b) gestützt werden. Hier korrelierten Abitur und (Hochschul-)Zwischenprüfungsnoten über alle Fächer hinweg gesehen mit  $r=.40$  (wobei allerdings bei einer nach Fächern getrennten Auswertung größere Unterschiede zu beobachten sind). Betrachtet man zusätzlich zu Rindermanns Ansatz den Zusammenhang zwischen (Hochschul-)Zwischenprüfungsnoten und (Hochschul-)Abschlußprüfungsnoten, der nach theoretischen Überlegungen (größere Ähnlichkeit der Prüfungsinhalte, teilweise oder gänzlich dieselben Prüfer) eine viel höhere Übereinstimmung aufweisen sollte, so ergibt sich insgesamt (bei ebenfalls größeren fachspezifischen Unterschieden) eine Korrelation von  $r=.62$ . Auch dieser Wert liegt deutlich unter den Interraterreliabilitäten für Kursmittel bei Lehrveranstaltungsbewertungen und unter den häufig in der Literatur geforderten Werten von  $r>.8$  (bzw.  $.7$ ). Zwischen Abitur und (Hochschul-)Abschlußprüfungsnoten besteht ein Zusammenhang von  $r=.22$ . An anderer Stelle in diesem Kapitel wurde im Zusammenhang mit der Diskussion

einigen Fällen begnügt man sich jedoch auch mit einem Wert von  $r=.70$  (vgl. z.B. Brosius 2002: 766).

Die Interraterreliabilitäten der in die Untersuchungen einbezogenen Instrumente kann demzufolge ab einer Anzahl von mindestens zehn Urteilern als hoch eingeschätzt werden. Rindermann (2001: 134) kommt unter diesen Voraussetzungen zu dem Schluß: „In allen Lehrverhaltens- und Lehreffektivitätsskalen ließen sich gute Interraterreliabilitäten feststellen.“ Da die Urteile eines einzelnen oder von sehr wenigen (unter 10) Studierenden (ähnlich wie auch bei einzelnen Gutachtern für wissenschaftliche Zeitschriften) stärker divergieren, sollten Einschätzungen von *einzelnen* Studenten nicht zur Bestimmung der Lehrqualität herangezogen werden.

### **Retestreliaibilität bei Meßwiederholung nach ca. 2 Monaten und über mehrere Semester**

Rindermann (1996: 36, ausführlicher 156 ff.) untersuchte die Stabilität/ Retestreliaibilität von Urteilen über dieselbe Veranstaltung zu verschiedenen „Meßzeitpunkten“ und stellte große Übereinstimmungen der Veranstaltungsmittelwerte fest.

Befragt wurden die Teilnehmer jeweils zur Mitte und zum Ende des Semesters in denselben Veranstaltungen. Hierbei gelangten nur diejenigen Befunde in die Auswertung, die zu beiden Meßzeitpunkten auf der Basis von mindestens 10 Teilnehmern erfolgten. Diese waren aber nicht immer identisch, da einige Veranstaltungen im Laufe des Semesters Teilnehmer verloren, andere hinzugewannen.

Jedoch konnten bei einem Vergleich der Urteile, bei denen die Teilnehmer nicht identisch waren, mit denjenigen, bei denen sie identisch waren, kaum Unterschiede festgestellt werden.

Dennoch könnte die hohe Stabilität der Urteile innerhalb des Semesters theoretisch auch darauf zurückzuführen sein, daß die Kurswahlentscheidung der Studierenden bereits zu Beginn des Semesters, spätestens jedoch nach bis zu drei Lehrveranstaltungsbesuchen fällt, so daß die Unzufriedenen gar nicht erst in die Untersuchung gelangt sind, weil sie die Lehrveranstaltungen nicht weiter besucht haben. Wöllte man dies untersuchen, stieße man aber auf das Problem, daß sich die Studierenden nach einigen wenigen Veranstaltungs-

---

von (Abschluß-)Noten als objektive Indikatoren bereits auf die großen Unterschiede in der Notenvergabepraxis zwischen verschiedenen Hochschulen hingewiesen, die letztgenannten geringen Zusammenhang zumindest teilweise erklären könnten.

besuchen noch nicht so kompetent über die Veranstaltung äußern können. Die Befragung zu Semesteranfang wäre nur sinnvoll, wenn man die Erwartungen der Studierenden erfassen will (vgl. z.B. Kromrey 1993). Auch ist kein Fall bekannt, daß Einschätzungen über die Qualität von Lehrveranstaltungen lediglich auf Basis von Befragungen zu Veranstaltungsbeginn erfolgten. So bleibt dieser Einwand rein hypothetisch.

Nach einer Übersicht über Retestrelisibilitäten innerhalb eines Semesters bei 4 Studien (vgl. Rindermann 2001: 137) läßt sich festhalten, daß Studierende auch bei zweimaliger Untersuchung innerhalb eines Semesters (Mitte und Ende) weitestgehend wieder zu den gleichen Urteilen kommen.<sup>185</sup> Die Korrelationskoeffizienten lagen bei drei der Studien zwischen  $r=.71$  und  $.92$ , bei einer um  $r=.50$  (HILVE I). Unterschiede gab es v.a. bei Veranstaltungen, die zum ersten Meßzeitpunkt relativ schlecht beurteilt wurden, und die nach Feedbackbesprechung und einigen Änderungen im Dozentenlehrstil besser beurteilt wurden (ausführlicher vgl. Rindermann 1996: 159). Aber auch diese einzelnen Abweichungen sprechen nicht unbedingt gegen die Stabilität studentischer Urteile, da ihre Richtung plausibel und schließlich als Ziel der Evaluationsmaßnahme auch erwünscht war. Offen bleibt bei der Untersuchung zu verschiedenen Meßzeitpunkten innerhalb eines Semesters jedoch, inwieweit Studierende verschiedener Jahrgänge dieselben Dozenten anders beurteilen. Deshalb soll dies ebenfalls noch untersucht werden.

Bis vor einiger Zeit ließ sich die Stabilität studentischer Urteile über mehrere Semester für bundesweite Untersuchungen nur anhand sehr weniger verfügbarer Daten ermitteln. So wiesen beispielsweise Hornbostel & Daniel (1996) darauf hin, daß diejenigen Items, die sie aus ihrer 1989 für den „Spiegel“ durchgeführten Untersuchung übernommen hatten, in ihrer Bewertung relativ konstant geblieben waren. Jedoch sind die Untersuchungen des „Spiegel“ 1989 und 1993 wegen Frageumformulierungen und des großen Abstandes nur bedingt vergleichbar. Aber auch Bargel u.a. (1996, 2003) berichteten von ihren regelmäßig durchgeführten Studierendenbefragungen ähnliches. Allerdings ist bei Bargel einzuschränken, daß diese Untersuchungen sich fachbereichsübergreifend auf die Ebene der Studienbedingungen insgesamt bezogen, über die Stabilität der Urteile zur Qualität der Lehre an einzelnen Fächern und Hochschulen läßt sich deshalb kaum etwas ableiten, sondern lediglich über die Einschätzung der Studienbedingungen in Deutschland insgesamt.

---

<sup>185</sup> Rindermann bezog insgesamt 16 Studien in seine Übersicht ein. Jedoch wurde bei den meisten Studien die Retestrelisibilität innerhalb eines Semesters nicht untersucht.

Untersuchungen, die Vergleiche von Lehrveranstaltungsbeurteilungen und Aussagen zur Stabilität über mehrere Semester hinweg zulassen, sind die seit 1994 (vgl. detaillierter Rindermann 1996 und 2001) im Zusammenhang mit dem HILVE vorgelegten. Rindermann nahm seine Untersuchungen über mehrere Semester hinweg für jeweils die gleichen Lehrenden vor und analysierte, wie stabil die studentischen Urteile über den Zeitraum von ein oder anderthalb Jahren hinweg seien. Trotz einiger Einschränkungen (z.B. beteiligten sich nicht alle Dozenten mehrfach) zeigen die Ergebnisse der Untersuchung über die Semester hinweg eine relativ große Übereinstimmung.

Ein Maß um festzustellen, inwieweit die Bewertungen zusammenhängen, ist die Korrelation zwischen den Kursmittelwerten in den verschiedenen Semestern (vgl. Rindermann 2001: 138). Relativ stark hängen hier die Bewertungen bei den Aspekten Struktur (.78) Lehrkompetenz ( $r=.77$ ), Dozentenengagement (.79) und bei der Bewertung der Dimension „Thema“ (.79) zusammen. D.h., wenn die Studierenden eines Jahrgangs die Lehrtätigkeit eines Dozenten als „strukturiert“ „kompetent“ und „engagiert“ einschätzen, werden es mit hoher Wahrscheinlichkeit auch die Studierenden des darauffolgenden Jahrganges in der gleichen Relation tun. Ähnlich ist es bei der Einschätzung zum „Thema“. Die hohe Korrelation bei „Thema“ kann aber auch darauf zurückgeführt werden, daß viele Dozenten zu den jeweiligen Meßzeitpunkten Veranstaltungen zu den gleichen bzw. ähnlichen Themen durchführten.<sup>186</sup> Im arithmetischen Mittel über die verschiedenen Dimensionen des HILVE I hinweg lag die Retestreliaibilität über mehrere Semester bei .70 und kann damit unter Berücksichtigung des Umstandes, daß hier viele andere Einflußfaktoren eine Rolle spielen dürften, als befriedigend gelten.

Wenn Dimensionen oder Skalen hoch korrelieren, kann dies aber auch heißen, daß es („Parallel“-) Verschiebungen im Anspruchsniveau gegeben haben könnte. Rindermann spricht hier von evtl. möglichen „kollektiven Mittelwertverschiebungen“ (Vgl. Rindermann 1996: 193 ff.) Deshalb müsse zur Beurteilung der Übereinstimmung auch das Mittelwertprofil herangezogen werden. Dieses zeigte aber ebenfalls große Übereinstimmungen. Kleinere Unterschiede der Mittelwerte konnten lediglich hinsichtlich Fleiß, Beteiligung und Diskussion beobachtet werden.<sup>187</sup> Dies sind jedoch typische Studenten-Selbsteinschätzungs-

---

<sup>186</sup> Einen starken Zusammenhang gibt es auch jeweils kreuzweise zwischen Lehrkompetenz und Dozentenengagement (.62 bzw. .76 zu den verschiedene Meßzeitpunkten), nicht jedoch zwischen Lehrkompetenz und Thema sowie Dozentenengagement und Thema (alle  $r < .40$  – vgl. Rindermann 1996: 193 ff.)

<sup>187</sup> Auch bei Verarbeitung gab es Unterschiede, hier war jedoch die Streuung relativ hoch. Der einzig signifikante Mittelwertunterschied fand sich bei Überforderung. Hier wurde jedoch die Fragestellung zur

Dimensionen, die bei den anderen Vergleichen ebenfalls stärker schwankten. Bei der Untersuchung Rindermanns kann man also von Stabilität studentischer Urteile über mehrere Semester hinweg ausgehen, die lediglich bei studentischen Selbsteinschätzungen eingeschränkt ist.

Die Ergebnisse werden auch von anderen Untersuchungen gestützt (Übersicht vgl. Rindermann 2001: 137). Beispielsweise Marsh und Hocevar (1991) kamen bei ihren Untersuchungen an einem ähnlichen, aber englischsprachigen Inventar anhand von 195 Dozenten aus 31 Studienfächern ebenfalls zu dozentenbezogenen Stabilitäten von  $r = .76$ . Sie hatten dabei sowohl gleiche als auch verschiedene Veranstaltungsthemen einbezogen (vgl. ausführlicher Rindermann 1996). Bei 5 weiteren Studien, die Rindermann (2001) in seine Übersicht einbezog, lagen die Retestrelisabilitäten (über mehrere Semester gesehen) für die dozentenbezogenen Skalen zwischen  $.59$  und  $.76$ . Bei einer Studie, die allerdings retrospektive Einschätzungen beinhaltetete, lag der Wert sogar bei  $.83$ . Berechnet man das Mittel über alle genannten Studien, so liegt dieses bei  $.69$ . Über verschiedene Instrumente hinweg gesehen, kann die Retestrelisabilität damit ebenfalls als befriedigend gelten.

Da für Deutschland bislang außer Rindermanns Analysen kaum Untersuchungen zu Retestrelisabilitäten über mehrere Semester bekannt sind, sollten – um die Ergebnisse auch für Deutschland generalisieren zu können – hier noch weitere Untersuchungen an verschiedenen Hochschulen und Fachrichtungen vorgenommen werden.<sup>188</sup> Aber auch wenn für weitere eingesetzte Instrumente in Deutschland insgesamt eine ausreichende Retestrelisabilität belegt werden kann, scheint jedoch weiterhin noch etwas Vorsicht geboten, wenn schwerwiegendere Schlußfolgerungen aus der Verwendung der Ergebnisse von Studentenforschungen in Indikatorensystemen erwachsen sollen. Hierfür sollten mindestens zwei oder drei Lehrbewertungen abgewartet werden, bevor auf der Grundlage solcher Ergebnisse irgendwelche Entscheidungen getroffen werden, da es bei einzelnen Lehrveranstaltungen immer besondere Umstände geben kann, die ggf. zu „verzerrten“ Ergebnissen führen könnten (siehe auch Abschnitt zur Vergleichbarkeit der Ergebnisse von Studentenforschungen).

---

leichteren Bearbeitung des Fragebogens stärker verändert, wodurch sowohl der Unterschied als auch die Richtung des Unterschieds als Methodeneffekt plausibel begründet wurden (Vgl Rindermann 1996: 196).

<sup>188</sup> Grundsätzlich möglich wäre dies z.B. auch anhand der Ergebnisse der Lehrveranstaltungsbewertungen an der TU Dresden anhand eines Vergleiches von drei aufeinanderfolgenden Semestern vom Sommersemester 1999 bis Sommersemester 2000 (vgl. Krempkow/ Winter 2000). Mangels personeller Ressourcen war hier leider nur eine überblicksartige Auswertung in Form von Mittelwertprofilen ohne detailliertere Analysen möglich. Die Mittelwertprofile derselben Fächer/ Fakultäten stimmen allerdings weitgehend überein.

### **5.4.3 Vergleichbarkeit der Ergebnisse von Studentenerhebungen**

Auch wenn die studentischen Urteile nach den vorliegenden empirischen Untersuchungen sehr viel zuverlässiger sind als bisher angenommen: Den Ansprüchen an eine Vergleichbarkeit studentischer Urteile zumindest über die gleichen Fachbereiche verschiedener Hochschulen genügen die studentischen Urteile jedoch nicht ohne weiteres. Hierzu müßten sie nach Marsh (1982b) bzw. Rindermann (1996) frei sein von "externen" Einflußfaktoren, die ihrerseits nicht mit der Lehr- und Studienqualität zusammenhängen. Zwar konnte anhand der Studie von Hage (1996b) für einen repräsentativen Querschnitt von Hochschulen<sup>189</sup> der Bundesrepublik gezeigt werden, daß weder unterschiedliche Sachkenntnis und Reife der Studierenden, noch das Geschlecht der Studierenden die Lehrbewertung in inakzeptabler Weise verzerrt.

### **Sachkenntnis, Reife und Geschlecht der Studierenden als potentielle Einflußfaktoren**

Eine Möglichkeit, die Urteilskompetenz der Studierenden zu prüfen, ist die Analyse, ob Studierende mit höherer Sachkenntnis z.B. aufgrund ihrer Hilfskrafttätigkeit, Studienintensität, eines Hochschulwechsels oder höheren Semesterzahl die Lehre anders bewerten als solche, von denen man eine geringere Sachkenntnis erwartet. Jedoch weder die studentischen Hilfskräfte, noch die Studierenden mit einer höheren Anzahl an Semesterwochenstunden, und auch nicht diejenigen, die zuvor schon eine andere Hochschule besucht hatten, beurteilten die Lehr- und Studiensituation insgesamt nennenswert anders als ihre Kommilitonen (vgl. Hage 1996b). Einzig die Anzahl der Fachsemester hatte einen geringen Einfluß. Aber hier ist es ebenfalls nicht wie ursprünglich vermutet, daß Studierende mit "höherem Einsichtsvermögen" die Lehre besser bewerten, sondern umgekehrt: Höhere Fachsemester urteilen im allgemeinen kritischer als niedrigere. Die Unterteilung in Grund- oder Hauptstudium hat dagegen keine bedeutsamen Unterschiede erbracht. Die wesentliche Änderung der studentischen Perspektive auf das Studium erfolgt bereits in den ersten beiden Semestern (vgl. Hage 1996b). Auch das Alter als

---

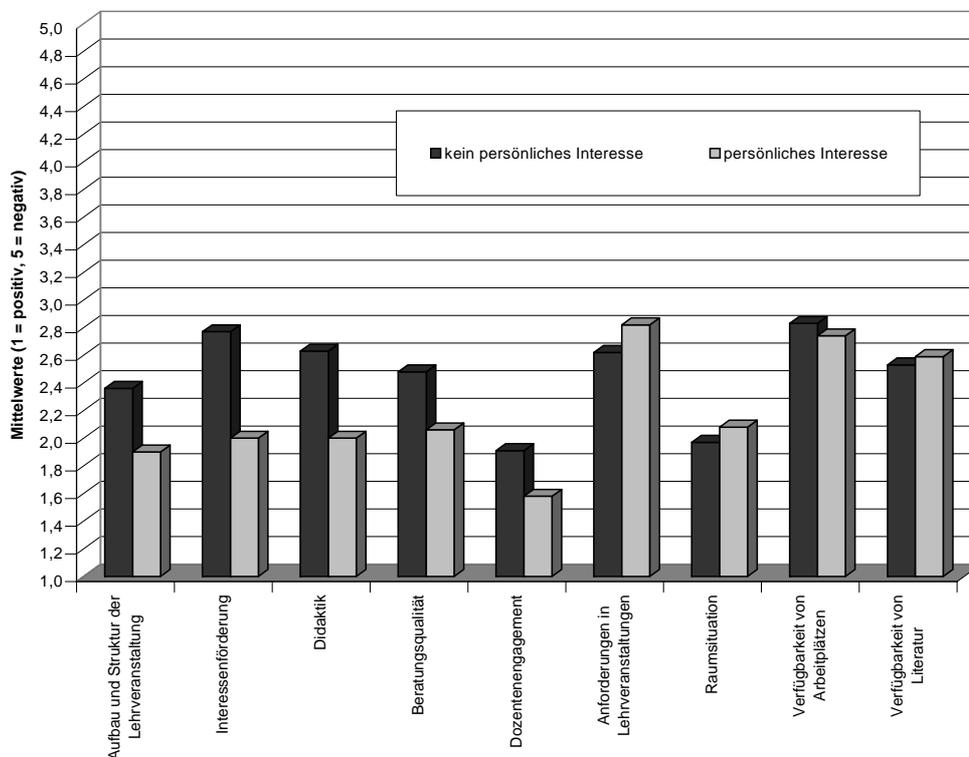
<sup>189</sup> Von den Konstanzer Sozialforschern um Bargel wurde versucht, ein möglichst umfassendes Bild der Situation an den repräsentativ ausgewählten Fachbereichen zu zeichnen. Die von Bargel verwendeten Indikatoren wurden in der Nachfolgeuntersuchung von Hage (1996b) auf ihre Validität überprüft. Zusammengefaßt wurden dabei drei Elemente als wesentlich für die Beschreibung der Studienqualität angesehen: Aufbau und Gliederung des Studienganges, Art und Durchführung der Lehrveranstaltungen sowie die inhaltliche Ausrichtung des Lehrangebots, Beratung und Betreuung der Studierenden (Vgl. Bargel, u.a. 1996: 111-126).

einzigster verfügbarer Indikator für Reife hatte keinen nennenswerten Einfluß. Hinzuzufügen ist, daß sich die Studierenden im Gegensatz zu den Befunden früherer weniger differenzierter Studien nicht generell vor allem vom Unterhaltungswert beeinflussen lassen. Dies ist nur dann der Fall, wenn sie in keiner Weise motiviert sind, der Lehrveranstaltung zu folgen und keinen Nutzen in der Lehrveranstaltung erkennen können (vgl. Marsh 1982a, 1997). Aufgrund eines internationalen Überblicks über Studien zu diesen potentiellen Einflussfaktoren kommt auch Rindermann (2001: 183) zu dem Schluß: „Generell können über alle Studien und Reviews hinweg studentische und dozentenbezogene biographische Variablen (wie Alter, Geschlecht, Semesterzahl) als Verzerrungsvariablen ausgeschlossen werden. Auch Persönlichkeitsmerkmale (...) korrelieren nur schwach mit der Veranstaltungsbeurteilung.“

### **Studieninteresse, Noten, Leistungsbereitschaft und Kursgröße als Einflußfaktoren**

Für weitere vermutete Einflußfaktoren, wie das Studieninteresse, die Noten/ Notenerwartung, die Leistungsbereitschaft und die Kursgröße wurden unterschiedlich große Zusammenhänge mit der studentischen Lehrbewertung festgestellt. Der stärkste Zusammenhang wurde bei der Untersuchung des Studieninteresses gefunden, und zwar sowohl auf Fachbereichsebene, als auch auf Ebene einzelner Lehrveranstaltungen. Auf Fachbereichs- und Hochschulebene wurde dies für Einschätzungen zur Qualität der Lehre bundesweit (vgl. Hage, 1996b) sowie an einzelnen Hochschulen untersucht (vgl. Krempkow/ Heldt 2000):

## Besuchsgrund „persönliches Interesse“ als Einflußfaktor auf Lehrbewertungen<sup>190</sup>



Quelle: Krempkow/ Winter 2000: 32

Auch für Einschätzungen zu Studienbedingungen wurde dies exemplarisch anhand eines Faches zu mehreren Zeitpunkten untersucht. Hierbei ergab sich zum ersten Zeitpunkt ein Einfluß von  $r=0,35$ , der sich aber für den zweiten Untersuchungszeitpunkt nicht bestätigen ließ (vgl. Krempkow 2003b: 267).

<sup>190</sup> Die Auswertung bezieht sich hier auf eine Analyse der im Sommersemester 2000 an der TU Dresden flächendeckend durchgeführten Lehrveranstaltungsbewertungen auf der Basis von insgesamt 19544 befragten Studierenden (vgl. Krempkow/ Winter 2000). Aufgrund dieser hohen Fallzahl ist gewährleistet, daß auch in den kleineren der analysierten Subgruppen zuverlässige Aussagen erfolgen. Das Signifikanzniveau liegt in allen Fällen unter einer Fehlerwahrscheinlichkeit von 0,05. Gleiches gilt für die hier nicht dargestellten Analysen der Ergebnisse der an der TU Dresden durchgeführten Lehrveranstaltungsbewertungen im Wintersemester 1999/ 2000 (23000 befragte Studierende) und im Sommersemester 1999 (17000 Studierende), für die bereits ähnliche Ergebnisse berichtet wurden.

Die Auswahl der Lehrveranstaltungen oblag laut Beschluß des akademischen Senates der TU Dresden den je zur Hälfte aus Studierenden und Lehrenden zusammengesetzten Studienkommissionen. Dabei war sicherzustellen, daß alle Lehrveranstaltungstypen angemessen berücksichtigt werden. Die Auswahl sowie die ihr zugrundeliegenden Auswahlkriterien waren fakultätsintern bekanntzugeben und im Jahresbericht der Studienkommission und im Lehrbericht der Fakultät darzustellen (vgl. Krempkow/ Heldt 2000a).

Auf Lehrveranstaltungsebene wurde der Einfluß des Interesses als Besuchsgrund der Lehrveranstaltung als Maß für das Studieninteresse untersucht (vgl. z.B. Kromrey 1996: 4ff.<sup>191</sup>, Rindermann 2001: 179ff.). In diesen Fällen wurden Einflüsse von  $r=.19$  bis  $.54$  festgestellt. Jedoch hegen einige Autoren Zweifel daran, ob das Studieninteresse als „echte“ Biasvariable zu bezeichnen wäre (vgl. Spiel/ Gössler 2000, Daniel 1998). Andere Autoren führen zudem aus, daß die Wirkungsrichtung wechselseitig sei, da z.B. gute Lehre das Interesse am Thema fördert (vgl. Rindermann 2001: 183). Auch Schwarzkopf (2001) führt hierzu aus, daß einerseits durch eine interessante Gestaltung auch Interesse geweckt werden kann und andererseits durch eine adäquate Information der Studienanfänger und - Interessenten „falsche“ Erwartungen bereits im Ansatz vermieden werden könnten. Das Interesse der Studierenden sollte aber in jedem Fall erfasst und bei stärkerem Einfluß in die Ergebnisinterpretation einbezogen werden.

Der am häufigsten untersuchte Einflussfaktor ist wohl der Einfluß der Noten bzw. der Notenerwartung auf die Ergebnisse von Studentenforschungen zur Lehrqualität. Um eine „Verzerrung“ und Beeinträchtigung der Vergleichbarkeit festzustellen, muß dies jedoch über mehrere Lehrveranstaltungen eines Faches hinweg betrachtet werden. Sowohl ältere Untersuchungen (Übersicht vgl. Krempkow 1997) als auch inzwischen erfolgte neuere Untersuchungen hierzu konnten aber entweder keine (Daniel 1998, Schwarzkopf 2001) oder nur relativ kleine Zusammenhänge feststellen (Krempkow/ Heldt 2000, Rindermann 2001, Krempkow 2003b: 267). Außerdem wird darauf hingewiesen, daß sich Zusammenhänge mit

---

<sup>191</sup> Kromrey fand bei der Untersuchung der Daten zur Lehrbewertung an der Ruhr-Universität Bochum einen deutlichen Zusammenhang: Studierende, die angaben, die Vorlesung „aus persönlichem Interesse“ zu besuchen, beurteilen die Lehre positiver und umgekehrt. Um sich nicht nur auf einen einzelnen Indikator zu verlassen, bildete Kromrey aus den Gründen des Veranstaltungsbesuchs (aus rein persönlichem Interesse, zur Auffrischung/ Wiederholung von Kenntnissen, Wahlveranstaltung usw.) einen Index zur Beschreibung der vorherrschenden Motivation der Hörschaft. Mit diesem Index untersuchte er die Unterschiede zwischen Lehrveranstaltungen, an denen vorwiegend extrinsisch Motivierte (mehr als 80% nur-Pflicht-Hörer); und vorwiegend intrinsisch Motivierte (mehr als 70%) teilnahmen. Die waren beachtlich: Wenn alle anderen Merkmale gleich sind (bzw. ihr Einfluß statistisch kontrolliert wurde), unterscheiden sich die Zufriedenheitsurteile um 1,42 Skaleneinheiten (Skala von -2 [=Erwartungen „gar nicht erfüllt“] bis +2 [=„voll erfüllt“]) (Vgl. Kromrey 1996: 4).

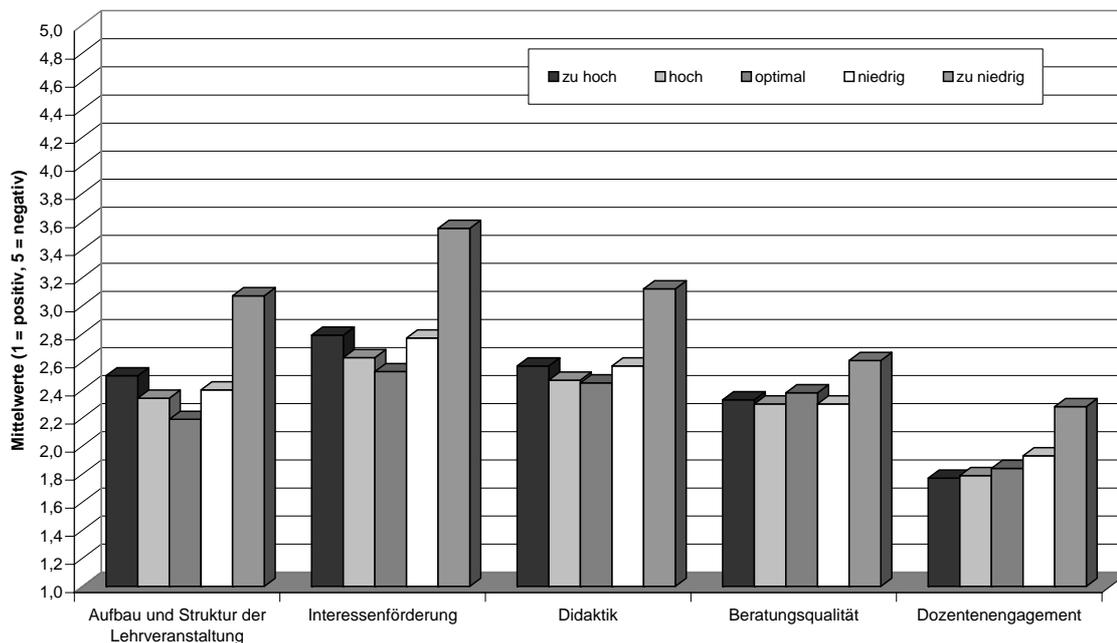
Allerdings muß die von Kromrey gewählte Darstellung der Ergebnisse dahingehend hinterfragt werden, warum er gerade 70% und 80% der Teilnehmer als Grenze zu „durchmischte“ motivierten Lehrveranstaltungen wählte. Durch diese hohe definierte Schwelle zu einer eindeutigen Motivationslage werden Aussagen über „intrinsisch“ Motivierte nur auf der Basis etwa eines Zwanzigstels der insgesamt Befragten getroffen. Zu vermuten ist, daß bei Vergrößerung der Subgruppe dadurch, daß man Lehrveranstaltungen mit einer einfachen Mehrheit, also über 50% intrinsisch Motivierter, mit den übrigen Lehrveranstaltungen vergliche, die Unterschiede nicht mehr so beeindruckend ausfielen, wenn nicht sogar insignifikant werden. Kromrey begründete seine Grenzziehung leider nicht genauer.

Noten nicht automatisch als Bias definieren lassen, „da günstige Noten auch aufgrund einer guten Lehrveranstaltung erwartet werden können“ (vgl. Rindermann 2001: 184).

Und schließlich wurde in mehreren Untersuchungen noch der Einfluß von Schwierigkeitsgrad bzw. Anforderungen der Lehrveranstaltung untersucht. Beispielhaft soll hier eine Untersuchung erwähnt werden, die von einem verzerrenden Einfluß ausgeht (Esser 1997), und eine, die diesen Einfluß relativiert (Krempkow/ Heldt 2000). Oft wurde der Schwierigkeitsgrad von Lehrveranstaltungen (wie bei Esser) mittels einer Skala erfragt, die nur eine positive (höhere Anforderungen) und eine negative Bewertung ermöglichen (niedrigere Anforderungen). Hier lassen sich dann auch z. T. die erwarteten Zusammenhänge finden und als Verzerrung der Ergebnisse interpretieren.

Untersucht man den Einfluß der Leistungsanforderungen auf einer Skala, die zusätzlich zu hohen und niedrigen Anforderungen auch die Antwortvorgabe „optimal“ (z.B. in der Mitte der Skala) enthält, so zeigt sich ein anderes Bild. Gegenübergestellt werden kann dann jeweils in Gruppen die Einschätzung der Lehrqualität von Studenten, die die Anforderungen als (zu) hoch, optimal, oder (zu) niedrig empfinden. Zwar werden auch hier Aufbau und Struktur, Interessenförderung und in geringem Maße die Didaktik der Lehrveranstaltung bei als „zu hoch“ empfundenen Anforderungen etwas schlechter (0,2 Skalenpunkte, Skala 1 bis 5) bewertet als bei hohen Anforderungen. Deutlich schlechtere Bewertungen (0,6-0,8 Skalenpunkte) erhalten die Lehrveranstaltungen jedoch bei als „zu niedrig“ empfundenen Anforderungen. Die besten Bewertungen erhalten die Lehrveranstaltungen immer dann, wenn die Anforderungen als „optimal“ empfunden werden. Dieser Effekt wurde bereits 1999 festgestellt (vgl. Krempkow/ Heldt 2000) und hat sich inzwischen mehrfach bestätigt. Das Ergebnis deckt sich im übrigen mit Erkenntnissen der Lernpsychologie. Optimale Lehr- und Lernerfolge können demnach nur erzielt werden, wenn auch die Anforderungen auf die Lernenden abgestimmt sind.

## Einschätzung der Anforderungen als Einflussfaktor für Lehrbewertungen



Quelle: Krempkow/ Winter 2000: 33

Schließlich soll als letzter der am häufigsten genannten potentiellen Einflußfaktoren noch die Kursgröße diskutiert werden, die als Rahmenbedingung der Lehrveranstaltungen deren Bewertung beeinflussen kann. In der Überblickarbeit von Hage (1996b) zeigte sich der Einfluß der Kursgröße jedoch als inkonsistent. Bei Rindermann (1996) zeigten sich zwar nennenswerte signifikante Zusammenhänge, jedoch wurden lediglich Klima, Beteiligung und Diskussion bei zunehmender Kursgröße kritischer beurteilt. Bei Spiel und Gössler (2001) wurde die Größe der Studieneinrichtung als potentieller Einflußfaktor untersucht; es fand sich allerdings kein nennenswerter Zusammenhang. Auch bei den von Rindermann in einer Übersicht dargestellten internationalen Studien fanden sich dort, wo der Einfluß der Kursgröße untersucht wurde, höchstens bei Interaktionsvariablen nennenswerte Zusammenhänge (vgl. Rindermann 2001: 181ff.).

Schwarzkopf (2001) untersuchte zusätzlich zu den hier dargestellten noch ein ganzes Set weiterer Rahmenbedingungen der Lehre und des Studiums, die Einfluß auf die Ergebnisse von Studentenbefragungen haben könnten. Sie kam jedoch zu dem Schluß, daß lediglich die Einschätzung der Organisation des Studiums und der Bereitschaft der Lehrenden, sich mit fachlichen und organisatorischen Problemen der Studierenden zu beschäftigen, mit den

Lehrbewertungen nennenswert in Zusammenhang stehen. Diese von Schwarzkopf (2001) gefundenen Zusammenhänge stehen aber ihrerseits wiederum sehr stark in Zusammenhang mit Aspekten, die nach Marsh (1982b) selbst als Kriterien von Lehr- und Studienqualität aufgefaßt werden können.

Letztlich kann die Entscheidung, ob diese Aspekte Bestandteil der Lehr- und Studienqualität sind oder „verzerrende“ Einflußfaktoren, erst anhand der konkreten Ziele getroffen werden. Rindermann (2001: 199) spricht daher in Abhängigkeit von den Zielen der Bewertung von „Veranstaltungsvalidität“ bzw. „Dozentenvalidität“, wobei er das Nichtvorhandensein von Biaseinflüssen als Indiz für Validität auffaßt. Mit „Veranstaltungsvalidität“ ist gemeint, ob die Qualität einer Lehrveranstaltung adäquat durch Studentenbefragungen umschrieben werden kann. „Dozentenvalidität“ meint, ob die Lehrfähigkeit eines Dozenten adäquat erfasst wird. Da Lehre durch Rahmenbedingungen wie z.B. Überfüllung beeinflusst werden kann, können Ergebnisse von Studentenbefragungen zwar das Geschehen in Lehrveranstaltungen realistisch beschreiben, für die Lehrfähigkeit eines Dozenten gilt dies jedoch nur eingeschränkt. Insbesondere nicht-dozentenbezogene Skalen sind anfälliger für verzerrende Einflussfaktoren, während klar dozentenbezogene Skalen relativ unanfällig sind. Die Dozentenvalidität wurde aber unabhängig von Untersuchungen der Einflussfaktoren auch durch Variation von Veranstaltungsthemen und Dozenten über verschiedene Veranstaltungen hinweg untersucht. Die dozentenbezogenen Skalen erwiesen sich dabei als dozentenattribuierbar, sofern mindestens fünf Veranstaltungen herangezogen werden (vgl. Rindermann 1996, 2001: 199).

Da Rindermann dies an verschiedenen Hochschulen und über eine große Veranstaltungsanzahl hinweg untersuchte und seine Analysen (die in den Ursprüngen auf das Jahr 1994 zurückgehen) trotz z.T. intensiver Diskussion in den vergangenen Jahren in den wesentlichen Punkten nicht widerlegt werden konnten, wird davon ausgegangen, daß für Lehrveranstaltungsbewertungen mittels HILVE bei Beachtung der Voraussetzungen zuverlässige Urteile möglich sind.

Bislang nicht derart intensiv untersucht wurde dies für Studiengangs- bzw. Fachbereichsbewertungen. So könnte das unterschiedliche Studieninteresse der Studierenden an verschiedenen Fachbereichen durch seinen Zusammenhang mit der Bewertung der Studiensituation Qualitätsunterschiede in der Lehre suggerieren, obwohl die Ursachen für unterschiedliche Bewertungen eventuell außerhalb des Verantwortungsbereichs der einzelnen Fachbereiche liegen. Andererseits könnte bei einer stärkeren Übertragung von Verant-

wortung für die Studieninformation bis hin zur Auswahl von Studierenden an die einzelnen Fachbereiche eine hohe Studienmotivation durchaus als integrierter Bestandteil optimaler Lehr- und Lernprozesse betrachtet werden. Unter diesen Voraussetzungen wäre ein höheres oder geringeres Studieninteresse kein das studentische Urteil verzerrender Bias-Effekt, sondern selber ein Indikator für "gute Lehre", die gute Studienbedingungen einschließt.

Letztlich bleibt es eine Frage des Gesamtkonzeptes der Evaluation bzw. Leistungsbewertung, ob auf ggf. vorhandene Zusammenhänge der studentischen Bewertungen mit dem Studieninteresse, den Noten oder auch der Kursgröße bei der Interpretation der Ergebnisse hingewiesen wird und dies bei Entscheidungen als integrierter Bestandteil der Studiensituation Berücksichtigung findet oder ob dies als Verzerrung gesehen und deshalb bei Vergleichen von Befragungsergebnissen "herausgerechnet" werden sollte. Erhoben und bei Vergleichen berücksichtigt werden sollten solche potentiellen Einflußfaktoren in jedem Fall.

#### **5.4.4 Weitere Anforderungen an Studentenbefragungen als Leistungsindikator**

Als weitere Anforderungen an Ergebnisse von Studentenbefragungen als möglicher Leistungsindikator sind – wie an andere Indikatoren auch – Zeitnähe, sinnvolle Anreizwirkung, Vermeidung unerwünschter Anpassungseffekte bzw. Fehlanreizwirkungen und ein vernünftiges Verhältnis von Erhebungsaufwand und erwartetem Nutzen zu nennen.

Zur Zeitnähe kann festgehalten werden, daß die Ergebnisse von Studentenbefragungen zur Qualität der Lehre – sofern Lehrveranstaltungsbewertungen wie empfohlen bereits in der Mitte des Semesters durchgeführt werden, in den allermeisten Fällen noch im laufenden Semester zur Verfügung stehen. Selbst wenn Befragungen zu Studienbedingungen von externen Institutionen wie z.B. für Studienführer wie dem des CHE durchgeführt werden, stehen deren Ergebnisse nach einigen Monaten bis ca. einem halben Jahr der Öffentlichkeit und den Hochschulen (diesen auch als separat anzufordernde detaillierte Sonderauswertungen) zur Verfügung. Damit sind sie deutlich vor den sogenannten objektiven Kennziffern bzw. Indikatoren verfügbar, die häufig erst ein Jahr oder später nach Abschluß des Erhebungszeitraumes komplett aufbereitet und publiziert sind. Im Vergleich zu anderen subjektiven Bewertungen wie z.B. Absolventenbefragungen ist zu beachten, daß deren retrospektive Bewertungen der Studiensituation sich auf einen Zeitraum beziehen, der meist schon mehrere Jahre zurückliegt. Selbst wenn (z.B. zugunsten einer intensiveren

Auswertungs- und Nachbereitungsphase bzw. ein umfassendes „Follow-up“) Studentenerbefragungen nicht jedes Semester, sondern nur einmal im Jahr oder alle eineinhalb Jahre durchgeführt würden, wäre Studentenerbefragungen insgesamt immer noch eine vergleichsweise hohe Zeitnähe zu bescheinigen.

Zur Anforderung, daß Studentenerbefragungen eine sinnvolle Anreizwirkung zur Nutzung der finanziellen Freiräume entfalten sollten, ist darauf hinzuweisen, daß sie häufig mit genau diesem Ziel zur Ergänzung vorhandener oder geplanter Leistungsbewertungs- und Anreizmodelle in die Diskussion gebracht werden. Beispielhaft soll hier zur Illustration der Aspekt der Betreuung herausgegriffen werden: Bisher erfolgt die Erfassung der Betreuungssituation häufig mit einem quantitativen bzw. objektiven Indikator - der Betreuungsrelation - und kann daher nur das Potential für eine gute Betreuung anzeigen. Einen Anreiz, tatsächlich eine gute Betreuung zu leisten, kann mit diesem objektiven Indikator nicht erzielt werden. Vielmehr kann eine geringe durchschnittliche Anzahl Studierender je Professur anstelle der Nutzung dieses Potentials für eine intensivere Betreuung (als in anderen Studiengängen) auch dazu genutzt werden, lediglich einen absoluten Mindeststandard in der Betreuung zu erfüllen und stärker als andere in der Forschung aktiv zu sein (was bei den derzeit „real existierenden“ Anreizwirkungen an deutschen Hochschulen durchaus rational scheint).

Zieht man dagegen die von den Studierenden empfundene subjektive Einschätzung der Betreuungssituation hinzu, definiert eine positive Einschätzung als anzustrebendes Ziel und faßt die - ggf. unterschiedlichen - objektiven Betreuungsrelationen als Ausgangsbedingungen auf (anhand derer die subjektiven Bewertungen zu interpretieren und an die ggf. die Zielstellungen anzupassen sind), so könnte durch die Einbeziehung des subjektiven Aspektes in vorhandene Anreizmodelle auch ein Anreiz zur Nutzung des vorhandenen besseren Betreuungspotentials geschaffen werden. Ergänzend zur Studierbarkeit und Studiendauer können Einschätzungen, ob es aus Sicht der Studierenden/ Absolventen in der Studienorganisation Gründe gab, die eine Einhaltung der Regelstudienzeit erschwerten (und ggf. um welchen Zeitraum), hilfreiche Anreize setzen, die von den Studierenden wahrgenommenen Hindernisse und Schwierigkeiten ernst zu nehmen und zu vermindern. Ähnlich wäre dies auch für andere Aspekte denkbar, wie z.B. subjektive Einschätzungen zum Aufbau und zur Strukturierung der Lehre und zur didaktischen Gestaltung als Beispiele für weitere Indikatoren der Prozeßqualität, die einen Anreiz zur Schaffung günstiger Bedingungen für einen höheren Lehr-/ Lernerfolg schaffen könnten. Damit wären solche

Indikatoren der (Prozeß-) Qualität eine sinnvolle Ergänzung von objektiven Indikatoren wie z.B. Anzahl der Absolventen (ggf. in Relation zum entspr. Studienanfängerjahrgang), die lediglich „Stückzahlen“ erfassen können, aber nicht, inwieweit die Bedingungen hierfür günstig waren.

Selbst für Einschätzungen zu Leistungsanforderungen scheint dies denkbar. (Sind diese zu niedrig oder viel zu hoch, wird das Lernpotential ebenfalls nicht ausgeschöpft und sollte daher möglichst leicht über dem als Optimum empfundenen Anforderungsniveau liegen.) Hierzu sind aber sicherlich weitere Erfahrungen mittels empirischer Untersuchungen zu sammeln, ob hierbei nicht evtl. unerwünschte Anpassungseffekte auftreten.

Die Vermeidung solcher Anpassungseffekte ist eine weitere Anforderung an Indikatoren. Insgesamt sollte es darum gehen, mittels Einbeziehung subjektiver Indikatoren neben den objektiven Indikatoren ein ausgewogenes Verhältnis von Quantität und Qualität herzustellen. Auf diese Weise könnte einem möglichen unerwünschten Anpassungseffekt entgegengewirkt werden, daß ein Studiengang oder eine Hochschule die denkbar höchste (z.B. Wissenschaftlernachwuchs-)Qualität ohne Rücksicht auf den quantitativen Absolventenbedarf an in der Gesellschaft als Ziel definiert und nur sehr wenige Absolventen ausbildet.

Ein vernünftiges Verhältnis von Erhebungsaufwand und erwartetem Nutzen wurde als weitere Anforderung an Indikatoren formuliert. Bei Studierendenbefragungen, die ohnehin durchgeführt werden (sofern sie den formulierten methodischen Standards entsprechen), ist ein sehr günstiges Verhältnis schon allein dadurch gegeben, daß die vorhandenen Ergebnisse genutzt werden können. Ist dies nicht der Fall, muß eine Abwägung der Kosten zum erwarteten Nutzen geschehen. Allerdings sollte eine solche Abwägung nicht allein der Hochschule oder den Hochschulen überlassen werden, für die sich u.U. aufgrund der existierenden Rahmenbedingungen keine günstige Kosten-Nutzenrelation ergeben kann. Vielmehr müssten unter Berücksichtigung von zu erwartenden Folgekosten etwaiger Fehlanreizwirkungen für die Volkswirtschaft die Rahmenbedingungen durch die Legislative und/ oder Exekutive so gesetzt werden, daß die Erhebung von volkswirtschaftlich wichtigen Aspekten mittels qualitativer Indikatoren in einer günstigen Kosten-Nutzen-Relation stehen.<sup>193</sup>

---

<sup>193</sup> Beispielsweise sind Absolventenbefragungen zum beruflichen Verbleib und zur retrospektiven Einschätzung des Studiums relativ aufwendig, werden andererseits aber von vielen Autoren als einzige Möglichkeit genannt, die Qualität der Ausbildung anhand ihrer Wirkung zu erfassen.

#### **5.4.5 Fazit zur Eignung von Studentenforschungen**

Die Diskussion zur Validität und Reliabilität läßt sich folgendermaßen zusammenfassen: Den vorliegenden Ergebnissen empirischer Untersuchungen zufolge müssen studentische Urteile zur Lehr- und Studiensituation als weitgehend valide und zuverlässig bzw. reliabel eingeschätzt werden. Zunächst konnte bei den Untersuchungen zur Validität festgestellt werden, daß die Urteile von Lehrenden, Fremdgutachtern und Studierenden bei Anlegen gleicher Kriterien und konkreter Fragen hierzu weit besser übereinstimmen, als häufig angenommen wird. Auch zur Zuverlässigkeit kann ähnliches berichtet werden: Sowohl bei Meßwiederholung innerhalb eines Semesters, als auch über mehrere Semester hinweg sind die Urteile der Studierenden über das Lehrverhalten ihrer Dozenten weitgehend stabil. Gleichwohl sollten immer mehrere (mindestens fünf) Lehrveranstaltungen zur Beurteilung des Lehrverhaltens von Dozenten herangezogen werden und an einigen Punkten besteht noch weiterer Untersuchungsbedarf – z.B. zur Stabilität der Ergebnisse von Studentenforschungen an deutschen Hochschulen, die über das Heidelberger Inventar (HILVE) hinausgehen. Will man weitergehende Aussagen zur Eignung von Studentenforschungen als ein Qualitätskriterium in Leistungsvergleichen treffen, dann sollten aber noch weitere Aspekte berücksichtigt werden. So zählte neben den am Anfang dieses Kapitels genannten Autoren wie Ziegele auch Rindermann das Nichtvorhandensein von „verzerrenden“ Bias-Variablen bzw. Einflußfaktoren mit zu seinen Eignungskriterien.

Als die häufigsten genannten Einflußfaktoren, die die Vergleichbarkeit beeinträchtigen könnten, werden Sachkenntnis und Reife, Geschlecht, Noten, Leistungsanforderungen, Interesse und Kursgröße genannt. Hierbei haben unterschiedliche Sachkenntnis und Reife von Studierenden (gemessen an evtl. Hilfskrafttätigkeit, Studienintensität, Hochschulwechsel oder Alter) und Geschlecht keinen nennenswerten Einfluß. Lediglich die Semesterzahl hat einen geringen Einfluß. Allerdings urteilen Studierende in höheren Fachsemestern (also höherer Sachkenntnis) eher kritischer. Auch der Einfluß der Noten bzw. Notenerwartung – der zur Vermeidung eines Zirkelschlusses (durch den Einfluß von realen, durch unterschiedliche Lerneffektivität innerhalb eines Kurses verursachten Leistungsunterschieden) nicht innerhalb einer Lehrveranstaltung, sondern über mehrere Lehrveranstaltungen hinweg untersucht werden muß – fiel in verschiedenen Studien eher gering aus oder war nicht nachweisbar.

Für die Leistungsanforderungen war im Einklang mit Erkenntnissen der Lernpsychologie festzustellen, daß keineswegs bei den niedrigsten Leistungsanforderungen die besten Bewertungen vergeben wurden. Vielmehr wurden diese bei als optimal empfundenen und etwas höheren Anforderungen erzielt. Die niedrigsten Leistungsanforderungen gingen dagegen mit den schlechtesten Bewertungen einher. Lediglich der Einfluß des Interesses der Studierenden entsprach den Vermutungen und zeigte sich damit im Verhältnis zu den anderen potentiellen Einflußfaktoren als stärkste „Bias“-Variable. Studierende mit höherem Fachinteresse beurteilen Lehrveranstaltungen besser als Studierende mit geringerem.

Neben den Anforderungen an die Datenqualität wurden Zeitnähe, sinnvolle Anreizwirkung, Vermeidung unerwünschter Anpassungseffekte bzw. Fehlanreizwirkungen und das Verhältnis von Aufwand und Nutzen genannt. Hierzu kann festgehalten werden, daß Studentenbefragungen meist eine vergleichsweise hohe Zeitnähe haben, daß sie in Verbindung mit quantitativen bzw. objektiven Indikatoren eine sinnvolle Anreizwirkung entfalten und dabei unerwünschte Anpassungseffekte gerade vermeiden könnten. Da Studentenbefragungen zur Lehrqualität ohnehin gesetzlich gefordert sind und meist auch durchgeführt werden, ist (wenn diese den Anforderungen entsprechen) ein günstiges Verhältnis von Aufwand und Nutzen festzustellen. Wenn Studentenbefragungen den methodischen Standards nicht entsprechen, sollten sie in jedem Falle entsprechend fortentwickelt werden, so daß sie eine zuverlässige Grundlage für die durchzuführende Bewertung der Lehre darstellen.

Bei Indikatoren für Leistungsbewertung und Leistungsanreize spielt aber auch ihre Akzeptanz eine wesentliche Rolle (vgl. Abschnitt zu Leistungsanreizmodellen in dieser Arbeit). Diese Akzeptanz der Einbeziehung von Studentenbefragungen in Anreizmodelle ist im Vergleich z.B. zu Absolventenbefragungen geringer (vgl. Ergebnisse eigener Lehrendenbefragungen in dieser Arbeit).<sup>194</sup> Zudem eignen sie sich nicht zur Erfassung langfristiger Wirkungen des Studiums. Daher sollen neben Studentenbefragungen nachfolgend auch Absolventenbefragungen zu ihrer Eignung für Indikatoren diskutiert werden.

---

<sup>194</sup> Andererseits wird trotz der von verschiedenen Seiten geäußerten Kritik (vgl. auch die Diskussion von (Abschluß-)Noten als objektive Indikatoren in diesem Kapitel) in anderen Zusammenhängen selten verlangt, gleich hohe Anforderungen z.B. an die Reliabilität von Hochschulprüfungen zu stellen (vgl. z.B. Kriz 1994: 11) und „dann zu verlangen, wenn diese Bedingungen nicht erfüllt werden, Prüfungen abzuschaffen“ (vgl. Rindermann 2001: 204). Wie bei Prüfungen, für die ihre praktische und gesellschaftliche Notwendigkeit eine größere Rolle spiele als Genauigkeits- und Validitätsfragen, müsse auch für Studentenbefragungen entschieden werden, ob deren Nutzen mögliche Ergebnisgenauigkeiten aufwiegt, argumentiert Rindermann (2001: 204).

## **5.5 Die Eignung von Absolventenbefragungen zur Messung der Wirkungen der Hochschulausbildung**

In den vorangehenden Abschnitten wurde festgestellt, daß sogenannte objektive Indikatoren Aspekte von Qualität nicht in ausreichendem Maße erfassen können. Auch Studentebefragungen zur Qualität der Hochschullehre sind trotz zahlreicher Belege ihrer Validität und Zuverlässigkeit (vgl. Abschnitt zur Datenqualität von Studentebefragungen) nach wie vor umstritten und können in der Tat potentiell durch einige Biasfaktoren wie das Studieninteresse beeinflußt sein, die eine Vergleichbarkeit erschweren und daher zu berücksichtigen sind. Daher kamen Anfang der 1990er Jahre einige Autoren zu dem Schluß, daß einem Teil der Einwände gegen eine Evaluation von Hochschulen mittels objektiver Indikatoren und subjektiver Bewertungen mittels Studentebefragungen mit Absolventenstudien begegnet werden kann (vgl. Teichler 1993: 92). Inzwischen wird der berufliche Erfolg von Hochschulabsolventen häufiger als Indikator für die Wirkung der Hochschulausbildung genannt (vgl. z.B. Ziegele 2002: 11). Nicht immer wird allerdings in diesem Zusammenhang ausgeführt, woran genau dieser gemessen werden soll.

Nach einer grundsätzlichen Diskussion der Möglichkeiten und Grenzen von Absolventenstudien und Erörterung der Anforderungen an die Datenqualität von Absolventenbefragungen soll vor allem diskutiert werden, welche Ergebnisse evtl. geeignet sind, adäquat den Berufserfolg von Hochschulabsolventen als Indikator für die Wirkung der Hochschulausbildung zu erfassen. Daneben werden jedoch auch weitere Anforderungen wie Vergleichbarkeit, Zeitnähe und mögliche Anpassungseffekte in die Diskussion einbezogen.

### **5.5.1 Wie können Wirkungen der Hochschulausbildung mittels Absolventenstudien „gemessen“ werden?**

Absolventenstudien erfreuen sich – nach einer Konjunktur in der 1970ern und enttäuschten Erwartungen in den 1980ern – seit den 1990er Jahren im Zusammenhang mit dem verstärkten Interesse an Evaluation wieder wachsender Beliebtheit (vgl. Teicher/Schomburg 1997: 243ff.). Absolventenstudien umfassen insgesamt nicht nur Absolventenbefragungen, sondern auch Arbeitgeberbefragungen (z.B. DIHK 2004) oder Analysen von Personalakten in Unternehmen (vgl. Teichler/Schomburg 1997: 246ff.). Nachfolgend sollen jedoch lediglich Absolventenbefragungen diskutiert werden, da nur diese von den

Hochschulen direkt und ohne ggf. problematische Abstimmungen mit weiteren Institutionen durchführbar sind.

Die Möglichkeiten von Absolventenbefragungen werden vor allem darin gesehen, daß sie „ihren Stellenwert in erster Linie für die Messung der Erträge der Hochschulausbildung (‘outputs’ und ‘outcomes’) haben“ (vgl. Teicher/ Schomburg 1997: 246). Diesen Stellenwert können andere Indikatoren zur Hochschullehre (wie die in diesem Kapitel diskutierten objektiven Indikatoren und Studentenforschungen) nicht einnehmen. Lediglich ein Testdesign für Studenten, das die Hochschuleingangsfähigkeiten im Vergleich zu den Fähigkeiten bei Hochschulabschluß erfasst, könnte dies theoretisch leisten. Allerdings würde auch dies nur Teilaspekte erfassen können, da z.B. unklar bleiben müsste, welche dieser Fähigkeiten/ Kompetenzen besonders wichtig für den Berufseinstieg und beruflichen Erfolg sind.

Die Möglichkeiten von Absolventenbefragungen erschöpfen sich jedoch nicht in der Bewertung der Erträge im engeren Sinne entsprechend der intendierten Ziele („output“) bzw. der Wirkungen in einem weiteren Sinne („outcomes“). In einer systematischen Evaluation mittels Absolventenbefragungen ist es nach Teichler/ Schomburg (1997: 245) möglich:

- den Einfluß von Studienvoraussetzungen zu bestimmen, um nicht bestimmte Erträge (oder Probleme) fälschlicherweise der Hochschule zuzuschreiben
- den Einfluß von Ressourcen und Rahmenbedingungen zu bestimmen, z.B. räumliche Bedingungen, Ausstattung der Bibliotheken bzw. Labore
- den Einfluß der (Lehr- und Lern-)Prozesse zu bestimmen, z.B. Lehrveranstaltungen, Beratung usw.
- bezüglich der Erträge Unterscheidungen vorzunehmen, insbesondere erstens zwischen den Kompetenzen bei Studienabschluß (wie oben bereits angesprochen) und der symbolischen Dokumentation von Kompetenzen in Zertifikaten sowie zweitens zwischen unmittelbaren Ergebnissen des Studiums und späteren Aktivitäten und Leistungen im Beruf (die durch Ergebnisse des Studiums mitverursacht sind).

Da die Beschränkung der Fragestellung auf eine reine Ertragsmessung den Hochschulen neben der Feststellung (und ggf. Einordnung ihrer Leistungen im Vergleich zu anderen) nicht in ausreichendem Maße ermöglicht, Ansatzpunkte für Veränderungen zu gewinnen, sollten Absolventenbefragungen „immer auch die Prozesse an den Hochschulen so weit einbeziehen, daß sie zur Erklärung der Erträge beitragen können“ (vgl. Teichler/ Schomburg

1997: 247). Auch eine Beschränkung von Absolventenstudien auf die Feststellung des beruflichen Erfolges und die Zufriedenheit mit dem Studium wie bei dem größten Teil der Absolventenstudien, die die Hochschulen in den USA durchführen (vgl. auch Reinfeldt/ Frings 2003: 282), schließe aus „zu prüfen, ob bestimmte Erfahrungen mit bestimmten Erfolgen im Beruf verknüpft sind“ (vgl. Teichler/ Schomburg 1997: 247). Darüber hinauszugehen, impliziert Analysen zu Determinanten bzw. Bestimmungsgründen beruflichen Erfolges, wie sie von Teichler/ Schomburg (2001) sowie relativ umfassend auch im Rahmen der Dresdner Absolventenstudien durchgeführt wurden, die in einem der nachfolgenden Abschnitte auch unter dem Aspekt der Vergleichbarkeit der Ergebnisse von Absolventenbefragungen angesprochen werden (vgl. Krempkow/ Popp 2003a).

Eine Analyse der Bestimmungsgründe beruflichen Erfolges setzt eine Klärung voraus, welche Kriterien für den beruflichen Erfolg herangezogen werden sollen. Zunächst ist beruflicher Erfolg abzugrenzen von Studienergebnissen, da es nur bei getrennter Analyse möglich ist, das Spannungsverhältnis zwischen Ergebnissen des Studiums und deren beruflicher Belohnung zu betrachten. Teichler und Schomburg (1997: 247) betonen, daß es darum gehe, den Hochschulen Informationen hierzu bereitzustellen, nicht jedoch um eine „Subsumption der Hochschulen unter die Anforderungen des Beschäftigungssystems“. Die Autonomie der Hochschulen in der Setzung ihrer Ziele werde gerade nicht gefördert, wenn die Hochschulen nicht über Informationen zum Schicksal ihrer Absolventen verfügen.

Entsprechend den jeweils gesetzten Zielen gebe es ein breites Spektrum an Themen, die bei bereits durchgeführten Befragungen von Hochschulabsolventen erhoben wurden und als Maße zu den beruflichen Erträgen des Studiums herangezogen werden können. Teichler und Schomburg (1997: 248)<sup>195</sup> faßten diese zu insgesamt vier Aspekten zusammen:

### **Mögliche Kriterien beruflichen Erfolges nach Teichler/ Schomburg**

- Objektive Maße für den Übergang vom Studium in den Beruf
- Objektive Maße für den Berufserfolg
- Subjektive Maße für den Berufserfolg
- Einschätzungen zum Zusammenhang von Studium und Beruf

---

<sup>195</sup> Von Teichler erschienen in nachfolgenden Jahren weitere Veröffentlichungen zu Absolventenstudien, in denen Aspekte des beruflichen Erfolges genannt wurden und die z.T. in ihrer Gliederung variierten (vgl. z.B.

Als objektive Maße für den *Übergang vom Studium in den Beruf* nennen Teichler und Schomburg die Dauer der Beschäftigungssuche bzw. die Zeitspanne bis zur ersten regulären Beschäftigung, die Anzahl vergeblicher und erfolgreicher Bewerbungen, längerfristige Arbeitslosigkeit sowie Status- und Beschäftigungsmerkmale der ersten beruflichen Tätigkeit. Einschränkend wird angemerkt, daß der Berufseinstieg nicht immer wegweisend für die Verwendungschancen des Erlernenen sind.

Als *objektive Maße für den Berufserfolg* nennen Teichler und Schomburg Einkommen und betriebliche Position. Andere Autoren nennen hierzu auch weitere Merkmale wie Umfang der beruflichen Tätigkeit (Wochenarbeitszeit) und Befristung (vgl. Rostampour/ Lembert 2003: 163). Während das Einkommen von den meisten Absolventenstudien herangezogen wird, werden bei der betrieblichen bzw. beruflichen Position z.T. Probleme bei der Zuordnung beispielsweise von Selbständigen und Freiberuflern (vgl. Plicht/ Schreyer 2002: 533ff.) sowie im Hochschulbereich gesehen. Auch die Befristung und der Umfang der Erwerbstätigkeit eignen sich nur bedingt, da in einigen Bereichen, wie z.B. im öffentlichen Dienst oder im Hochschulbereich zu einem großen Teil nur noch befristete Verträge bzw. häufig Teilzeitstellen abgeschlossen werden.<sup>196</sup>

Als *subjektive Maße für den Berufserfolg* werden berufliche Zufriedenheit, und dabei als einzelne Aspekte u.a. die Selbständigkeit bei der Arbeitsgestaltung (Autonomie), Kooperation mit Kollegen und die Sicherheit des Arbeitsplatzes genannt (vgl. Teichler/ Schomburg 1997/ 2001, Rostampour/ Lembert 2003: 163). Andere Autoren nennen auch das Kriterium der beruflichen Eingliederung bzw. Einbindung (vgl. Zimmermann 1999).<sup>197</sup>

---

Teichler 2002: 13ff., Burkhardt u.a. 2000). Es wurden im Wesentlichen dieselben Einzelaspekte genannt, diese allerdings nicht so ausführlich diskutiert wie in Teichler/ Schomburg (1997).

<sup>196</sup> Rostampour/ Lembert (2003: 163ff.) kamen in ihren Regressionsanalysen zum Berufserfolg von Diplompädagogen zu dem Ergebnis, daß „eine Vollzeittätigkeit unabhängig vom Einkommen für die Befragten keinen Erfolg darstellt“. Auch die Befristung zeigte keinen Einfluß auf den subjektiv empfundenen Berufserfolg.

<sup>197</sup> Zimmermann formulierte hier eine pädagogische Operationalisierung unter Rückgriff auf die Selbstbestimmungstheorie und ausgehend von der zentralen Annahme, daß der Mensch von Natur aus Regulationsmechanismen der sozialen Umwelt verinnerlicht, um sich selbst als Mitglied in diese einzufügen. Demnach ließen sich drei dem Menschen angeborene Grundbedürfnisse identifizieren: Die Erfahrung eigener Kompetenz, Autonomie und sozialer Eingebundenheit in die Gesellschaft (bzw. in den Betrieb im beruflichen Kontext). Ein erfolgreicher Verbleib ist daher lediglich dann gegeben, wenn das Individuum nach der Ausbildung eine Erwerbstätigkeit findet, in der es sich kompetent, autonom und sozial eingebunden fühlt (vgl. Zimmermann 1999: 48ff). Während die Autonomie hier bereits explizit bei der beruflichen Zufriedenheit angesprochen wurde und das Kompetenzerleben als ein Teilaspekt der Ausbildungsadäquanz betrachtet werden kann (z.B. der Qualifikationsadäquanz), wird die soziale Eingebundenheit häufig nur indirekt und relativ allgemein über die Einschätzung des Arbeits- und Betriebsklimas erfaßt (vgl. z.B. Teichler/ Schomburg 2001, Krempkow 2001 und Folgejahre). Gleichwohl kann man nicht davon sprechen, daß dieser Aspekt völlig unberücksichtigt bleibt, da er in die Einschätzung des Arbeits- und Betriebsklimas einfließen dürfte.

Für Einschätzungen zum *Zusammenhang von Studium und Beruf* nennen Teichler und Schomburg als Beispiele das Ausmaß der Qualifikationsverwendung generell und die berufliche Nützlichkeit verschiedener Studieninhalte. Die Einschätzungen zum Zusammenhang von Studium und Beruf werden z.T. unterschiedlich benannt und operationalisiert (vgl. Brüderl/ Reimer 2002). Andere Autoren bezeichnen sie auch als Adäquanz der Akademikerbeschäftigung (vgl. Plicht/ Schreyer 2002) bzw. kürzer Ausbildungsadäquanz (vgl. z.B. Büchel/ Weißhuhn 1997). Die Aussagekraft von Einschätzungen zum Zusammenhang von Studium und Beruf wird nach wie vor sehr unterschiedlich eingeschätzt. Während einige Autoren zumindest einzelne Aspekte durchaus als Maß beruflichen Erfolges sehen (vgl. Schomburg/ Teichler 1997: 249), ist es für andere Autoren eine die Absolventen überfordernde „Expertise zum Durchschauen solcher Zusammenhänge“ (vgl. Burkhardt u.a. 2000: 17). Dennoch werden in vielen Absolventenstudien Fragen zum Zusammenhang von Studium und Beruf verwendet. Da eine „gelungene Anbindung des Studiums an die berufliche Praxis eine der wesentlichen Forderungen (ist), die in den vergangenen Jahren die Reform von Studium und Lehre an den deutschen Hochschulen prägten“ (vgl. Reinfeldt/ Frings 2003: 279), müssen sich Hochschulen jedenfalls mit dieser Forderung auseinandersetzen. Eine adäquate Beschäftigung steht aber auch nach theoretischen Überlegungen mit dem Erleben von Kompetenz als menschlichem Grundbedürfnis (Vgl. Zimmermann 1999) in Zusammenhang. Daher soll die Ausbildungsadäquanz hier ausführlicher diskutiert werden. Plicht/ Schreyer (2002) unterscheiden hierbei nach vertikaler und horizontaler Ebene. Dabei beschreibt die vertikale Ebene die positionale Passung des Ausbildungsniveaus mit der beruflichen Stellung und kann z.B. anhand der Abschlußadäquanz gemessen werden (Wird der Hochschulabschluss von den Absolventen als nützlich für die berufliche Karriere erachtet?). Teilweise wird für die vertikale Ebene auch die Statusadäquanz herangezogen (z.B. über die subjektive Einschätzung der Angemessenheit der beruflichen Position). Die horizontale Ebene der Ausbildungsadäquanz umfasst die inhaltliche Passung der in der Ausbildung erworbenen Qualifikationen mit der ausgeübten Tätigkeit und kann anhand der Fachadäquanz (Verwendbarkeit der Studieninhalte im Beruf) und Qualifikationsadäquanz (gemessen z.B. anhand von beim Berufsstart aufgedeckten Qualifikationsdefiziten) untersucht werden (vgl. Krempkow/ Pastohr 2004).<sup>198</sup>

---

<sup>198</sup> Die Adäquanz der Beschäftigung unterliegt methodischen Abgrenzungsproblemen. So muß eine Diskrepanz zwischen im Studium erworbenen und in der Beschäftigung geforderten Qualifikationen nicht prinzipiell als (fachliche/ qualifikatorische bzw. horizontale) Inadäquanz gewertet werden (vgl. Plicht/

Zu den verschiedenen Teilaspekten der Ausbildungsadäquanz muß einschränkend festgestellt werden, daß Inadäquanz in einem Aspekt mit einer hohen Adäquanz in einem anderen Aspekt einhergehen kann. So könnte beispielsweise Einkommen und Statusadäquanz eines Germanisten in einem Kleinverlag sehr niedrig, die Fach- bzw. Qualifikationsadäquanz jedoch sehr hoch sein (vgl. Plicht/ Schreyer 2002). Wenn solchen Absolventen Aspekte wie Karriere, Einkommen und Status relativ unwichtig, eine interessante Tätigkeit jedoch deutlich wichtiger sind, werden sie sich neben dem Erreichen einer hohen beruflichen Zufriedenheit auch als beruflich erfolgreich sehen.<sup>199</sup> Teichler und Schomburg (1997: 249) formulieren hierzu: „Eine Hochschule oder ein Fachbereich kann dann als erfolgreich bezeichnet werden, wenn viele der Absolventen fachlich affine Aufgaben übernehmen, selbst wenn damit nicht der höchste berufliche Status verbunden ist.“ Teichler und Schomburg sehen damit eine hohe Fachadäquanz allein als hinreichendes Kriterium beruflichen Erfolges. Andererseits ist dem entgegenzuhalten, daß sie mit ihrer Formulierung davon ausgehen, daß *allen* Absolventen Aspekte wie Karriere, Einkommen und Status relativ unwichtig und eine interessante Tätigkeit wichtiger sind, was – wie z.B. in den wirtschaftswissenschaftlichen Fächern – nicht der Fall sein muß. Hinzu kommt, daß sich in der Phase der Familiengründung die Bedeutung des Einkommens gegenüber anderen Aspekten der Tätigkeit erhöhen dürfte (insbesondere wenn Haushalte mit Kindern z.B. wegen fehlender Kinderbetreuungsmöglichkeiten auf nur ein Einkommen angewiesen sind). Auch Brüderl/ Reimer (2002: 209) weisen darauf hin, daß Erfolg zeitveränderlich ist. Daher

---

Schreyer 2002: 531). Einige Akademikerberufe weisen überdies eine geringere (positionale bzw. vertikale) betriebliche Statusadäquanz auf als andere und sind außerdem historisch-gesellschaftlich variabel, während andere scheinbar dauerhaft statusadäquat sind. (vgl. ebd.). Neben Berufen mit rigiden Zugangsregelungen (z.B. Apotheker) gibt es auch sogenannte „Mischberufe“ (z.B. Künstler, Versicherungsberufe, Abgeordnete), deren Zugangsregelungen weniger klar abgegrenzt sind (Abschlussadäquanz) (vgl. Plicht/ Schreyer 2002: 533ff.). Solche methodischen Abgrenzungsprobleme erschweren eine eindeutige graduelle Zuordnung zwischen beruflichem Erfolg und Mißerfolg. Um dem Phänomen Rechnung zu tragen, daß „Adäquatheit“ unterschiedlich aufgefasst und bewertet wird, kann die subjektive Einschätzung der Amortisierung der Hochschulausbildung im Beruf herangezogen werden. Diese ist zwar von anderen Faktoren (wie der Einstellung zur Arbeit) nicht unbeeinflusst, ermöglicht jedoch zumindest eine Darstellung der Wahrnehmung der betroffenen Absolventen. So können Berufsanfänger beispielsweise temporäre Einschränkungen bei Status und Einkommen in Kauf nehmen, wenn die Beschäftigung interessant ist und inhaltlich mit den Studienschwerpunkten konform geht (vgl. Teichler u.a. 1998: 150).

<sup>199</sup> Tatsächlich wird der Aspekt „interessante Tätigkeit“ beim Berufsstart von der Mehrheit der Absolventen z.B. an der TU Dresden in fast allen Fächern als wichtiger eingeschätzt als der Aspekt „Hohes Einkommen und Karrierechancen“ (vgl. Krempkow/ Popp 2003b: 125). Auch bundesweit zeigten sich bei den meisten Fächern ähnliche Ergebnisse. Ausnahmen bilden demnach aber insbesondere die Absolventen der wirtschaftswissenschaftlichen Fächer, die überwiegend „statusorientiert“ seien (vgl. Teichler/ Schomburg 2001: 151).

wird hier davon ausgegangen, daß sich hohe und niedrige Adäquanz in verschiedenen Aspekten der Ausbildungsadäquanz *gegenseitig* kompensieren können.

Unter dieser Voraussetzung würde es dann Sinn machen, Teilaspekte additiv zu einem Gesamtscore der Adäquanz zu verrechnen (vgl. für Fachadäquanz z.B. Brüderl/ Reimer 2002: 210). Diese Vorgehensweise würde der Tatsache gerecht, daß für die Beurteilung der Ausbildungsadäquanz einer Beschäftigung nicht alle Kriterien (voll) ausgeprägt sein müssen und unterschiedlich gewichtet sein bzw. das Fehlen anderer ausgleichen können, wie es Teichler u.a. (1981: 57ff.) zu einem früheren Zeitpunkt formulierten. Beispielsweise könnten die Variablen zur Ausbildungsadäquanz zu einem Gesamtindex aggregiert werden, indem man die drei bzw. vier Variablen „Abschlussadäquanz“ (und ggf. Statusadäquanz), „Fachadäquanz“ und „Qualifikationsadäquanz“ additiv zu einer Variable „Ausbildungsadäquanz“ zusammenführt (vgl. Krempkow/ Pastohr 2004). Durch die Berechnung der „mittleren“ Adäquanz können sich die einzelnen (untergeordneten) Variablen zumindest teilweise gegeneinander aufwiegen. Auf diese Weise lassen sich verschiedene Aspekte in eine Gesamtdarstellung der Ausbildungsadäquanz integrieren. Möglicherweise lassen sich bei entsprechend differenzierter Erfassung und Auswertung von Einschätzungen zum Zusammenhang von Studium und Beruf auch die Skeptiker der Ausbildungsadäquanz überzeugen, daß ihre Einbeziehung einen Beitrag zu einer umfassenden Betrachtung und Erklärung beruflichen Erfolges liefern kann. Hierzu sind aber sicherlich weitere Analysen erforderlich.

Jedenfalls könnten neben einer differenzierteren Erfassung und Integration von verschiedenen Aspekten der Ausbildungsadäquanz durch eine Indexbildung zugleich auch die Daten zuverlässiger werden, da ein solcher Index (Eindimensionalität vorausgesetzt) als Dimension ähnlich wie bei den Analysen zur Zuverlässigkeit bei Studentenforschungen aufgefaßt werden kann, bei der sich die Fehlervarianzen durch die Zusammenfassung reduzieren würden (vgl. Rindermann 2001). Aspekte der Datenqualität sollen jedoch im folgenden Abschnitt diskutiert werden.

Zusammenfassend kann zu Möglichkeiten der „Messung“ der Wirkungen der Hochschulbildung mittels Absolventenstudien formuliert werden, daß in verschiedenen Studien bereits eine Vielzahl von Kriterien eingesetzt wurde, die hierfür herangezogen werden können. Die Kriterien sollten sich dabei an den ggf. unterschiedlichen Zielen der einzelnen Hochschulen bzw. Fachbereiche orientieren. Sie sollten dabei sowohl objektive Kriterien wie Einkommen enthalten als auch subjektive wie berufliche Zufriedenheit bzw.

Einschätzungen des Zusammenhanges von Studium und Beruf. Objektive Kriterien wie das Einkommen wären als alleiniges Erfolgskriterium „sicherlich nicht hinreichend“ (vgl. Brüderl/ Reimer 2002: 209). Wird neben Aspekten der beruflichen Zufriedenheit der Zusammenhang von Studium und Beruf hinzugezogen, der auch als Ausbildungsadäquanz bezeichnet wird, sollte er jedoch mittels mehrerer Einzelaspekte erfaßt werden, da diese für die Absolventen unterschiedlich wichtig sind und sich gegenseitig kompensieren können. Gegebenenfalls können diese Einzelaspekte dann additiv zu einem Gesamtindex zusammengefasst werden.

Sofern Absolventen relativ kurz nach Studienabschluß befragt werden, können auch Kriterien des Überganges von der Hochschule in den Beruf herangezogen werden, wie z.B. die Dauer der Stellensuche oder die Anzahl der Bewerbungen. Diese sollten dann allerdings lediglich als Kriterien des Berufseinstiegs, nicht jedoch als Kriterien beruflichen Erfolges aufgefasst werden, da ihre Aussagekraft für den langfristigen beruflichen Erfolg begrenzt ist.

### **5.5.2 Die Datenqualität der Ergebnisse von Absolventenbefragungen**

Möglichkeiten zur Überprüfung der Validität und Reliabilität von Befragungen wurden bereits im Zusammenhang mit Studentenbefragungen ausführlicher angesprochen. Prinzipiell lassen sich Absolventenbefragungen nach demselben Muster überprüfen. Allerdings sind Absolventenbefragungen erheblich weniger umstritten. Daher soll hier kürzer und eher überblicksartig darauf eingegangen und v.a. Besonderheiten von Absolventenbefragungen dargestellt werden. Grundsätzlich werden Ergebnisse von Absolventenbefragungen v.a. in zweierlei Hinsicht genutzt: Einerseits geschieht dies zur retrospektiven Bewertung der Situation von Lehre und Studium, z.B. zu Lehrqualität und Studienbedingungen, teilweise aber auch zur retrospektiven Bewertung des Stellenwertes von Lehrinhalten im Studium (vgl. z.B. Meinefeld 2002: 78). Andererseits werden sie zur Informationsgewinnung über den beruflichen Verbleib bzw. Erfolg der Absolventen genutzt.

Zu ersterem Verwendungszweck formulieren einige Autoren Einschränkungen. So liegen nach Teichler und Schomburg (1997) Grenzen der Absolventenstudien, die grundsätzlich ihre Aussagekraft bezüglich bestimmter Aspekte einschränken, darin, daß Prozesse von Lehre und Studium nur retrospektiv zu erfassen sind, wodurch die Tiefe und Verlässlichkeit von Informationen beeinträchtigt sein kann. Allerdings wurde dies nicht mit konkreten

Ergebnissen von Absolventenbefragungen belegt, so daß diese Einschränkung vermutlich eher allgemeiner Natur ist und sich generell auf retrospektive Einschätzungen unabhängig von Absolventenbefragungen bezieht. Im Zusammenhang mit der Diskussion der Kriteriumsvalidität von Studentenbefragungen zur Lehrqualität wurden Beispiele anhand von Vergleichen der Ergebnisse von Studenten- und Absolventenbefragungen genannt, die zeigen, daß diese Einschränkung in unterschiedlichem Ausmaß gelten kann. So schätzten Absolventen rückblickend die Situation von Lehre und Studium tendenziell positiver ein, ohne daß sich allerdings die Rangfolge in der Bewertung der einzelnen Aspekte grundsätzlich unterschied. Nach Rindermann (2001) könnte man hier von einer „Parallelverschiebung“ der Urteile sprechen. Dies schränkt allerdings lediglich die absolute Aussagekraft der Ergebnisse ein, die kritisch bewerteten Aspekte bleiben dieselben. Die bisher vorgelegten Ergebnisse sind jedoch noch nicht verallgemeinerbar. Wirklichen Aufschluß über diese Frage könnten erst Ergebnisse von Panelanalysen derselben Befragten als Studenten und Absolventen geben, die bislang nicht vorliegen. Zwar wurden im Zuge der Dresdner Absolventenstudien mit der ersten Welle von Wiederholungsbefragungen im Jahre 2004 auch Panelanalysen zur Stabilität der Einschätzungen zur Situation von Lehre und Studium an den einzelnen Fakultäten durchgeführt und in diesen zeigten sich bis auf einzelne Items keine signifikanten Unterschiede zwischen einer Bewertung kurz nach Studienabschluß und mit mindestens 5 Jahren Berufserfahrung. Es gab lediglich eine leichte Tendenz, daß das Studium mit größer werdendem zeitlichen Abstand positiver gesehen wird. Damit fanden sich hier zunächst keine Belege für eine starke Verzerrung der Einschätzungen, so daß von einer generellen Beeinträchtigung der Ergebnisse nicht gesprochen werden kann. Die Untersuchungsergebnisse können aber noch nicht verallgemeinert werden. Schomburg und Teichler sprechen sich in Ihrem Fazit allerdings auch nicht generell gegen eine Nutzung der Ergebnisse von Absolventeneinschätzungen zur Situation von Lehre und Studium aus, sondern betonen: „Absolventenstudien verweisen auf den Stellenwert von Lehre und Studium, und sie räumen mit den Absolventen denjenigen eine qualifizierte Stimme in der Evaluation ein, die die Stärken und Schwächen der Angebote und Bedingungen seitens der Hochschulen in besonders schicksalhafter Weise erfahren haben.“ (vgl. Teichler/ Schomburg 1997: 257).

Insgesamt können Absolventenbefragungen zur Einschätzung der Studiensituation trotz einiger aufgrund ihrer Retrospektivität vorzunehmenden Einschränkungen grundsätzlich als geeignet gelten. Während ihre Aussagekraft kurz nach Abschluß des Studiums durch den

erstmaligen möglichen Gesamtüberblick bei geringen Retrospektivitätseffekten als am höchsten eingeschätzt werden kann, sinkt sie mit größer werdendem zeitlichem Abstand (nach mehreren Jahren oder gar Jahrzehnten) durch Erinnerungsunschärfen oder Verklärungen.

Zum weiter oben angesprochenen zweiten Verwendungszweck von Absolventenbefragungen – zur Informationsgewinnung über den beruflichen Verbleib bzw. Erfolg der Absolventen – werden bezüglich der Datenqualität in den vorliegenden Studien zu Absolventenbefragungen keine grundsätzlichen Bedenken geäußert. Der berufliche Verbleib ist ein Bereich unmittelbaren Erlebens der Absolventen, für den offenbar erwartet wird, daß sie über die entsprechende Expertise verfügen. Geeignete Befragungsinstrumente und Forschungsdesigns vorausgesetzt, die den Anforderungen der empirischen Sozialforschung genügen, werden daher Absolventenbefragungen für den genannten Zweck als geeignet eingeschätzt. So formulieren z.B. Reinfeldt/ Frings (2003: 285) in einem Überblicksartikel zum Stand der Absolventenbefragungen resümierend zur Datenqualität der Ergebnisse zum beruflichen Verbleib bzw. Erfolg der Absolventen, es gebe „Ansätze wie die Kasseler Studien, die europäische Absolventenbefragung und die Schweizer Studien, deren Fragenkataloge zum Übergang vom Studium in den Beruf sich in vielen Bereichen überschneiden, bereits über einen längeren Zeitraum genutzt werden und als validiert gelten können“.

Einschränkungen, die zu nennen wären, sind lediglich die bei der Auswertung von Befragungen generell geltenden, die hier nicht ausführlicher diskutiert werden sollen. Als ein Beispiel ist die Beachtung von erhöhtem Item-Nonresponse bei einigen Fragen, die als sensibel gelten könnten, zu nennen. Insbesondere bei der Frage zum Einkommen tritt z.T. erhöhter Item-Nonresponse auf, der die Aussagekraft dieses Kriteriums einschränken könnte.

Bezüglich möglicher Retrospektivitätseffekte ist festzuhalten, daß diese lediglich auftreten können, wenn Absolventen erst viele Jahre nach ihrem Studienabschluß und (im Normalfall zügig darauffolgenden) Berufseinstieg zu ihren Erfahrungen beim Berufseinstieg bzw. den ersten Berufserfahrungen befragt würden. Bezüglich des *aktuellen* beruflichen Verbleibes bzw. Erfolges können von vornherein keine Retrospektivitätseffekte auftreten. Auch wenn der Berufsverlauf nur wenige Jahre rückblickend erfragt wird, dürften solche Effekte gering sein. So kommt Becker (2001: 52) anhand einer vergleichenden Reliabilitätsanalyse von Paneldaten zum Berufsverlauf zwischen der Privatwirtschaft und dem Öffentlichen Dienst zu dem Schluß: „Den empirischen Befunden zufolge kann insgesamt von einer hohen

Reliabilität retrospektiver Berufsverlaufsdaten und damit von einer guten Datenqualität solcher Längsschnittdaten ausgegangen werden.“ Lediglich für Beamte zeigte sich, daß diese „offensichtlich Schwierigkeiten haben, die einzelnen Tätigkeiten voneinander abzugrenzen. Allerdings gab es auch deutliche Hinweise, daß die unzuverlässigeren Angaben der Beamten erhebungstechnisch bedingt sind“ (vgl. Becker 2001: 52). Als Schlussfolgerung sieht Becker dann auch nur für Befragungen mit einer Schwerpunktsetzung auf Beamte den Bedarf einer Reorganisation der Erhebungsinstrumente.

Keine Einschränkungen der Datenqualität aus engerer methodischer Sicht, aber eine Anmerkung zur Begrenztheit der Aussagekraft von Absolventenbefragungen allgemein soll noch bezüglich ihrer Verwendbarkeit für die Weiterentwicklung der Studieninhalte vorgenommen werden. Auch wenn die Angaben der Absolventen zu ihren Tätigkeiten korrekt sind und mit größer werdendem zeitlichen Abstand bzw. zunehmender Berufserfahrung von Absolventen auch ein zunehmendes Urteilsvermögen bezüglich der Studieninhalte unterstellt wird, können sie natürlich nicht als Maß aller Dinge betrachtet werden. Absolventen gehen von ihrer subjektiven Situation aus und können nur genauso gut oder schlecht wie andere Menschen künftige Entwicklungen in ihrem Tätigkeitsgebiet abschätzen. Daher sollten Ergebnisse von Absolventenbefragungen bezüglich der Studieninhalte nur ein Kriterium unter mehreren darstellen. So kann die Forschungsausrichtung eines Faches ebenfalls eine wesentliche Rolle spielen. Ohnehin wird hier davon ausgegangen, daß sich Studienangebote nicht vorrangig am kurzfristigen (und teilweise stark schwankenden) Bedarf orientieren sollten, sondern die Entwicklung über einen längeren Zeitraum betrachtet werden sollte. Wenn Absolventenbefragungen wie häufig an den Hochschulen z.B. im Zusammenhang mit Fachevaluationen alle 3 bis 6 Jahre durchgeführt werden, könnten für langfristige Entscheidungen auch die Ergebnisse mehrerer Absolventenbefragungen herangezogen werden.

Wenn sich allerdings über einen längeren Zeitraum und über mehrere Kriterien hinweg gesehen zeigt, daß grundlegende Veränderungen stattgefunden haben, können Absolventenbefragungen durchaus wertvolle Anregungen zu konkreten Weiterentwicklungsmöglichkeiten enthalten. So wird im Bericht für 2003 im Rahmen des Berichtssystems zur technologischen Leistungsfähigkeit Deutschlands formuliert: „Es sollte allerdings erwartet werden, daß eine über Jahre erkennbare signifikante strukturelle Veränderung im Bedarf und

damit der Nachfrage nach Arbeitskräften auch zu entsprechenden Veränderungen der Produktionsbedingungen für Arbeitsqualifikation führt.“ (vgl. HIS 2003: 5).

### **5.5.3 Vergleichbarkeit der Ergebnisse von Absolventenbefragungen**

Ähnlich wie bei Studentenbefragungen gilt als grundlegende Voraussetzung für die Vergleichbarkeit der Ergebnisse von Absolventenbefragungen, daß sie frei sein sollen von "externen" Einflußfaktoren, die ihrerseits nicht als Bestandteil des zu erfassenden Inhaltes gelten können. Da beispielsweise die Abschlußnoten hier den Studienergebnissen zugeordnet und nicht als Kriterium beruflichen Erfolges angesehen werden, könnte ein Zusammenhang von Abschlußnoten mit dem beruflichen Erfolg (v.a. wenn die Verantwortlichen für die Studiengänge die Noten selbst bestimmen können) als potentiell verzerrender Einflussfaktor gesehen werden. Dies wäre auch mit Blick auf zu vermeidende unerwünschte Anpassungseffekte plausibel.

Ein anderes Problem für die Vergleichbarkeit der Ergebnisse von Absolventenbefragungen wäre, wenn die soziale Herkunft in starkem Maße mit den beruflichen Erfolgchancen zusammenhängt. In einem solchen Fall wäre z.B. eine Hochschule mit einem hohen Anteil an Studierenden mit niedrigerer sozialer Herkunft im Nachteil gegenüber einer Hochschule, die vor allem die Kinder „Bessergestellter“ ausbildet.<sup>200</sup>

Auch die Studienvoraussetzungen (gemessen z.B. an der Abiturnote) könnten für die Vergleichbarkeit problematisch sein, da die Situation für Studiengänge, die viele Studierende mit schlechteren Studienvoraussetzungen haben, schwieriger sein könnte. Auch ist darauf hinzuweisen, daß die beruflichen Erfolgchancen von Absolventen verschiedener Fächer höchst unterschiedlich sind und sich daher ein Vergleich von Absolutwerten verschiedener Fächer oder ganzer Hochschulen (mit unterschiedlicher Fächerstruktur) von vornherein nicht zweckmäßig ist. Für (hier nicht angesprochene) hochschulübergreifende Vergleiche über mehrere Bundesländer hinweg müßten außerdem noch Indikatoren für die regionale Arbeitsmarktlage (z.B. Arbeitslosenzahl von Absolventen des jeweiligen Faches, Vermittlungsquoten o.ä., länder- oder regionalspezifisch ausgewertete Statistiken der Bundesagentur für Arbeit einbezogen werden - vgl. z.B. Plicht/ Schreyer 2002: 531).

---

<sup>200</sup> Im Zusammenhang mit der Diskussion von Schulqualität und der Leistungsfähigkeit von Schulen wird der Einfluß der sozialen Herkunft auf die Schulleistungen häufig angesprochen. In Finnland werden die Schulleistungen daher nicht an einem absoluten Ziel, sondern unter Berücksichtigung eines „Erwartungswertes“ gemessen, der die soziale Herkunft der Schüler berücksichtigt (vgl. Kahl 2002). Es erscheint naheliegend, daß die soziale Herkunft auch im Hochschulbereich Einfluß auf die Leistungsfähigkeit hat.

Nachfolgend sollen die häufigsten genannten und einige weitere plausibel erscheinende potentielle Einflussfaktoren auf den beruflichen Erfolg systematisch analysiert werden. Eine Entscheidung, welche Aspekte als die Ergebnisse „verzerrend“ aufgefasst werden müssen und welche nicht, ist dabei nicht einheitlich möglich, da diese auch von den Zielen und Rahmenbedingungen der jeweiligen Ausbildung abhängt. In die Modelle wurden daher alle am häufigsten genannten Einflussfaktoren aufgenommen. So muß z.B. ein Zusammenhang des beruflichen Erfolges mit den Abschlussnoten in bestimmten Fächern mit Staatsprüfungen nicht gegen die Eignung von Absolventenbefragungsergebnissen sprechen, da die Staatsprüfungen für alle Hochschulen (eines Landes) gleich durchgeführt werden. Auch unterschiedliche Studienvoraussetzungen müssen nicht als verzerrender Einflussfaktor aufgefasst werden, wenn die Hochschulen ihre Studierenden selbst auswählen.

Grundsätzlich soll in Anlehnung an den Fairnessgedanken davon ausgegangen werden, daß alle Aspekte, auf die eine Hochschule keinen oder nur sehr begrenzten Einfluß hat, als potentiell verzerrender Einflussfaktor aufgefaßt werden kann.

Zusammenhänge des beruflichen Erfolges mit Aspekten, auf die die Hochschule Einfluß hat, würden dagegen für die Eignung der Erfolgskriterien sprechen. Nachfolgend sollen die Ergebnisse von Analysen zu Einflussfaktoren auf den beruflichen Erfolg exemplarisch für die Absolventen der TU Dresden vorgestellt werden.

### **Analysen zu Einflussfaktoren anhand der Absolventenbefragungen an der TU Dresden**

Seit dem Jahr 2000 wurden am Lehrstuhl für Mikrosoziologie der TU Dresden Absolventenbefragungen im Auftrag der Universitätsleitung durchgeführt. Der Fragebogen umfasste die Themenkomplexe Studium, Tätigkeitsverlauf, Weiterqualifikationen, Stellensuche, Berufsstart, Kompetenzen, Anforderungen und Zufriedenheit, berufliche Pläne und Orientierungen sowie Angaben zur Person und zu den Eltern. Außerdem wurde festgelegt, die Absolventenstudien an der TU Dresden als Panelstudien bzw. Langzeituntersuchungen durchzuführen. Die befragten Absolventen sollen im Abstand von drei bis vier Jahren nochmals befragt werden, um über einen möglichst langen Zeitraum den beruflichen Integrationsprozeß in Erfahrung zu bringen. Dieses Forschungsdesign wurde seit dem Jahr 2000 an allen Fakultäten der TU Dresden eingesetzt. Die Rücklaufquoten der Befragungen waren mit ca. 50 Prozent sehr zufriedenstellend, so daß sie den Vergleich zu bundesweiten Absolventenstudien (30 bis 40 Prozent Rücklaufquote) keineswegs zu scheuen brauchen

(vgl. HIS 2000, Teichler/ Schomburg 2001). Auch ein Vergleich von Merkmalen der Befragtenpopulation mit der Grundgesamtheit bezüglich Jahrgangsstärken, Studiendauer, Abschlußnoten, Alter und Geschlecht ergab nur sehr wenige signifikante Abweichungen in den einzelnen Fachrichtungen/ Fakultäten. Eine detailliertere Erläuterung des methodischen Vorgehens findet sich in den jeweiligen Abschnitten zur Anlage und Durchführung der Studie in den Abschlußberichten der Dresdner Absolventenstudien (vgl. Krempkow 2001 u. Folgejahre).

Mit den Analysen zu Einflussfaktoren wurde an vorhandene Studien angeknüpft. Ähnliche Analysen existieren beispielsweise mit denen von Brüderl (1996), Teichler/ Schomburg (2001) sowie für Promovierte von Enders/ Bornmann (2002). Angelehnt an diese soll hier eine umfassende Analyse solcher potentieller Einflussfaktoren auf den beruflichen Erfolg anhand der Dresdner Absolventenstudien der Jahre 2000 bis 2003 zusammenfassend vorgestellt werden. Dabei wurde ein Hochschulgesamtdatensatz mit Daten von 1957 Absolventen verwendet und wohl erstmals in Deutschland alle Fakultäten einer großen Universität in eine derartige Untersuchung einbezogen. Neben Korrelationsanalysen (die zuvor auch auf Fachrichtungsebene durchgeführt wurden, was hier nicht umfassend dargestellt werden kann) wurde v.a. multiple Regressionsanalysen durchgeführt (vgl. Krempkow/ Pastohr 2004). Als Erfolgskriterien gelten hierbei nicht nur die Suchdauer beim Berufseinstieg, das Bruttomonatseinkommen (zu verschiedenen Zeitpunkten) und die berufliche Zufriedenheit, sondern auch, inwieweit die Tätigkeit zur Ausbildung adäquat ist (Ausbildungsadäquanz). Dies sind Erfolgskriterien, wie sie häufig in Absolventenstudien herangezogen wurden, wobei z.T. nicht alle, sondern nur einzelne der hier genannten Kriterien verwendet wurden (vgl. z.B. Rostampour/ Lemberg 2003, Brüderl/ Reimer 2002, Meinefeld 2002, Teichler/ Schomburg 2001).

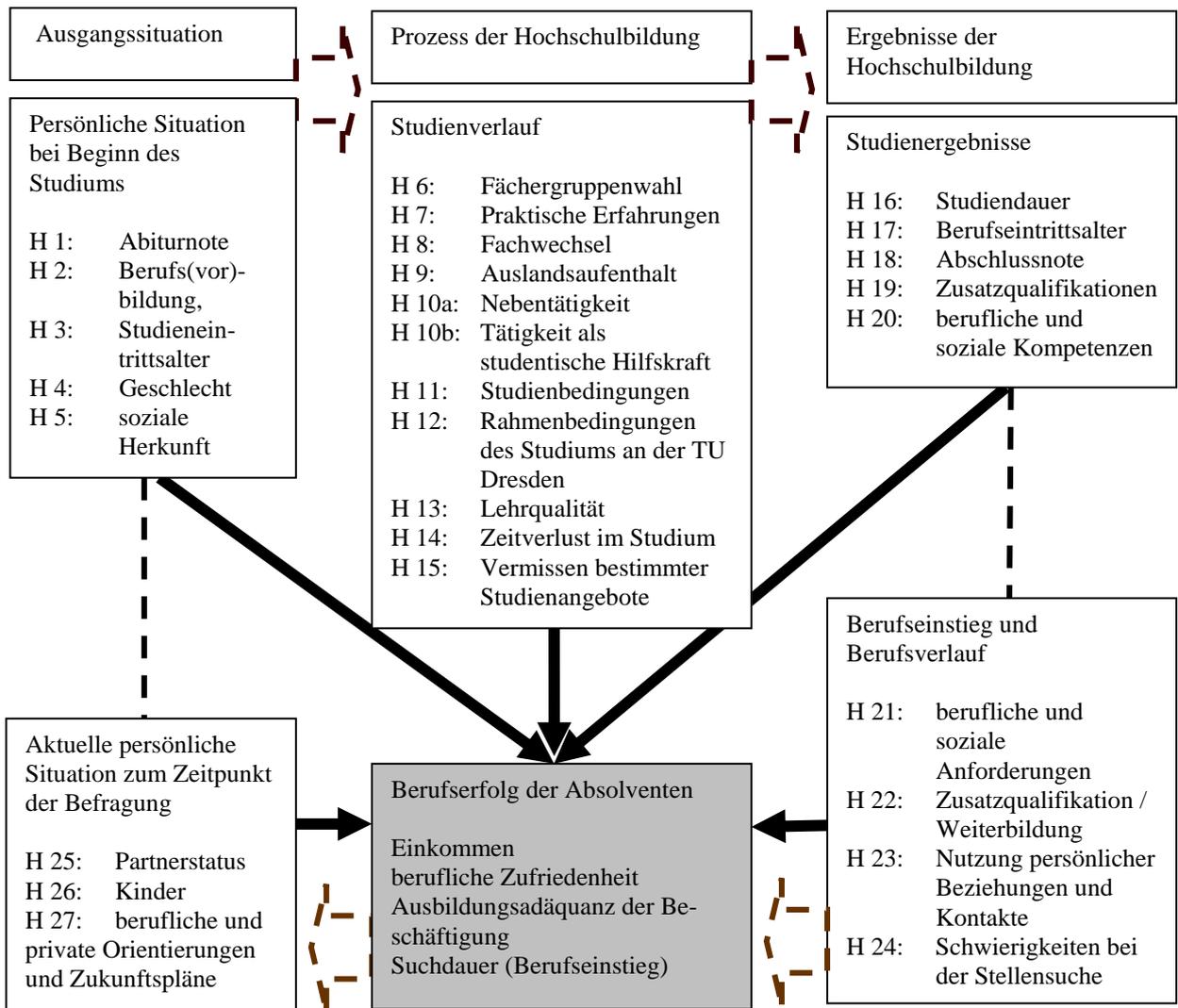
In nachfolgender Abbildung werden die Beziehungen in einem umfassenden Modell graphisch dargestellt. Dabei wurden die einzelnen Variablen der Ausgangssituation (Input), Prozess oder Ergebnis (Output) in der Hochschulbildung sowie einzelnen Etappen der Berufslaufbahn zugeordnet. Die Gliederung entspricht in ihrer Grundstruktur im wesentlichen auch dem Dimensionenmodell von Reinfeld/ Frings (2003: 286), welches wiederum eine modifizierte Fassung des Kasseler Modells darstellt. Wie auch bei den Kriterien beruflichen Erfolges verbergen sich hinter den Variablen, die hier als potentielle Einflussfaktoren beruflichen Erfolges angeführt sind, z.T. Subvariablen, also untergeordnete Items. Die in die Untersuchung einbezogenen Variablen wurden zuvor bereits Faktoren-

analysen unterzogen (vgl. Krempkow/ Popp 2003 bzw. Krempkow u.a. 2000 – 2002). Hierbei zeigte sich jedoch, daß eine Dimensionenreduktion nur in wenigen Fällen ohne größere Verluste der Aussagekraft möglich war. Teilweise stellen die einbezogenen Variablen auch bereits die mit Hilfe von Faktorenanalysen deutlich reduzierten Itemlisten dar (z.B. berufliche und private Orientierungen).

---

<sup>215</sup> Zur Klärung der Zusammenhänge wurden die jeweils relevanten Koeffizienten berechnet und in der ausführlicheren Publikation ausgewiesen (vgl. Krempkow/ Pastohr 2004). Meist war dies der Rangkorrelationskoeffizient nach Spearman. Untersuchungen mittels Kolmogorov-Smirnov-Tests und trendbereinigter Q-Q-Diagramme zeigten, daß die Einstiegseinkommen der Absolventen mehrerer Fakultäten bzw. Fachrichtungen nicht normalverteilt sind, so daß für diese die Berechnung des Korrelationskoeffizienten nach Pearson nicht zulässig ist (vgl. Brosius 2002: 377ff., 813ff.).

## Übersicht zu potentiellen Einflußfaktoren beruflichen Erfolges



Quelle: eigene Darstellung, angelehnt an Krempkow/ Pastohr 2004

Bei der Datenanalyse wurde in mehreren Schritten vorgegangen. Zunächst wurden mittels Korrelationsanalysen die Stärke (und gegebenenfalls die Richtung) des wechselseitigen Zusammenhangs zweier Variablen auf Hochschul- und Fakultäts- bzw. Fachrichtungsebene ermittelt.<sup>215</sup> Nur die unabhängigen Variablen, bei denen fächerübergreifend ein sehr signifikanter Zusammenhang mit der jeweiligen abhängigen Variablen besteht, wurden in die Analyse einbezogen. Dies rechtfertigt sich dadurch, daß der primäre Fokus dieser Untersuchung auf der Hochschulebene lag und hier sehr zuverlässige Zusammenhänge als erforderlich angenommen werden. Da diese allerdings durch Einflüsse anderer Variablen verstärkt oder unterdrückt und somit systematisch Zusammenhänge verborgen bleiben können, wurden schließlich auch signifikante Korrelationen, die nicht bei allen, aber mindestens bei zwei Fakultäten bzw. Fachrichtungen auftreten, berücksichtigt.

In einem zweiten Schritt werden die (möglicherweise intervenierenden) unabhängigen Variablen zur Bestimmung der abhängigen Variable in eine multiple Regressionsanalyse einbezogen. In der folgenden Tabelle sollen nun die signifikanten Zusammenhänge zwischen ausgewählten Kriterien beruflichen Erfolges (abhängige Variablen) und den als Einflussfaktoren auf den beruflichen Erfolg zu untersuchenden (unabhängigen) Variablen entsprechend den Ergebnissen der auf Hochschulebene durchgeführten Regressionsanalysen dargestellt werden. Hierbei wurde zur Wahrung der Übersichtlichkeit das Einkommen lediglich zu einem Zeitpunkt (12 Monate nach Studienabschluß) ausgewiesen.<sup>216</sup> Die Analysen wurden für mehrere Zeitpunkte durchgeführt, es ergaben sich aber relativ wenige Unterschiede, so daß dies vertretbar erscheint. Außerdem wurden Zukunftspläne der Absolventen, die ebenfalls erfragt und in die Analyse einbezogen wurden, hier nicht als potentielle Einflußfaktoren aufgefasst. Es wird davon ausgegangen, daß z.B. der Wunsch nach beruflichen Veränderungen aus der beruflichen Situation resultiert und nicht umgekehrt und die Zukunftspläne der Absolventen daher kaum einen Beitrag zur Erklärung des beruflichen Erfolges leisten können.

Nachfolgend erfolgen noch einige kurze Erläuterungen, die das Verständnis der Angaben in der Tabelle erleichtern sollen. So bezeichnet die Angabe „korr.  $R^2$ “ unter dem jeweiligen Erfolgskriterium im Tabellenkopf das „Bestimmtheitsmaß“ des Regressionsmodells für das jeweilige Kriterium des beruflichen Erfolges (vgl. Brosius 2002: 531ff.). Es misst den Anteil der erklärten Varianz an der gesamten Streuung, der grundsätzlich zwischen 0 und 1 liegen kann. Beispielsweise für das Einkommen (nach einem Jahr) beträgt  $R^2=.25$ <sup>217</sup> oder für die Ausbildungsadäquanz  $R^2=.43$ . Nach Backhaus/ Erichsen u.a. (2000: 24) berücksichtigt das korrigierte  $R^2$  im Gegensatz zum  $R^2$  auch die Anzahl der einbezogenen Variablen.

---

<sup>216</sup> Die Untersuchung des Einkommens wurde in der zugrundeliegenden Analyse zu insgesamt 3 Zeitpunkten vorgenommen: Einstiegseinkommen, Einkommen 12 Monate und 24 Monate nach Studienabschluß. Die anderen Erfolgskriterien berufliche Zufriedenheit und Ausbildungsadäquanz spiegeln die subjektiven Einschätzungen der Absolventen zum Befragungszeitpunkt wider. Die Suchdauer bis zur 1. Stelle nimmt als Erfolgskriterium für den Berufseinstieg ohnehin eine Sonderstellung ein, die Angaben der Absolventen beziehen sich hier aber auch auf Ihre Einschätzung zum Zeitpunkt der Befragung. Absolventen, die noch keine Stelle gefunden hatten, wurden bei der Berechnung der Suchdauer bis zur ersten Stelle nicht berücksichtigt.

<sup>217</sup> In einer ähnlichen Analyse zum Einkommen der bezüglich ihrer Anstellungsverhältnisse weniger heterogenen Gruppe von Promovierten wurden in verschiedenen Modellen 18 bzw. 21% erklärte Varianz erzielt (vgl. Enders/ Bornmann 2001: 188). Eine Langzeituntersuchung ehemaliger Gymnasiasten ergab 17% Erklärungskraft (vgl. Meulemann 1995: 258).

**Übersicht zu potentiellen Einflussfaktoren auf ausgewählte Kriterien beruflichen Erfolges bzw. des Berufseinstiegs (mit Betrag der Beta-Koeffizienten)**

	Potentielle Einflußfaktoren auf den beruflichen Erfolg	Einkommen (nach 1 Jahr) (korr.R <sup>2</sup> = .25)	berufliche Zufriedenheit (korr.R <sup>2</sup> = .28)	Ausbildungsadäquanz (korr.R <sup>2</sup> = .43)	Suchdauer bis zur 1. Stelle (korr.R <sup>2</sup> = .12)
psl. Situation bei Studienbeginn	<b>H 1: Abiturnote</b>	-	-	-	-
	<b>H 2: Berufs(vor)bildung</b> (binär <sup>218</sup> )	-	-	-	-
	<b>H 3: Studieneintrittsalter</b>	-	-	-	-
	<b>H 4: Geschlecht:</b> (männlich) binär	-	-	.06	-
	<b>H 5: soziale Herkunft</b>	-	.12	.05	-
Studienverlauf	<b>H 6: Fächergruppenwahl</b> Ingenieur- / Mathematisch- / Naturwissensch. Ausrichtung	.11	-	-	-
	<b>H 7: praktische Erfahrungen</b> praktische Tätigkeiten (Praktika und/ oder Berufsausbildung - binär)	.09	-	.06	-
	<b>H 8: Fachwechsel</b>	-	-	.06	-
	<b>H 9: Auslandsaufenthalt</b>	-	-	-	-
	<b>H 10: Nebentätigkeit (SHK)</b>	-	-	.04	-
	<b>H 11: Studienbedingungen im Fach:</b> Aufbau und Struktur des Studienganges	-	-	.07	-
	Breite des Lehrangebots	-	-	.07	-
	fachliche Spezialisierungs- und Vertiefungsmöglichkeiten	.13	.07	-	-
	zeitliche Koordination des Lehrangebotes	.07	-	.05	-
	Praxisbezug der Lehre	-	-	.28	-
	ausreichende Anzahl von Plätzen in Lehrveranstaltungen	-	-	.05	-
	<b>H 12: Rahmenbedingungen des Studiums an der TU Dresden:</b> Angebot an Mensen und Cafeterias	-	-	.07	-
	Wohnsituation in Dresden	-	-	.06	-
<b>H 13: Lehrqualität im Fach:</b> Einübung in selbständiges Lernen/ Arbeiten	-	.05	.06	-	
Vermittlung von Fachwissen	-	-	.10	-	
Studienergebnisse	<b>H 14: Zeitverluste im Studium aufgrund organisatorischer Regelungen</b>	-	-	-	-
	<b>H 15: bestimmte Studienangebote vermisst?</b>	-	-	-	-

<sup>218</sup> Hier bedeutet „binär“, daß eine Variable für die Analyse in die Werte „0“ und „1“ umkodiert bzw. dichotomisiert wurde, um sie in die Regressionsanalyse einbeziehen zu können (vgl. Brosius 2002: 547).

	Potentielle Einflußfaktoren auf den beruflichen Erfolg	Einkommen (nach 1 Jahr) (korr.R <sup>2</sup> = .25)	berufliche Zufriedenheit (korr.R <sup>2</sup> = .28)	Ausbildungsadäquanz (korr.R <sup>2</sup> = .43)	Suchdauer bis zur 1. Stelle (korr.R <sup>2</sup> = .12)
Berufseinstieg und Berufsverlauf	<b>H 16: Studiendauer:</b> Fachsemester	-	-	-	-
	Hochschulsemester	-	-	-	.11
	<b>H 17: Berufseintrittsalter</b>	.14	-	-	-
	<b>H 18: Abschlussnote</b>	-	.06	-	-
	<b>H 19: Zusatzqualifikationen:</b> Englisch	.05	.05	-	-
	Textverarbeitung	-	.06	-	-
	Internetnutzung	-	-	-	.13
	Programmiersprachen	.23	.06	-	-
	<b>H 20: berufliche und soziale Kompetenzen:</b> Kreativität	-	.08	-	-
	Lernfähigkeit	-	-	.12	-
	Zeiteinteilung	-	-	.06	-
	Fachkenntnisse	-	-	.07	-
	wirtschaftliches Denken	-	-	-	.12
	<b>H 21: berufliche und soziale Anforderungen:</b> Fachkenntnisse	.08	-	.13	-
Kreativität	-	-	.04	-	
Fremdsprachen	.06	.06	-	-	
wirtschaftliches Denken	.11	-	-	-	
Teamarbeit	.10	.05	-	-	
Kritisches Denken	-	.08	-	-	
Verhandlungsgeschick	-	.05	-	-	
Anpassungsfähigkeit	-	.07	.06	.09	
<b>H 22: Zusatzqualifikation / Weiterbildung:</b> Abschluss einer Promotion	-	-	.05	-	
Besuch von Weiterbildungsveranstaltung(en): methodische Kompetenzen	-	.08	-	-	
betriebswirtschaftliches Wissen	-	.09	-	-	
<b>H 23: Nutzung persönlicher Beziehungen und Kontakte:</b> Nutzung der Kontakte während eines Jobs/ Praktikums im Studium	-	-	.04	-	
<b>H 24: Schwierigkeiten bei der Stellensuche</b> keine Probleme	-	.09	.11	.13	
anderer Studienschwerpunkt gesucht	-	.08	-	-	
fehlende Spezialkenntnisse	-	.08	-	.13	
Aktuelle psl. Situation	<b>H 25: Partnerstatus</b> (mit/ ohne)	-	-	-	-
	<b>H 26: Kinder</b> (Vorhandensein)	-	-	-	-
	<b>H 27: berufliche und private Orientierungen</b> hohes Einkommen und Karrierechancen	.12	-	-	-

Quelle: Krempkow/ Pastohr 2004

Einflussfaktoren, die als Biasvariablen aufgefasst werden können, sind vor allem in der Ausgangssituation zu suchen. Diese können im Gegensatz zu Studienverlauf und Studienergebnissen derzeit kaum durch die Hochschule (mit)beeinflusst werden. Darüber hinaus kann aber auch die aktuelle persönliche Situation der Absolventen, z.B. das Vorhandensein von Kindern oder eine Mobilitätseinschränkung durch eine/n feste/n Partner/in, den beruflichen Erfolg beeinflussen, ohne daß die Hochschule hierauf Einfluß hätte. Der Berufseinstieg und z.T. der Berufsverlauf bei Weiterqualifikationen kann bereits durch entsprechende Angebote der Unterstützung von Absolventen zur Verringerung von Schwierigkeiten bei der Stellensuche, wie sie inzwischen an vielen Hochschulen existieren, beeinflusst werden, wenngleich hier der Einfluß sicherlich nach wie vor begrenzt ist.

Betrachtet man zunächst die erfassten Aspekte der Ausgangssituation, so zeigt sich, daß weder unterschiedliche Vorkenntnisse (gemessen an Abiturnoten oder vorhergehenden Berufsausbildungen), noch das Studieneintrittsalter einen Einfluß auf die untersuchten Kriterien des beruflichen Erfolges haben. Lediglich zum Geschlecht und zur sozialen Herkunft finden sich Zusammenhänge mit subjektiven Indikatoren des beruflichen Erfolges. Auch in anderen Studien zeigte sich z.T. ein deutlicher Einfluß des Geschlechts und der sozialen Herkunft (vgl. z.B. Enders/ Bornmann 2001: 188 ff.). Daß es im Gegensatz zu anderen Studien in den eigenen Untersuchungen keinen Zusammenhang von Geschlecht und Einkommen (als objektives Kriterium beruflichen Erfolges) gibt, liegt sicherlich daran, daß dies sich bei Untersuchung der Einkommen 12 Monate nach Studienabschluß noch nicht niederschlagen kann, weil diese Effekte erfahrungsgemäß verstärkt erst mit Eintritt in die Familienphase auftreten (vgl. Kerst/ Minks 2004: 13, Rostampour/ Lemberg 2003: 168ff., Teichler/ Schomburg 2001: 179). Hierzu wäre die Einkommensentwicklung über einen längeren Zeitraum zu betrachten, wie es im Rahmen der Dresdner Absolventenstudien mittels Wiederholungsbefragungen im Abstand von einigen Jahren auch für alle Fächer geplant ist. Zeigen sich hierbei deutliche Einkommensunterschiede und betrachtet man das Einkommen als Indikator für den Erfolg der Hochschulausbildung, so würden Hochschulen bzw. Studiengänge mit einem höheren Frauenanteil benachteiligt. Ein ähnliches Problem zeigt sich bei Betrachtung des Zusammenhanges der sozialen Herkunft (hier: Bildungsherkunft) mit der beruflichen Zufriedenheit. Soll die berufliche Zufriedenheit als Indikator des Erfolgs der Hochschulausbildung gelten, so wären Hochschulen bzw. Studiengänge mit einem höheren Anteil von Absolventen aus bildungsfernen Schichten benachteiligt. Daher

sollten sowohl Frauenanteil als auch soziale Herkunft bei Vergleichen des beruflichen Erfolges von Absolventen berücksichtigt werden.

In ähnlicher Weise gilt dies auch für die aktuelle persönliche Situation der Absolventen: Die Untersuchungen konnten zwar keine Einflüsse von Kindern oder Partnerstatus belegen. Es ist aber mit zunehmenden zeitlichem Abstand vom Studienabschluß eine stärkere Beeinflussung des beruflichen Erfolges der Absolventen zu erwarten (vgl. z.B. Rostampour/Lembert 2003: 167ff., Hartmann/ Kopp 2001: 449). Der einzige Zusammenhang, der hier belegt werden konnte, war der von Karriereorientierung der Absolventen und dem erzielten Einkommen. Dies deutet darauf hin, daß stärker in eine bestimmte Richtung motivierte Absolventen auch tatsächlich tendenziell größere Erfolge in dieser Richtung erzielen. Setzt man voraus, daß ein bestimmter Anteil von Studierenden vom Beginn des Studiums an eine stärkere Karriereorientierung hat und geht man zusätzlich davon aus, daß eine solche Orientierung durch die Hochschule nicht beeinflusst werden kann, müsste man dies als einen die Ergebnisse verzerrenden Einflussfaktor auffassen. Häufig wird jedoch gerade das Gegenteil beklagt, daß nämlich die Studierenden am Anfang des Studiums meist keine klare Orientierung haben. Sollte dies doch einmal der Fall sein, wird es zumindest z.T. der guten Studienberatung zugeschrieben. Außerdem haben inzwischen die Hochschulen die Möglichkeit, ihre Studierenden zu einem großen Teil selbst auszuwählen und könnten dabei durch Berücksichtigung der Motivation der Studierenden künftig stärker Einfluß auf den Anteil von Studierenden mit einer bestimmten Orientierung im jeweiligen Studiengang nehmen. Dennoch wäre eine Betrachtung des Zusammenhangs von Karriereorientierung und Einkommen als verzerrender Einflussfaktor nicht von der Hand zu weisen, wenn z.B. ein Studiengang aufgrund der regional oder in diesem Fach schwierigen Bewerbersituation de facto keine Möglichkeit zur Bewerberauswahl hat, weil dies auch durch offensive Studentenwerbung nicht immer zu verändern ist. Außerdem müssen die ausgewählten Studierenden entsprechend den Zielen eines Faches nicht die mit Karriereorientierung sein. Unter solchen Bedingungen erscheint es durchaus sinnvoll und notwendig, auch die berufliche Orientierung der Absolventen bei Vergleichen zu berücksichtigen.

Ähnlich diffizil ist die Frage zu entscheiden, inwieweit die Einbeziehung von Aspekten des Berufseinstieges bzw. Berufsverlaufes bei Vergleichen der Ergebnisse zum Berufserfolg von Absolventen sinnvoll oder geboten ist. Bei den Wegen der Stellenfindung und bei der Überwindung von Schwierigkeiten bei der Stellensuche kann die Hochschule durchaus Hilfestellung geben, wie verschiedene Initiativen an den Hochschulen zeigen (vgl. Teichler

2002: 11ff.). Auch auf die *vorhandenen* beruflichen und sozialen Kompetenzen zum Zeitpunkt des Studienabschlusses haben die Hochschulen grundsätzlich Einfluß.

Dagegen haben sie kaum Einfluß auf die *geforderten* beruflichen und sozialen Kompetenzen (hier zur besseren Abgrenzung als Anforderungen bezeichnet), die von den Arbeitgebern an die Absolventen gestellt werden. Andererseits wird vielerorts an den Hochschulen diskutiert, daß ein Hochschulstudium sehr wohl nicht nur fachliche, sondern auch fachübergreifende bzw. fachunabhängige Kompetenzen vermitteln soll, zu denen viele der geforderten Kompetenzen zählen, die Einfluß auf den beruflichen Erfolg haben. Zu nennen wären hier beispielsweise kritisches Denken, Teamfähigkeit, Kreativität oder Anpassungsfähigkeit, die häufig auch unter dem Oberbegriff soziale Kompetenzen zusammengefaßt werden. Bei der Diskussion um fachübergreifende bzw. fachunabhängige Kompetenzen wird häufig davon ausgegangen, daß Absolventen nur Gelegenheit bekommen, sich im Beruf höheren Anforderungen an bestimmte Kompetenzen zu stellen, wenn sie bereits über relativ gut ausgeprägte Kompetenzen verfügen und daher die Differenz zwischen bei Studienabschluß vorhandener und im Beruf geforderter Ausprägung dieser Kompetenzen möglichst gering sein sollte. Wenn dies aber ein Ziel der Hochschulausbildung ist und hierfür die mittels Absolventenbefragungen erhobenen Anforderungen eine Grundlage sein sollen, können die Anforderungen nicht zugleich als verzerrender Einflussfaktor betrachtet werden.<sup>219</sup> Daher können auch die Anforderungen an berufliche und soziale Kompetenzen nicht generell als Biasvariable aufgefasst werden. Wenn im Einzelfall jedoch Einigkeit darüber herrscht, daß ein bestimmter Aspekt des Berufseinstiegs nicht durch die Hochschule oder den Studiengang beeinflusst werden kann, ist es sicher sinnvoll und geboten, diesen Aspekt beim Vergleich von Ergebnissen zum beruflichen Erfolg von Absolventen zu berücksichtigen.

Zusammenfassend kann formuliert werden, daß Ergebnisse von Absolventenbefragungen grundsätzlich vergleichbar sind. Hierbei sollten jedoch Aspekte berücksichtigt werden, die

---

<sup>219</sup> Der Ausprägung von fachlichen und fachunabhängigen Kompetenzen bzw. ihrer „Passung“ zu den beruflichen Anforderungen wird von mehreren Autoren eine steigende Bedeutung beigemessen. Dabei wird davon ausgegangen, daß künftig angesichts schnell veraltender Wissensbestände tendenziell weniger konkretes Fachwissen gefordert ist, sondern stärker z.B. die Kompetenz, sich schnell und kompetent in neue Inhalte einzuarbeiten. Bereits jetzt ließ sich exemplarisch für die Absolventen der TU Dresden zeigen, daß Absolventen sowohl erfolgreicher sind, wenn sie zum Studienabschluß stärker über bestimmte Kompetenzen verfügen, als auch, wenn sie in Tätigkeiten arbeiten, in denen an sie bezüglich bestimmter Kompetenzen höhere Anforderungen gestellt werden (vgl. Krempkow 2004). An einigen Hochschulen wurden hierzu neue Lehrformen ausgearbeitet (z.B. Projektstudium oder das sogenannte problemorientierte Lernen), die

eine Hochschule nicht zu beeinflussen in der Lage ist. So kann insbesondere die Ausgangssituation der Absolventen schon am Beginn ihres Studiums bezüglich sozialer Herkunft und Geschlecht, die über das gesamte Studium hinweg bis zum beruflichen Verbleib wirken und auch Einfluß auf den beruflichen Erfolg haben. Aber auch die aktuelle persönliche Situation der Absolventen zum Befragungszeitpunkt sollte grundsätzlich berücksichtigt werden. Zwar konnte relativ kurze Zeit (1 bis 2 Jahre) nach Studienabschluß in den hier durchgeführten Regressionsanalysen kein Einfluß z.B. durch das Vorhandensein von Kindern gezeigt werden. Ein solcher wäre jedoch mit größerem zeitlichen Abstand zum Studium oder einem höheren Anteil von Absolventen mit Kindern zu erwarten. Lediglich für einen Aspekt der beruflichen Orientierung (Karriereorientierung) konnte ein Einfluß auf das Einkommen aufgezeigt werden. Allerdings ist es für die Hochschulen möglich, diesen Aspekt zu beeinflussen, so daß er nur unter bestimmten Bedingungen als verzerrender Einflussfaktor aufgefasst werden kann. Aspekte aus anderen Dimensionen von Absolventenbefragungen, wie Studienverlauf, Studienergebnisse und z.T. auch der Berufseinstieg, können durch die Hochschule bzw. den Studiengang grundsätzlich in stärkerem Maße beeinflusst werden (als die Ausgangssituation oder die persönliche Situation der Absolventen zum Zeitpunkt der Befragung) und stellen daher keine potentiell verzerrenden Einflußfaktoren dar, die die Vergleichbarkeit beeinträchtigen könnten. Ergebnisse von Absolventenbefragungen zum beruflichen Verbleib bzw. Erfolg sind daher unter Berücksichtigung der Ausgangsbedingungen bei Studienbeginn und der persönlichen Situation der Absolventen zum Befragungszeitpunkt vergleichbar.

#### **5.5.4 Weitere Anforderungen**

Als weitere Anforderungen an objektive wie auch an subjektive Indikatoren wurden Zeitnähe, sinnvolle Anreizwirkung und Vermeidung unerwünschter Anpassungseffekte sowie ein vernünftiges Verhältnis von Erhebungsaufwand und erwartetem Nutzen genannt.

Zur Anforderung an die Zeitnähe wurde bereits angesprochen, daß diese bei retrospektiven Bewertungen der Situation von Lehre und Studium nicht gegeben ist, bei der Einschätzung des beruflichen Verbleibes und Erfolges aber sehr wohl.

Als eine weitere Anforderung an Ergebnisse von Absolventenbefragungen soll hier zusätzlich noch einmal formuliert werden, was eigentlich bei allen Befragungen gilt, aber

---

tatsächlich aus Sicht der Absolventen in Hinblick auf die Ausbildung einiger solcher Kompetenzen besser

bei Absolventenbefragungen aufgrund der z.T. geringen Anzahl von Absolventen eine besondere Bedeutung erhält. So kann eine Vollerhebung eines Absolventenjahrganges eines bestimmten Faches bereits mit einer Handvoll Absolventen geschehen. Diese Fallzahl genügt sicherlich für einen vollständigen Überblick über den beruflichen Verbleib, da die Anforderung an Repräsentativität bei Übereinstimmung aller Merkmale der Befragten per definitionem mit denen der Grundgesamtheit gegeben ist. Häufig kann eine Vollerhebung jedoch nicht vollständig verwirklicht werden. Die Zuverlässigkeit der Ergebnisse und Resistenz gegenüber Schwankungen, die nichts mit Veränderungen der Qualität zu tun haben müssen, ist bei so geringen Fallzahlen jedoch selbst bei hohen Rücklaufquoten nicht gegeben, da bereits der Irrtum, ein Eingabefehler oder gar ein bewusster Beeinflussungsversuch durch extrem positive oder extrem negative Antworten bei einem einzelnen Absolventen die Ergebnisse beträchtlich verschieben könnten. Wenn man z.B. den Anteil regulär erwerbstätiger Absolventen oder Einkommensklassen in Prozent darstellen und ggf. zur Einordnung mit ähnlichen Fächern derselben Hochschule oder mit demselben Fach an einer anderen Hochschule oder dem Bundesdurchschnitt in diesem Fach vergleichen möchte, empfiehlt sich eine Fallzahl von wenigstens 50, besser 100<sup>220</sup> Absolventen, die z.T. erst durch Zusammenfassen mehrerer Absolventenjahrgänge erzielt werden kann (vgl. Teichler 2002: 15). Solche Vergleiche werden bei Absolventenstudien als besonders wichtig eingeschätzt, da „bei Absolventenstudien – anders als beispielsweise bei Evaluationen in Forschung, Studium und Lehre, bei denen in der Regel eine vergleichende Perspektive in Form eines peer reviews durch Gutachter hergestellt wird - für die angemessene Interpretation häufig der kontrastierende Hintergrund fehlt“ (vgl. Reinfeld/Frings 2003).

Der Anforderung, sinnvolle Anreize zu schaffen, ohne unerwünschte Anpassungseffekte zu erzielen, kann durch Absolventenbefragungen sehr gut entsprochen werden, da sie häufig als einzige Möglichkeit gesehen werden, nicht nur die Quantität bzw. „Stückzahl“ der Absolventen zu erfassen, sondern auch deren Qualität. Eine hohe Qualität der Ausbildung ist nicht nur gesetzliches Ziel der Hochschulen. Eine durch evtl. Fehlanreizwirkungen deutlich sinkende Ausbildungsqualität hätte unabsehbare Folgen für die Volkswirtschaft, da „der

---

bewertet werden (vgl. auch Schaeper/ Briedis 2004).

<sup>220</sup> Diese Fallzahlen entsprechen dem optimalen Stichprobenumfang von Signifikanztests für Mittelwertdifferenzen bei mittleren Effektstärken (n=50), Korrelationen bei mittleren Effektstärken (n=64) bzw. für Häufigkeitsdifferenzen bei einem Freiheitsgrad (n=87) bzw. zwei Freiheitsgraden (n=107), ebenfalls bei mittleren Effektstärken (vgl. Bortz/ Döring 2002: 613).

Ausbildung an Universitäten und Hochschulen eine entscheidende Bedeutung für die zukünftigen Innovationspotenziale der deutschen Wirtschaft zukommt“ (vgl. HIS 2003: 5). Solche Fehlanreizwirkungen gilt es daher in gesamtgesellschaftlichem Interesse zu vermeiden.

Vor diesem Hintergrund sollte auch die Abwägung von Erhebungsaufwand und erwartetem Nutzen geschehen. Zwar ist der Erhebungsaufwand trotz der Versuche, einen größeren Anteil der Befragungen über das Internet abzuwickeln, wegen der derzeit zum Erreichen von Repräsentativität<sup>221</sup> häufig noch ergänzend notwendigen postalischen Befragung und der damit anfallenden Portokosten relativ groß. Im Vergleich zu anderen Befragungsverfahren wie face-to-face und telefonischen Befragungen sind sie jedoch kostengünstiger und ebenso brauchbar (vgl. Reuband 2001: 329, 2000: 201ff.). Bei Einbeziehung der vorliegenden Erfahrungen zur Durchführung solcher Befragungen können sowohl bei postalischen Bevölkerungsumfragen (50% bis 78% - vgl. Reuband 2001: 308) als auch bei postalischen Absolventenbefragungen (durchschnittlich 50% - vgl. Krempkow u.a. 2000, 2001, 2002, 2003) zumindest gleichhohe Teilnahmeraten erzielt werden wie bei face-to-face und telefonischen Befragungen (45% bis knapp über 50% - vgl. Reuband 2001: 308).

Kleinere Hochschulen oder Studiengänge dürften einerseits vom finanziellen Erhebungsaufwand, andererseits aber auch bei der notwendigen Qualifikation des Personals schnell überfordert sein. Für sie könnte eine finanzielle und/ oder institutionelle Unterstützung seitens des Landes oder bei größeren Hochschulen seitens der Hochschulleitung (für die einzelnen Studiengänge) erfolgen. Auf diese Weise kann der finanzielle und organisatorische Aufwand einer Erhebung von für die Erfassung der Qualität notwendigen Aspekten z.B. mittels Absolventenbefragungen „auf viele Schultern verteilt“, erhebliche Portorabatte genutzt und damit letztlich auch für kleinere Hochschulen/ Studiengänge tragbar werden. Auf längere Sicht dürfte es möglich sein, Absolventenbefragungen (zumindest zu größeren Anteilen) mittels Onlinebefragungen über das Internet abzuwickeln und damit den Erhebungsaufwand deutlich zu senken. Bei der Konzeption und den Auswertungen solcher Befragungen kann der personelle Aufwand allerdings nicht ohne weiteres verringert werden, „weil die einfacheren Lösungen nicht lediglich etwas bescheidener, sondern deutlich eher sachlich irreführend sind“ (vgl. Schomburg/ Teichler

---

<sup>221</sup> Zum Begriff vgl. Kromrey (1998: 259) bzw. Fußnote 119 in dieser Arbeit.

1997: 257). Auch Reinfeld/ Frings (2003: 291) weisen darauf hin, daß solche Untersuchungen „voraussetzungsvoll und methodisch anspruchsvoll“ sind. Andererseits kann dem vergleichsweise hohen Aufwand für die hochschulweite Durchführung von Absolventenstudien „nicht nur in einem evaluativen Sinne, sondern auch im Hinblick auf die Alumniarbeit ein hoher Ertrag entgegenstehen“ (vgl. Reinfeld/ Frings 2003: 292). So könne der mit der Befragung verbundene Aufbau eines Adressenpools auch anderweitig genutzt werden. Neben der Absolventenbetreuung kann dieser auch für die Öffentlichkeitsarbeit einer Hochschule, z.B. zur Versendung eines elektronischen Newsletters oder zur Offerierung von Stellenangeboten und Weiterbildungsmöglichkeiten genutzt werden. Durch Analysen zu Hochschulbindung mit Hilfe von Absolventenbefragungen (vgl. Krempkow/ Pastohr 2003) könnten außerdem die Ergebnisse solcher Studien auch für ein zielgruppenspezifisches Hochschulmarketing bis hin zur Unterstützung von Fundraisingaktivitäten genutzt werden.

Insgesamt steht unter Berücksichtigung der durch die Nutzung der Ergebnisse von Absolventenbefragungen zu vermeidenden Fehlanreizwirkungen, der kostensparenden Weiterentwicklungsmöglichkeiten in Richtung Online-Befragungen und der weiteren Nutzungsmöglichkeiten der aus Absolventenbefragungen resultierenden aktuellen Adreßpools dem vergleichsweise hohen Aufwand auch ein vergleichsweise hoher Ertrag gegenüber.

### **5.5.5 Fazit zur Eignung von Absolventenbefragungen**

Absolventenbefragungen können trotz einiger aufgrund ihrer Retrospektivität vorzunehmenden Einschränkungen der Aussagekraft zur Einschätzung der Studiensituation grundsätzlich als geeignet gelten. Während kurz nach Abschluß des Studiums durch den erstmals möglichen Gesamtüberblick bei geringen Retrospektivitätseffekten ihre Aussagekraft als am höchsten eingeschätzt werden kann, kann sie mit größer werdendem zeitlichem Abstand (nach mehreren Jahren oder gar Jahrzehnten) durch Erinnerungsunschärfen oder Verklärungen sinken. Vergleiche der Ergebnisse von Studentenbefragungen und Absolventenbefragungen an der TU Dresden sowie der Ergebnisse von Absolventenbefragungen kurz nach Studienabschluß und mit mindestens 5 Jahren Abstand zum Studienabschluß konnten aber lediglich die Tendenz zu einer etwas

positiveren Einschätzung mit größer werdendem zeitlichen Abstand und keine grundsätzlich anderen Bewertungsmuster aufzeigen.

Fragenkataloge zum Übergang vom Studium in den Beruf, die wie z.B. die Kasseler Studien bereits über einen längeren Zeitraum genutzt werden, können als validiert gelten.

Zur Vergleichbarkeit von Absolventenbefragungen kann formuliert werden, daß diese unter Berücksichtigung solcher Aspekte, die eine Hochschule nicht zu beeinflussen in der Lage ist, als gegeben angesehen werden kann. Durch eine Hochschule nicht beeinflusst werden kann insbesondere die Ausgangssituation der Absolventen schon am Beginn ihres Studiums bezüglich sozialer Herkunft und Geschlecht, die über das gesamte Studium hinweg bis zum beruflichen Verbleib wirken und Einfluß auf den beruflichen Erfolg haben kann. Aber auch die aktuelle persönliche Situation der Absolventen zum Befragungszeitpunkt wie z.B. das Vorhandensein von Kindern oder der Partnerstatus sollten grundsätzlich berücksichtigt werden, da sie ebenfalls Einfluß auf den beruflichen Erfolg haben können. Weiteren Anforderungen wie der Zeitnähe können Absolventenbefragungen nur bezüglich der Angaben zum beruflichen Verbleib und Erfolg genügen, da sich retrospektive Bewertungen des Studiums auf eine bereits länger zurückliegende Situation beziehen. Der Anforderung, sinnvolle Anreize zu schaffen, ohne unerwünschte Anpassungseffekte zu erzielen, kann durch Absolventenbefragungen sehr gut entsprochen werden, da sie häufig als einzige Möglichkeit gesehen werden, nicht nur die Quantität bzw. „Stückzahl“ der Absolventen zu erfassen, sondern auch deren Qualität. Absolventenbefragungen erfordern allerdings einen vergleichsweise hohen Aufwand, der erst mittelfristig durch kostensparende Weiterentwicklungsmöglichkeiten in Richtung Online-Befragungen etwas gesenkt werden könnte. Dem Aufwand stehen als Ertrag jedoch neben der alternativlosen Informationsgewinnung zum beruflichen Verbleib und der (derzeit leider nicht finanziell bezifferbaren) Möglichkeit zur Vermeidung von Fehlanreizwirkungen weitere Nutzungsmöglichkeiten der aus Absolventenbefragungen resultierenden aktuellen Adreßpools und damit auch ein vergleichsweise hoher Ertrag gegenüber.

## **5.6 Lehrendenbefragungen als Bestandteil der Qualitätsorientierung**

Wenn als Ziel von Leistungsbewertungen und Leistungsanreizen formuliert wird, daß hiermit (auch) die Motivation von Lehrenden erhöht werden soll, so liegt es nahe, über Befragungen von Lehrenden entweder als Begleitmaßnahme (z.B. zur Untersuchung der Akzeptanz) oder angelehnt an Mitarbeiterbefragungen in Unternehmen als Bestandteil von Modellen zur Förderung der Qualitätsorientierung nachzudenken. In diesem Abschnitt soll es lediglich um letzteren Aspekt gehen. Lehrendenbefragungen zum Zwecke der Akzeptanzuntersuchung und die evtl. daraus folgenden Möglichkeiten zur Akzeptanzsteigerung sollen in einem separaten Abschnitt diskutiert werden.

Bei Lehrendenbefragungen als Bestandteil der Qualitätsorientierung kann wiederum in zwei Aspekte unterschieden werden: Zum einen können Lehrende zu ihrer Einschätzung der Situation von Lehre und Studium befragt werden und diese Einschätzungen als subjektive Indikatoren aufgefasst werden. Zum anderen können Lehrende zu ihrer Arbeitszufriedenheit insgesamt bzw. zu ihrer Zufriedenheit mit einzelnen Aspekten ihrer beruflichen Situation befragt werden, so z.B. zur Zufriedenheit mit der Höhe und Leistungsgerechtigkeit des Gehalts, aber auch zur Zufriedenheit mit der beruflichen Anerkennung. Dies sind Aspekte, die direkt die Reformvorhaben zur Leistungsbewertung und Leistungsanreizen tangieren und auf die an im sechsten Kapitel dieser Arbeit noch zurückgekommen wird. Daher soll die Eignung von Lehrendenbefragungen nachfolgend nicht nur zur Einschätzung der Situation von Lehre und Studium, sondern für beide Zielstellungen diskutiert werden.

### **5.6.1 Die Datenqualität der Ergebnisse von Lehrendenbefragungen**

Die Validität der Ergebnisse von Lehrendenbefragungen zur Einschätzung der Situation von Lehre und Studium wurde bisher eher selten untersucht. Von den wenigen Autoren, die hierzu empirische Analysen durchführten und diese veröffentlichten, werden eher Einschränkungen formuliert. Häufig wird darauf hingewiesen, daß die Perspektive der Lehrenden vor allem durch den Forschungshintergrund geprägt sei (vgl. Müller-Böling u.a. 2001: 11, Rindermann 2001: 174 sowie darin Übersicht über neun weitere Studien). Die vorliegenden Publikationen beziehen sich dabei v.a. auf den Aspekt der Kriteriumsvalidität,

indem sie die Übereinstimmung der Urteile zwischen verschiedenen Befragtengruppe untersuchen. Zunächst soll jedoch die Inhaltsvalidität diskutiert werden.

### **Inhaltsvalidität**

Die Inhaltsvalidität wurde in den vorliegenden Publikationen kaum angesprochen. Lediglich von Rindermann (2001: 44ff.) ist eine kleine Studie (n=58) mittels offener Fragestellungen bekannt, mit der er versuchte, Merkmale „guter Lehre“ auch aus Sicht von Lehrenden zu erfassen. Hierbei stellte sich heraus, daß Lehrende im wesentlichen dieselben Aspekte nannten wie Studierende und diese Aspekte wiederum von den in Studierendenbefragungen verwendeten Fragen gut erfaßt werden.<sup>222</sup> Da zur Einschätzung der Situation von Lehre und Studium häufig dieselben bzw. nur leicht veränderte Fragestellungen wie in Studierendenbefragungen verwendet wurden und deren Inhaltsvalidität bereits im Zusammenhang mit der Eignung von Studierendenbefragungen diskutiert und als ausreichend befunden wurde, kann für solche mit Studierendenbefragungen vergleichbare Fragestellungen an dieser Stelle auf die entsprechende Literatur verwiesen werden. Da jedoch nur eine Studie hierzu bekannt ist, die Lehrende in größerem Umfang einbezog, und die Ergebnisse möglicherweise an anderen Hochschulen etwas anders ausfallen könnten, soll hier zusätzlich anhand eigener Überlegungen und hilfsweise unter Hinzuziehung „halboffener“ Fragestellungen aus den eigenen Lehrendenbefragungen die Inhaltsvalidität dieser Lehrendenbefragungen zur Einschätzung der Situation von Lehre und Studium diskutiert werden.

Als (wenngleich nur teilweise geeigneter) Indikator für die Inhaltsvalidität von geschlossenen Fragestellungen wird in empirischen Studien häufig eine „halboffene“ Fragestellung zu Hilfe genommen (z.B. „Sonstiges“). Wenn sehr viele oder gar eine Mehrheit der Befragten nicht die vorgegebenen Antwortkategorien nutzt, sondern „Sonstiges“ ankreuzt, wird dies als Indiz für eine geringe Treffgenauigkeit der vorgegebenen Antwortkategorien für die Befragten gesehen und kann damit als mangelnde Inhaltsvalidität interpretiert werden. Da in die eigenen Lehrendenbefragungen aus Ressourcengründen nur zwei Fragestellungen (zu Rahmenbedingungen von Lehrveranstaltungen und zum Studiengang) mit insgesamt lediglich 15 Items einbezogen werden konnten, würde hier über übliche halboffene Fragestellungen hinausgehend sogar explizit gefragt: „Gibt es Aspekte, die Ihnen wichtig sind und hier nicht berücksichtigt

wurden?“ Ein hoher Anteil und/ oder ein nicht mit Studierendenbefragungen vergleichbarer Inhalt der Antworten auf diese Frage könnte dann ebenfalls als Hinweis für eine mangelnde Inhaltsvalidität interpretiert werden.

Trotz der expliziten Aufforderung und den im Vergleich zu Studierendenbefragungen stark gekürzten Frageinhalten zur Situation von Lehre und Studium machte an allen drei Hochschulen, an denen eigene Lehrendenbefragungen stattfanden, nur ein Bruchteil der Befragten Gebrauch von dieser Möglichkeit (TU Dresden: 14%, EUV Frankfurt: 13%, HS Zittau-Görlitz: 17%). Zwar wird häufig argumentiert, daß der Anteil offener Antworten generell geringer ausfällt (als er evtl. „eigentlich“ ist), weil die Befragten sich die zusätzliche Mühe machen müßten, hier etwas selbst zu formulieren und einzutragen. In dem hier diskutierten Fall war es den Befragten jedoch auch möglich, einfach nur ein Kreuz zu setzen, da erst unter der Fragestellung die Aufforderung kam, dann ggf. die nicht berücksichtigten Aspekte zu nennen. Selbst wenn man einräumen muß, daß die in den eigenen Lehrendenbefragungen verwendeten verkürzten Fragestellungen zur Situation von Lehre und Studium von einem nennenswerten Anteil der Befragten als ergänzungsbedürftig angesehen werden, so sieht offenbar die Mehrheit von über 80% diese als nicht unbedingt ergänzungsbedürftig an. Eine alternative Interpretation wäre, daß den befragten Lehrenden die Lehre generell nicht wichtig genug sei, um hierzu zusätzliche Aspekte zu formulieren. Dies würde zugleich aber einen mit vier Fünfteln der Lehrenden extrem weit verbreiteten geringen Stellenwert der Lehre unterstellen.

Die zweite Möglichkeit, anhand der „halboffenen“ Fragestellung in den eigenen Lehrendenbefragungen deren Inhaltsvalidität zu diskutieren, war der Inhalt der Antworten. Wenn dieser kaum mit den aus Studierendenbefragungen adaptierten und in den eigenen Lehrendenbefragungen verwendeten (bzw. sonst in Studierendenbefragungen verwendeten und aus Ressourcen Gründen nicht übernommenen Fragestellungen) übereinstimmt, könnte dies als Indiz für mangelnde Inhaltsvalidität gesehen werden. Deshalb sollen hier exemplarisch für die Lehrendenbefragung an der TU Dresden die Antworten auf die halboffene Frage dokumentiert werden. An der TU Dresden machten 28 (von 206) Befragte Angaben hierzu.

---

<sup>222</sup> Am häufigsten wurden didaktische Kompetenzen und das Engagement der Lehrenden sowie die Strukturierung der Lehre genannt (vgl. Rindermann 2001: 44ff.).

Als weitere wichtige Aspekte zur Studiensituation wurden insgesamt genannt (alphabetisch geordnet):

<b>Hochschullehrer</b>	<b>Wissenschaftliche Mitarbeiter</b>
Aktualität der Lehrinhalte Anerkennung von Engagement in der Lehre Betreuungsrelation Bezug zu anderen Disziplinen Einbindung Praxisvertreter Finanzielle Ausstattung, z.B. für Diplomarbeiten Internationalisierung Internationalisierung/ Interdisziplinarität Kommunikationsbereitschaft der Studierenden Orientierung bzgl. Schwerpunktsetzung Personelle Absicherung des Lehrangebots	Abstimmung (zwischen Fächern) Abstimmung zwischen Fachbereichen Aktualität der LVA Anerkennung von Lehrleistungen das Studium ist nicht straff genug organisiert Drittmitteltätigkeit ist nicht Lehrtätigkeit ergänzende und wahlobligatorische VL flache Hierarchie Gestaltung der Prüfung Stoff-/Zeitverhältnis Studenten auswählen Teamwork/ Kommunikation/ Zukunftsvision/ Konzept/ Entwicklung Verfügbarkeit von Lehrenden Verhältnis Lehrkräfte vs. Studenten Voraussetzungen bei Studierenden (Allgemeinbildung, Engagement, Motivation) Vorlesungsgröße Zahlenverhältnis Lehrende – Studenten

Ein Vergleich mit den in eigenen Lehrendenbefragungen verwendeten Fragestellungen zeigte, daß dies in der Tat zusätzliche Aspekte sind. Vergleicht man die genannten weiteren Aspekte jedoch mit den sonst in Studierendenbefragungen verwendeten und aus Ressourcengründen hier nicht übernommenen Fragestellungen, offenbart sich aber, daß diese Aspekte weitgehend in Studierendenbefragungen enthalten sind bzw. durch objektive Indikatoren erfasst werden können (vgl. hierzu auch die Übersicht zu möglichen Indikatoren der Leistungsbewertung am Ende des 3. Kapitels). Beispielsweise die von Hochschullehrern genannte Betreuungsrelation bzw. das von wissenschaftlichen Mitarbeitern genannte Zahlenverhältnis Lehrende – Studierende (z.B. als durchschnittliche Anzahl der Studierenden je Professor im Fach) kann als sogenannter objektiver Indikator erfasst werden. Andere Aspekte, wie z.B. das von wissenschaftlichen Mitarbeitern genannte Stoff-/ Zeitverhältnis, die Abstimmung zwischen Fächern, die Aktualität der Lehre (z.B. in Bezug auf den Stand der Forschung) wurden bereits in Studentenbefragungen verwendet (vgl. Krempkow/ Winter

2000, Krempkow 2003b). Und schließlich wurden einzelne Aspekte wie (die Anerkennung von) Engagement in der Lehre sowohl in Studierendenbefragungen (in Form von Fragen zum Dozentenengagement in der Lehre – vgl. Faktorenstruktur von Studierendenbefragungen im Abschnitt 5.4.2 dieser Arbeit) als auch in eigenen Lehrendenbefragungen (in Form von einer Frage zur beruflichen Anerkennung im Bereich der Lehre) verwendet.

Zusammenfassend kann aus den wenigen vorliegenden Untersuchungen zur Inhaltsvalidität von Lehrendenbefragungen mit der gebotenen Vorsicht geschlossen werden, daß es keine Hinweise auf eine generell mangelnde Inhaltsvalidität gibt. Die vorliegenden Ergebnisse deuten vielmehr darauf hin, daß Lehrende im wesentlichen dieselben Aspekte als relevant ansehen, wie sie in Studierendenbefragungen zur Einschätzung der Situation von Lehre und Studium genutzt werden. Stärker verkürzte Befragungsinstrumente bringen jedoch tendenziell ein höheres Risiko mit sich, relevante Aspekte zu vernachlässigen.

Der zweite größere Themenbereich, für den Lehrendenbefragungen als Bestandteil der Qualitätsorientierung genutzt werden können, war deren Zufriedenheit mit der beruflichen Situation, wie z.B. zur Zufriedenheit mit der Höhe und Leistungsgerechtigkeit des Gehalts, aber auch mit der beruflichen Anerkennung. Da diese Aspekte für Reformvorhaben zur Leistungsbewertung und Leistungsanreizen hier als relevant angesehen und weil darauf im sechsten Kapitel dieser Arbeit noch zurückgegriffen werden soll, wird dieser Themenbereich hier ebenfalls diskutiert. Diese wurde hier in den eigenen Lehrendenbefragungen als Gesamteinschätzung und auch als Einschätzung von Einzelaspekten erfasst. Für die Einschätzung der Inhaltsvalidität ist die Untersuchung der Frage relevant, ob mit den erfragten Einzelaspekten möglichst alle Merkmale der Arbeitszufriedenheit erfasst werden oder ob aus Sicht der Befragten wichtige Merkmale fehlen.

Da keine offenen Befragungen von Hochschullehrenden zu dieser Frage bekannt sind, kann eine Einschätzung wieder lediglich hilfsweise über die Ergebnisse einer halboffenen Fragestellung erfolgen. Wie sich zeigte, machte von dieser Möglichkeit aber nur ein sehr geringer Anteil der Lehrenden Gebrauch (TU Dresden: 9%, EUV Frankfurt: 6%, HS Zittau-Görlitz: 8%), so daß hier offenbar seitens der Befragten kein großer Ergänzungsbedarf gesehen wurde. Im einzelnen wurden als zusätzliche wichtige Aspekte der Arbeitszufriedenheit genannt (innerhalb der Statusgruppen alphabetisch geordnet):

Hochschullehrer	Wissenschaftliche Mitarbeiter
Beziehung zum Ministerium Nachwuchsförderung Subsidiarität, d.h. Delegation von Entscheidungen so weit wie möglich nach unten Verhältnis der finanziellen Ausstattung bei "West-bzw. Ost- Berufung" (Chancengleichheit)	Arbeitsräume Dienstleistungen (Bibliothek, Fo-Förderung, Dienstreisen) Geschwindigkeit beim Bücherkauf (SLUB) halbe Stelle > volle Arbeit > keine Zukunft; Entscheidungstransparenz; fehlende Zentralisierung von homogenen Aufgaben kein Beamer im Lehrsaal, Fernseher Lehrstuhlleitung Leistungsgerechtigkeit des Gehalts im Vergleich zu Vorgesetzten Osttarif Selbstdarstellung der Universität nach außen soziale/ sonstige Leistungen des Arbeitgebers (z.B. ÖPNV-Ticket, Essen-Zuschuß) Verwaltungsaufwand bei Personalangelegenheit, sonst. Verwaltungsakten

Ein Vergleich dieser einzelnen Angaben mit den Vorgaben im Fragebogen zeigte, daß ein Teil der als zusätzlich wichtige Aspekte genannten Merkmale im weiteren Sinne auch den Vorgaben zugeordnet werden könnte. So könnte z.B. die Leistungsgerechtigkeit des Gehalts im Vergleich zu Vorgesetzten auch der Leistungsgerechtigkeit des Gehalts im Vergleich zu Kollegen zugeordnet werden (wenngleich die Hierarchie offenbar aus Sicht von wiss. Mitarbeitern an den Hochschulen stark ausgeprägt ist, ging es hier vorrangig um die Leistungsgerechtigkeit des Gehalts generell, so daß der Hierarchieaspekt an dieser Stelle ggf. vernachlässigt werden könnte). Ähnlich sprechen weitere genannte Merkmale, wie „halbe Stelle > volle Arbeit > keine Zukunft“ z.B. auch die Arbeitsplatzsicherheit an. Oder der „Verwaltungsaufwand bei Personalangelegenheiten, sonst. Verwaltungsakten“ kann der Art der Verwaltung in der Hochschule zugeordnet werden. Sicherlich ist es eine Abwägungsfrage, inwieweit man hierbei verallgemeinernde Zuordnungen treffen und dabei die Besonderheiten der einzelnen offenen Antworten vernachlässigen möchte. Grundsätzlich möglich erscheint dies jedoch, so daß hier zusätzlich zu dem ohnehin geringen Anteil der Antworten auf die halboffene Frage ein mehr oder weniger großer Teil der Antworten

zusätzlich noch den Vorgaben zugeordnet werden könnte (die wiederum auch bereits das Ergebnis vorangegangener Studien zur Arbeitszufriedenheit sind).

Insgesamt kann zu den erfragten Aspekten der Arbeitszufriedenheit festgehalten werden, daß aus den hier hilfsweise anhand der halboffenen Fragestellungen durchgeführten Untersuchungen keine Zweifel an deren Inhaltsvalidität resultieren.

### **Kriteriumsvalidität**

Bezüglich der Kriteriumsvalidität wurde bereits erwähnt, daß Autoren, die sich mit der Validität von Lehrendenbefragungen beschäftigten, die Perspektive der Lehrenden vor allem als durch den Forschungshintergrund geprägt sehen. Hinzu kommt das grundsätzliche Problem, daß es sich bei einigen für die Lehre wichtigen Aspekten in einer Befragung von Lehrenden um Selbsteinschätzungen handelt. Bereits weiter vorn im Zusammenhang mit der Kriteriumsvalidität von Studierendenbefragungen wurden Ergebnisse einer Studie von Hornbostel (2001: 9) angeführt, die bei am wenigsten im Verdacht einer Beeinflussung durch Selbsteinschätzungsprobleme stehenden Aspekten je nach Fach Korrelationskoeffizienten zwischen  $r=.8$  und  $.55$  ergaben. Im Durchschnitt der untersuchten Fächer betrug die Korrelation  $r=.69$ .

Für andere Aspekte, bei denen ein stärkerer Einfluß durch Selbstbewertungseffekte zu vermuten war, lagen die Korrelationen im Mittel zwischen  $r=.46$  und  $.38$ . Während für wenig oder nicht durch Selbsteinschätzungsaspekte beeinflussten Urteile also von einer relativ guten Übereinstimmung der Einschätzungen ausgegangen werden kann, fällt diese bei Selbsteinschätzungen deutlich geringer aus. Rindermann (2001: 129ff.) berichtete bezüglich Lehrveranstaltungsbewertungen noch geringere Korrelationen zwischen Lehrenden und Studierenden, während sie zwischen unabhängigen Fremdgutachtern und Studierenden höher ausfielen (vgl. Abschnitt zur Datenqualität von Studierendenbefragungen in dieser Arbeit).

Auch eigene Analysen anhand eines Vergleiches der Einschätzungen von Lehrenden und Studierenden zur Situation von Lehre und Studium an der TU Dresden zeigten, daß Lehrende eine andere Perspektive haben und die Studiensituation tendenziell deutlich weniger kritisch sehen – insbesondere bei Fragen, die Aspekte der Selbstbewertung beinhalten. Allerdings ist hierbei nach Status zu differenzieren (vgl. Krempkow 2001: 21).

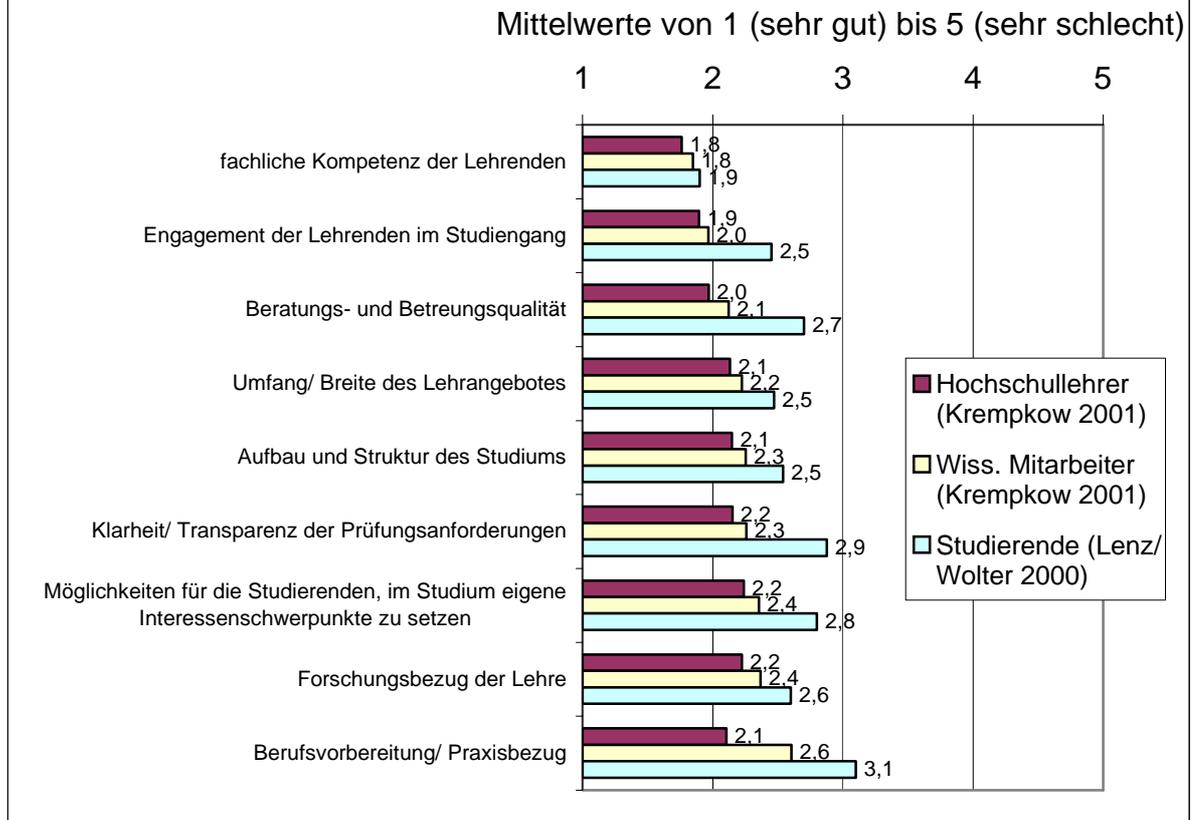
Dies soll nachfolgend anhand der Ergebnisse zur Einschätzung der Rahmenbedingungen von Lehre und Studium exemplarisch für die TU Dresden dargestellt werden. Die Einschätzungen von Lehrenden und Studierenden (vgl. Lenz/ Wolter/ Winter 2000: 83)<sup>223</sup> zur Situation von Lehre und Studium im jeweiligen Studiengang bzw. Fachbereich sind direkt vergleichbar, da weitgehend dieselben Fragestellungen verwendet wurden. Die Graphik kann daher aufzeigen, wo es besonders große Unterschiede zwischen den Einschätzungen von Lehrenden und Studierenden gibt und wo eher geringe. Auf eine Einbeziehung der Ergebnisse der Lehrendenbefragungen an den anderen beiden Hochschulen wird verzichtet, da hierzu keine Ergebnisse von Studierendenbefragungen vorliegen und sich die Ergebnisse der Lehrendenbefragungen nur in wenigen Aspekten deutlich unterscheiden (vgl. auch Krempkow 2002, Krempkow 2003).<sup>224</sup>

---

<sup>223</sup> Die Studierendenbefragung fand im Rahmen der Imagestudie der TU Dresden parallel bzw. in Abstimmung mit einer Bevölkerungsumfrage und Mitarbeiterbefragung statt. Hierbei wurden per Zufallsauswahl durch das Immatrikulationsamt der TU Dresden 2023 Studierende angeschrieben, von denen 1201 (59 Prozent) antworteten. Aufgrund der Fächerverteilung im Datensatz zur Imagestudie, die mit der vom Immatrikulationsamt veröffentlichten Verteilung sehr gut übereinstimmt (maximale Abweichung 2,7 Prozentpunkte bei SLW), sowie aufgrund der übereinstimmenden Anteile weiblicher und männlicher Studierender ist die Befragung als repräsentativ einzuschätzen. (Die Untersuchungen ergaben eine Beteiligung von 53,8 Prozent weiblichen und 46,2 Prozent männlichen Studierenden. Die Zahlen des Immatrikulationsamtes (Stand 1.12.1999) verweisen auf exakt denselben Frauenanteil, nämlich 53,8 Prozent.

<sup>224</sup> Die eigenen Lehrendenbefragungen an den drei Hochschulen mußten aus Ressourcengründen in verschiedenen Jahren durchgeführt werden. Unterschiede in den Ergebnissen können daher auch durch den zeitlichen Abstand mitverursacht sein. Es wurde jedoch eine zusätzliche Fragestellung aufgenommen, mit der evtl. Unterschiede in der Situation im Befragungssemester im Vergleich zu vorhergehenden Semestern eingeschätzt werden sollten. Wie sich zeigte, veränderte sich die Situation von Lehre und Studium für die Mehrheit im Vergleich zu vorhergehenden Semestern nicht (vgl. Krempkow 2001). Bei Kopp/ Weiß (1993) hatte sich nach Einschätzung der bundesweit befragten Professoren die Situation in ihrer Organisationseinheit in den letzten 5 Jahren ebenfalls für die Mehrheit nicht verändert. Daher wird davon ausgegangen, daß die Einschätzungen der Situation über wenige Jahre hinweg gesehen weitgehend stabil sind und ein Vergleich von Ergebnissen unterschiedlicher Befragungen über solche Zeiträume hinweg zumindest keinen allzu großen Verzerrungen unterliegen dürfte.

## Situation von Lehre und Studium im Studiengang/ Fachbereich



Quelle: Krempkow 2001

Die Situation von Lehre und Studium im Studiengang/ Fachbereich insgesamt, für den die Befragten in der Lehre überwiegend tätig waren, unterscheidet sich zwischen Hochschullehrern und Wissenschaftlichen Mitarbeitern in mehreren Aspekten, aber meist nur geringfügig. Lediglich bei der Berufsvorbereitung zeigte sich ein signifikanter Unterschied zwischen Hochschullehrern und Wissenschaftlichen Mitarbeitern. (Die Aspekte in der Graphik wurden nach dem Mittelwert über alle Lehrenden sortiert.) Auffällig ist bei dieser Darstellung vor allem, daß Hochschullehrer zu allen Aspekten tendenziell die beste Bewertung vornahmen, die wissenschaftlichen Mitarbeiter die zweitbeste und die Studierenden (meist mit größerem Abstand) die schlechteste. Während die Unterschiede zwischen Hochschullehrern und wissenschaftlichen Mitarbeitern meist nicht signifikant sind und daher nicht zu stark interpretiert werden sollten, sind die Unterschiede zwischen

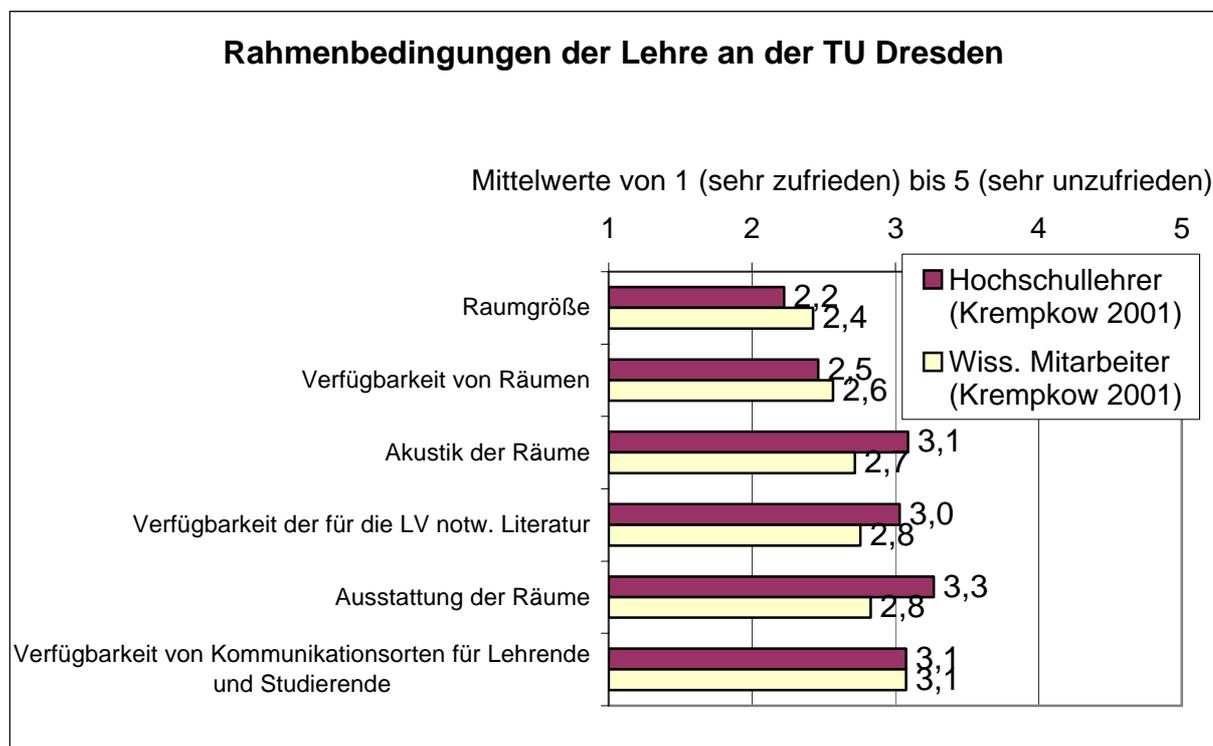
Lehrenden und Studierenden deutlich. Unter Zuhilfenahme der 95%-Konfidenzintervalle beider Befragungen sind sie auch durchweg als signifikant anzusehen.

Das könnte so interpretiert werden, daß es in der Tat zwischen Lehrenden und Studierenden einen systematischen Perspektivenunterschied gibt, wobei der Perspektivenunterschied (eventuell aufgrund des geringeren Statusabstandes oder aufgrund des geringeren Zeitabstandes zum eigenen Studium) zwischen wissenschaftlichen Mitarbeitern und Studierenden tendenziell etwas kleiner ausfällt. Besonders groß sind die Unterschiede zwischen Lehrenden und Studierenden beim Engagement der Lehrenden, der Beratungs- und Betreuungsqualität und den Prüfungsanforderungen. Deutlich geringer sind die Unterschiede der Einschätzungen bei Umfang/ Breite des Lehrangebotes, Aufbau und Struktur des Studiums, aber auch beim Forschungsbezug der Lehre. Selbst bei den Einschätzungen der fachlichen Kompetenz der Lehrenden, wo die Validität der Einschätzungen von Studierenden besonders fraglich ist, bestehen nur geringe Differenzen. Insgesamt zeigt sich beim Vergleich der Ergebnisse der eigenen Lehrendenbefragung mit Ergebnissen einer Studierendenbefragung, daß besonders große Unterschiede v.a. bei Aspekten auftreten, die eine Selbsteinschätzung der Lehrenden beinhalten, bei anderen Aspekten jedoch größere Übereinstimmung herrscht.

Werden die Korrelationen der mittleren Einschätzungen von Lehrenden und Studierenden betrachtet, so zeigt sich zwischen den Urteilen von wiss. Mitarbeitern und Studierenden ein deutlich höherer Zusammenhang ( $r=.80$ ) als zwischen Professoren und Studierenden ( $r=.50$ ). Faßt man Professoren und wiss. Mitarbeiter zusammen und vergleicht sie mit denen von Studierenden, so ergibt sich insgesamt eine Korrelation von  $r=.77$ .

Einschätzungen zu Rahmenbedingungen der Lehre, die auch als Arbeitsbedingungen gesehen werden und daher kaum Selbsteinschätzungsaspekte enthalten, liegen für Studierende leider nicht vergleichbar vor, daher können sie nur getrennt nach Professoren und wissenschaftlichen Mitarbeitern erfolgen. Bereits in vorhergehenden Befragungen wie auch in dieser hatte sich herausgestellt, daß zwischen diesen Befragtengruppen "Welten liegen" können (vgl. Enders/ Teichler 1995: 27). Auch hier unterscheiden sich Hochschullehrer und wissenschaftliche Mitarbeiter in ihren Bewertungen z.T. deutlich. Bei der Ausstattung und Akustik der Räume sowie bei der Verfügbarkeit der für die Lehrveranstaltungen notwendigen Literatur sind die Unterschiede signifikant. Da es sich bei den hier genannten Rahmenbedingungen nicht um Selbsteinschätzungsaspekte handelt, können die Unterschiede nicht darauf zurückgeführt werden. Ob es sich um reale

Unterschiede handelt, die aufgrund der Tatsache zustande kommen, daß wissenschaftliche Mitarbeiter z.B. eher Seminare/ Übungen und Hochschullehrer eher Vorlesungen abhalten und in größeren Hörsälen tatsächlich z.B. eine schlechtere Ausstattung und Akustik vorherrscht, kann an dieser Stelle nicht geklärt werden.



Quelle: Krempkow 2001

Generell wird bei Selbsteinschätzungsaspekten eine gewisse Skepsis bezüglich der Validität geäußert. Da auch die vorliegenden Ergebnisse zumindest bei Selbsteinschätzungsaspekten geringere Übereinstimmungen zeigen, erscheint es angemessen, diesbezüglich von eingeschränkter Validität von Lehrendenbefragungen zur Einschätzung der Situation von Lehre und Studium auszugehen.

### **Konstruktvalidität**

Die Konstruktvalidität wäre v.a. bedeutsam, wenn aus den Ergebnissen von Lehrendenbefragungen z.B. wie beim HILVE Dimensionen gebildet würden. In den verfügbaren Publikationen über Lehrendenbefragungen spielte die Konstruktvalidität bisher keine Rolle. Eine versuchsweise durchgeführte Faktorenanalyse zu den Ergebnissen der drei eigenen Lehrendenbefragungen (vgl. Krempkow 2001, 2002, 2003) über alle Items zur

Einschätzung der Situation von Lehre und Studium und der Arbeitszufriedenheit hinweg (Hauptkomponentenanalyse mit Varimax-Rotation) ergab drei Dimensionen, die allerdings nicht sehr trennscharf ausfallen. Diese könnte man bezeichnen als: Lehrveranstaltungen/ Zeitbudget, Kommunikationskultur/ Arbeitszufriedenheit, Lehrinhalte und Studienstruktur im eigenen Fach. Allerdings ist diese Faktorenstruktur an den einzelnen Hochschulen nicht stabil dieselbe, sondern fällt sehr unterschiedlich aus. Da die Validität solcher Konstrukte deshalb nicht als gegeben angesehen werden kann und es auch nicht das Ziel der Untersuchungen war, Dimensionen zu konstruieren, wird hier auf solche Dimensionen als übergreifende Konstrukte verzichtet.

Etwas anders sieht dies aus, wenn man nicht alle Items zur Einschätzung der Situation von Lehre und Studium und der Arbeitszufriedenheit, sondern lediglich die Aspekte der Arbeitszufriedenheit selbst betrachtet. Hierzu wurden in der Vergangenheit bereits aus theoretischen Überlegungen und empirischen Studien Dimensionen bzw. Kategorien der Arbeitszufriedenheit entwickelt (vgl. Abschnitt zu Determinanten der beruflichen Zufriedenheit im nachfolgenden Kapitel). Die Übertragbarkeit dieser Dimensionen auf Lehrende können anhand der eigenen Lehrendenbefragungen zumindest ansatzweise mittels Faktorenanalysen geprüft werden. Sowohl für die TU Dresden als auch für alle drei Hochschulen zeigt sich in der Tat eine weitgehend übereinstimmende Faktorenstruktur. Die Aspekte der Arbeitszufriedenheit lassen sich demnach vier Dimensionen zuordnen. Faktor 1 umfaßt eher ideelle Aspekte des Erfolgs: Aufstiegsmöglichkeiten und Möglichkeiten, eigene Ideen zu verwirklichen, berufliche Anerkennung und die Sicherheit des Arbeitsplatzes. Faktor 2 besteht aus den zwei Aspekten des materiellen Erfolges: Höhe und Leistungsgerechtigkeit des Gehalts. Faktor drei umschreibt einen Aspekt der Arbeit selbst (Lehrveranstaltungen) und das Verhältnis zu Kollegen. Faktor 4 umfaßt im wesentlichen Kontextaspekte: z.B. „Managementverhaltensweisen“ wie Informationen über Entscheidungen der Hochschulleitung, Mitbestimmungsmöglichkeiten und Art der Ressourcenzuteilung. Diese vier Faktoren erklären immerhin 59% der Varianz, obwohl ein wesentlicher Aspekt der Arbeit von Lehrenden, die Forschung, in der Befragung von vornherein nicht erfasst wurde. Die Dimensionen stimmen auch relativ gut mit den theoretischen Vorüberlegungen überein. Im Unterschied zu anderen, im o.g. Abschnitt dieser Arbeit beschriebenen Modellen der Arbeitszufriedenheit lässt sich der berufliche Erfolg nicht zu einer Dimension zusammenfassen, sondern besteht aus zwei Dimensionen (ideell und materiell). Dies

entspricht jedoch den im Zusammenhang mit Anreizmodellen formulierten Überlegungen, daß Lehrende nicht allein materiell, sondern in stärkerem Umfang auch ideell motiviert sind. Davon abgesehen ist lediglich noch der Aspekt Verhältnis zu Kollegen nicht (wie in den zugrunde gelegten Modellen der Arbeitszufriedenheit) der Dimension Kontext zuzuordnen, sondern bildet gemeinsam mit den Lehrveranstaltungen einen gemeinsamen Faktor. Außerdem wurde die Möglichkeit, eigene Ideen zu verwirklichen, dem ideellen beruflichen Erfolg zugeordnet und nicht der Arbeit selbst. Auch dies erscheint plausibel, da häufig die höhere Autonomie als wesentlicher Vorteil einer Hochschultätigkeit im Vergleich zu z.T. besser bezahlten Tätigkeiten in der Privatwirtschaft genannt wird. Ansonsten stimmt die Dimensionalität mit theoretischen Überlegungen und Modellen der Arbeitszufriedenheit überein. Zwar ist an dieser Stelle keine ausführlichere Diskussion der Dimensionalität von Aspekten der Arbeitszufriedenheit möglich. Dennoch kann aufgrund der Ergebnisse formuliert werden, daß die Dimensionalität sowohl für die exemplarisch herangezogene TU Dresden, als auch über mehrere Hochschulen hinweg betrachtet stabil ist und mit den zum Vergleich herangezogenen Modellen der Arbeitszufriedenheit weitgehend übereinstimmt. Daher lassen sich keine starken Hinweise auf eine mangelnde Konstruktvalidität finden .

### **Reliabilität**

Neben der Validität war auch die Reliabilität ein wichtiges Kriterium für die Einschätzung der Datenqualität. Auch zur Reliabilität von Lehrendenbefragungen wurden in den vorliegenden Publikationen von Lehrendenbefragungen bisher kaum Ergebnisse berichtet. Eine Möglichkeit zur Überprüfung z.B. der Retestreliabilität wäre ein Vergleich der Ergebnisse derselben Fragestellungen bei relativ kurz hintereinander durchgeführten Lehrendenbefragungen an einer Hochschule. An der TU Dresden ist dies teilweise möglich, da eine im Jahr 2000 durchgeführte Imagebefragung bei Hochschulmitarbeitern (vgl. Lenz/Wolter/ Winter 2000) z.T. ähnliche Fragestellungen enthielt wie die im Jahr 2001 durchgeführte eigene Lehrendenbefragung. Durch den relativ kurzen zeitlichen Abstand kann der Einfluß tatsächlicher Veränderungen der Einschätzungen, die im Zeitverlauf stattgefunden haben könnten, als gering eingeschätzt werden. Allerdings waren die Frageformulierungen nicht komplett identisch und daher ein korrelativer Vergleich nicht angemessen. Dennoch zeigten sich in der Grundtendenz zumindest dieselben Antwortmuster bezüglich bestimmter Aspekte. So war die Einschätzung der beruflichen Zufriedenheit

aufgrund einer anderen Frageformulierung und einer anderen Skala nicht vergleichbar. Es ergab sich jedoch das gleiche Resultat, daß Professoren deutlich zufriedener waren als wiss. Mitarbeiter (vgl. Lenz/ Wolter/ Winter 2000: 44, Krempkow 2001: 23). Auch bei anderen in etwa vergleichbaren Aspekten zeigte sich eine weitgehende Übereinstimmung der Aussagen. So wurde z.B. die Kooperation mit Kollegen desselben Faches/ Fachbereiches in beiden Befragungen von den Lehrenden deutlich besser eingeschätzt als die Kooperation mit Kollegen anderer Fachbereiche (vgl. Lenz u.a. 2000: 44ff., Krempkow 2001: 25).

Als weitere Möglichkeit zur Überprüfung der Reliabilität von Einschätzungen zur Situation von Lehre und Studium kann wie bei Studierendenbefragungen auch bei den eigenen Lehrendenbefragungen die Testhalbierungsmethode als Übereinstimmung des gemittelten Urteils über mehrere Personen (Interraterreliabilität) herangezogen werden. Die Interraterreliabilität ist damit auch ein Maß für die Genauigkeit von Mittelwerten eines Beurteilungsgegenstandes (in diesem Fall der Situation von Lehre und Studium eines Fachbereiches/ einer Fakultät). Das wurde am Beispiel der TU Dresden untersucht.<sup>225</sup> Für die einzelnen Fakultäten ergaben sich hierbei Interraterreliabilitäten von .65 bis .90. Im Mittel aller Fakultäten betrug die Interraterreliabilität .79. Da die Fallzahlen für die befragten Lehrenden je Fakultät etwa 10 betragen (meist 8 bis 13, lediglich in drei Fällen über 20<sup>226</sup>), erscheint dieser Wert am ehesten mit der Interraterreliabilität der Lehrveranstaltungsmittel für 10 Studierende vergleichbar. Diese betrug .81 und lag damit in einer ähnlichen Größenordnung (vgl. Rindermann (2001: 131). In der Literatur wird häufig ein Wert von mindestens 0,8 gefordert. In einigen Fällen begnügt man sich jedoch auch mit einem Wert von 0,7 (vgl. z.B. Brosius 2002: 766). Man kann also davon sprechen, daß ab einer Mindestanzahl von etwa 10 Lehrenden die Lehrenden im Mittel zu ähnlichen Einschätzungen gelangen und sich daher keine Belege für eine grundsätzlich mangelnde Interraterreliabilität finden lassen.

---

<sup>225</sup> Hierfür war eine Transponierung bzw. Drehung der Datensätze notwendig (vgl. Rindermann 2001: 130, Fußnote 44). Verwendet wurden Intraklassenkorrelationen für relative Übereinstimmung (vgl. Rindermann 2001: 123ff.).

<sup>226</sup> Zieht man Zufallsstichproben von jeweils 10 Lehrenden aus den drei Fakultäten und berechnet hierfür die Interraterreliabilitäten, so fallen diese etwas geringer aus. Das Gesamtmittel der Interraterreliabilitäten aller Fakultäten würde sich hierdurch jedoch lediglich auf .77 verringern und bewegt sich damit in derselben Größenordnung. Da die Fakultäten in der Realität jedoch verschieden groß sind und bei größeren Fakultäten zur Erlangung von für die Hochschule (nach Fächern) repräsentativen Ergebnissen auch eine größere Anzahl von Lehrenden befragt werden sollte, wurden hier die Fallzahlen nicht zur besseren Vergleichbarkeit mit den Ergebnissen Rindermanns „künstlich“ verringert.

### 5.6.2 Vergleichbarkeit

Probleme in der Vergleichbarkeit würden wie auch bei anderen Befragungsergebnissen dann auftreten, wenn diese durch externe Einflussfaktoren „verzerrt“ sind, die nicht inhaltlich mit dem zu betrachtenden Gegenstand zusammenhängen. Solche potentiellen Biasvariablen wie unterschiedliche Sachkenntnis und Reife, Geschlecht, Interesse oder Leistungsbereitschaft, wie sie im Zusammenhang mit Studentenbefragungen genannt wurden, erscheinen hier nicht relevant. Beispielsweise eine unterschiedliche Sachkenntnis könnte lediglich dann vorliegen, wenn Lehrende erst relativ kurze Zeit in einem Studiengang tätig sind. Aber Lehrende sind meist länger an einer Hochschule als Studierende und schon nach 1 bis 2 Semestern dürften ausreichend Erfahrungen vorliegen, so daß sich die grundsätzliche Einschätzung dann nicht mehr wesentlich ändert. Zum Teil wurde im Zusammenhang mit studentischen Lehrbewertungen auch vermutet, daß weibliche Lehrende z.B. aufgrund eines höheren Einfühlungsvermögens eventuell eine andere Sicht auf die Studiensituation haben als männliche Lehrende. Empirische Befunde hierzu sind jedoch nicht bekannt. In den meisten Fächern dürften statistisch zuverlässige Resultate aber aufgrund der geringen Anzahl von weiblichen Lehrenden auch kaum zu erbringen sein. Ein unterschiedliches Interesse von Lehrenden am Fach oder ihre Leistungsbereitschaft kommen hier ebenfalls nicht als Biasvariablen in Betracht, da sie als Bestandteil des Dozentenengagements aufgefasst werden. Denkbar wäre lediglich, daß Lehrende, die sich überdurchschnittlich in der Forschung engagieren, zu anderen Einschätzungen gelangen als Lehrende, die dies nicht tun bzw. sich überdurchschnittlich in der Lehre engagieren. Eine Überprüfung dieser Hypothese anhand der Angaben der Lehrenden zum Zeitaufwand für die Bereiche Lehre und Forschung zeigte jedoch, daß es keine signifikanten Unterschiede bei einer solchen Gruppierung nach Forschungsengagement oder Lehrengagement gibt.

Da in den vorliegenden Publikationen auch keine Biasvariablendiskussion geführt wurde, wird offenbar davon ausgegangen, daß es kaum Verzerrungen der Urteile gibt. Da außerdem die Kritiker bereits die Validität von Lehrendenbefragungen bezüglich bestimmter Einschätzungen zur Situation von Lehre und Studium in Frage stellen und Validität Voraussetzung für sinnvolle Vergleiche ist, wurde die Vergleichbarkeit aber evtl. auch als nachrangiges Problem eingestuft und daher nicht weiter diskutiert.

Bezüglich der Einschätzungen zur Arbeitszufriedenheit wurden in den vorliegenden Veröffentlichungen Einflussfaktoren auf die Arbeitszufriedenheit nicht unter dem Stichwort

Biasvariablen diskutiert. Einige, wie z.B. der Fächereinfluß, können jedoch als solche aufgefasst werden. So wird die Arbeitszufriedenheit durchaus über verschiedene Hochschulen hinweg betrachtet und dabei über verschiedene Fächer hinweg zusammengefaßt (vgl. Enders/ Teichler 1995). Ein Problem der Vergleichbarkeit solcher zusammengefaßten Einschätzungen kann darin bestehen, daß die Situation an den Hochschulen stark fächergebunden wahrgenommen wird. Wenn die eine Hochschule ein eher geisteswissenschaftliches Profil hat, eine andere jedoch ein eher ingenieurwissenschaftliches Profil, kann ein Unterschied in der Arbeitszufriedenheit eventuell auch auf die unterschiedliche Fächerstruktur zurückzuführen sein. Daher sollte bei Vergleichen die Fächerstruktur beachtet oder die Auswertungen nach Fächern bzw. Fächergruppen getrennt vorgenommen werden.

Auch bei Vergleichen über verschiedene Hochschularten hinweg sind ggf. ähnliche Besonderheiten zu beachten. Insgesamt kann relativ allgemeingültig formuliert werden, daß bei Vergleichen immer die Kontextgebundenheit der Ergebnisse zu beachten ist. Grundsätzlich werden in den vorliegenden Studien jedoch keine Einwände gegen Vergleiche der Ergebnisse von Lehrendenbefragungen formuliert.

### **5.6.3 Weitere Anforderungen**

Als weitere Anforderungen an objektive wie an subjektive Indikatoren sind wiederum Zeitnähe, sinnvolle Anreizwirkung und Vermeidung unerwünschter Anpassungseffekte sowie ein vernünftiges Verhältnis von Erhebungsaufwand und erwartetem Nutzen zu nennen.

Zur Zeitnähe ist festzuhalten, daß diese stark vom Erhebungsturnus der Lehrendenbefragung abhängt. Geht man vom Status quo aus, so finden diese nur eher sporadisch und meist aus „privatem“ Forscherinteresse statt. Daher haben sie eine eher geringe Zeitnähe. Als Erhebungsaufwand fallen neben der Arbeitszeit einer entsprechend empirisch ausgebildeten Person jedoch nur Papierkosten an. Im Vergleich zu Absolventenbefragungen ist der Aufwand eher gering, im Vergleich zu vorliegenden Hochschulstatistiken ist er höher.

Dennoch könnten Lehrendenbefragungen ohne größeren Aufwand durchaus auch im Abstand von einigen Jahren regelmäßig im Zusammenhang mit größeren Reformvorhaben oder externen Evaluationen durchgeführt werden. Sofern das Ziel der Erhöhung der Arbeitszufriedenheit (zum Beispiel in bestimmten Aspekten wie der Leistungsgerechtigkeit des Gehalts) verfolgt wird und tatsächlich eine Erhöhung belegt werden kann, wäre der Nutzen einer solchen Befragung in diesem Zusammenhang anschaulich aufzuzeigen. Leider

ist eine Quantifizierung der Kosten einer geringen Mitarbeiterorientierung zum Beispiel durch Kommunikationsdefizite und daraus resultierende Reibungsverluste im Hochschulbereich und damit eine Gegenüberstellung von Kosten und Nutzen der Lehrendenbefragungen schwer möglich. Solche Befragungen können aber neben reinen Kosten-Nutzen-Erwägungen auch ein Instrument der Kommunikation sein, da hiermit die Hochschulleitungen die Möglichkeit haben, Offenheit zur Einbeziehung aller Mitarbeiter sowie Aufgeschlossenheit für Kritik und Anregungen zu demonstrieren.<sup>227</sup> Wie sich in der Vergangenheit bei der Durchführung von Lehrendenbefragungen zeigte, machen die Betroffenen bei glaubhafter Zusicherung von Anonymität durchaus rege davon Gebrauch, Anregungen und Verbesserungsvorschläge zu formulieren. Diese sollten dann allerdings nicht nur durch eine unabhängige Person ausgewertet und in anonymisierter Form weitergegeben werden, sondern auch tatsächlich aufgegriffen werden. Wo dies möglich ist, und sei es in Einzelbeispielen, könnte dies dann kommuniziert werden und mit zu einer Qualitätskultur im Sinne des TQM-Grundgedankens beitragen.

Außerdem erbringen solche Befragungen Informationen, die auf andere Weise nicht zu erlangen sind. Durch weitere Analysen, z.B. Zusammenhangsanalysen, wären aus Lehrendenbefragungen auch Ansatzpunkte für Akzeptanz erhöhende Maßnahmen bezüglich bestimmter Reformvorhaben abzuleiten.

Unerwünschte Anpassungseffekte durch Lehrendenbefragungen sind nicht zu erwarten. Vielmehr stellen sie wie auch andere subjektive Indikatoren eine Möglichkeit dar, Fehlanreizwirkungen von objektiven Indikatoren zu vermindern. Beispielsweise könnte einem ggf. zu vermeidenden, zu starken Anreiz zu konkurrenzbetonter Arbeitsweise entgegengewirkt werden, wenn die subjektive Einschätzung der Kooperation mit anderen Fächern/ Fachbereichen oder mit Kollegen im eigenen Fach/Fachbereich in die Bewertung einbezogen würde. Auch in einigen Unternehmen der Privatwirtschaft ist es üblich, subjektive Bewertungen heranzuziehen. Insbesondere gilt dies für Bereiche, in denen keine objektiven Maße für die gewünschte Zusammenarbeit verschiedener Bereiche gibt, zum Beispiel im Dienstleistungsbereich. Schließlich ist hierbei noch zu beachten, daß die Bewertungseinheiten groß genug zu wählen sind, um die Anonymität zu wahren.

---

<sup>227</sup> In einer Lehrendenbefragung für das Land Nordrhein-Westfalen beurteilten die befragten Professoren die Akzeptanz bestimmter Reformvorhaben „um so positiver, je informativer und partizipativer der Einführungsprozeß verlaufen ist“ (vgl. Minssen/ Wilkesmann 2003: 113).

#### **5.6.4 Fazit zur Eignung von Lehrendenbefragungen**

Insgesamt kann zur Eignung von Lehrendenbefragungen festgehalten werden: Lehrende schätzen bezüglich der Situation von Lehre und Studium nach den vorliegenden Ergebnissen offener Befragungen und halboffener Fragestellungen zwar weitgehend dieselben Aspekte als relevant ein, wie sie in Studierendenbefragungen verwendet werden. Die Ergebnisse standardisierter Befragungen mit geschlossenen Fragestellungen zeigen jedoch, daß Lehrende die Lehre insbesondere bei Aspekten, die Selbsteinschätzungen beinhalten, anders bewerten als Studierende und unabhängige Fremdgutachter. Bei anderen Aspekten, z.B. zu Rahmenbedingungen der Lehre, stimmen die Bewertungen der verschiedenen Befragengruppen besser überein. Da bei Selbsteinschätzungen generell von eingeschränkter Validität gesprochen wird und die vorliegenden Ergebnisse dies bestätigen, ist auch bei Lehrendenbefragungen bezüglich bestimmter Aspekte nur von eingeschränkter Validität auszugehen. Ähnliche Ergebnisse zeigen auch Untersuchungen zur Dimensionalität der Befragungsergebnisse. Bei Untersuchungen unter Einbeziehung der Einschätzungen zur Situation von Lehre und Studium konnten mittels Faktorenanalysen keine stabilen Faktorenstrukturen gefunden werden. Anders sieht dies bei separater Analyse der verschiedenen Aspekte der Arbeitszufriedenheit aus. Die hierzu existierenden Modelle konnten unter Berücksichtigung der hochschulspezifischen Besonderheiten weitgehend bestätigt werden und zeigten sich auch bei Untersuchungen über mehrere Hochschulen hinweg als relativ stabil.

Untersuchungen zur Reliabilität, die hier in Ermangelung besserer Untersuchungsmöglichkeiten nur exemplarisch zu in etwa vergleichbaren Fragestellungen bei zwei unabhängig voneinander durchgeführten Lehrendenbefragungen an derselben Hochschule durchgeführt werden konnten, ergaben in der Grundtendenz ähnliche Ergebnisse und damit zumindest keine Anhaltspunkte für eine mangelnde Retestreliabilität. Auch eine andere Möglichkeit zur Überprüfung der Reliabilität nach dem Grundprinzip der Testhalbierungsmethode zeigte, daß die Urteilerübereinstimmung über mehrere Personen hinweg gesehen (Interraterreliabilität) ab einer Mindestzahl von 10 befragten Lehrenden in etwa ähnlicher Höhe lag wie bei Studierendenbefragungen, wo sie in dieser Höhe als akzeptabel eingeschätzt wurde.

Zur Vergleichbarkeit und dabei evtl. zu beachtenden „verzerrenden“ Einflussfaktoren wurden in den vorliegenden Publikationen bisher kaum Einschränkungen formuliert. Wie

generell bei Vergleichen ist jedoch die Kontextgebundenheit der Ergebnisse zu beachten und bei hochschulübergreifenden Vergleichen der Arbeitszufriedenheit von Lehrenden beispielsweise die Fächerstruktur der Hochschulen zu beachten, da diese Einfluß auf die Arbeitszufriedenheit haben könnte.

Weitere Anforderungen wie die Zeitnähe der Befragungsergebnisse sind derzeit kaum erfüllt, da Lehrendenbefragungen meist nur sporadisch durchgeführt werden. Der Erhebungsaufwand bei Lehrendenbefragungen hält sich jedoch in Grenzen, so daß grundsätzlich eine größere Zeitnähe möglich wäre. Wenn eine Hochschule sich für eine breite Einbeziehung aller Lehrenden in einen Prozeß zur Förderung der Qualitätsorientierung entschließen sollte, könnten Lehrendenbefragungen ähnlich wie Mitarbeiterbefragungen in Unternehmen aber u.a. auch als Kommunikationsinstrument und als Instrument zur Identifikation von Ansatzpunkten für akzeptanzerhöhende Maßnahmen für bestimmte Reformvorhaben dienen, so daß dem Erhebungsaufwand ein mehrfacher Nutzen gegenüberstände. Da auch unerwünschte Anpassungseffekte durch Lehrendenbefragungen nicht zu erwarten sind und ihre Ergebnisse eher zur Vermeidung oder Minderung von Fehlanreizwirkungen dienen als selbst Fehlanreizwirkungen auslösen könnten, erscheinen sie grundsätzlich als geeignetes Instrument zur Förderung der Qualitätsorientierung von Hochschulen. Hierbei ist die Datenqualität der Bewertung von Aspekten der Arbeitszufriedenheit den bisher vorliegenden Untersuchungen zufolge aufgrund der höheren Validität besser als bei Einschätzungen der Situation von Lehre und Studium. Die Untersuchungsergebnisse stellen jedoch z.T. nur erste Analysen dar und können noch nicht verallgemeinert werden.

Letztlich ist die Entscheidung über den Einsatz von Lehrendenbefragungen aber noch stärker als bei anderen Befragungen auch eine Frage des Gesamtkonzeptes der Qualitätsorientierung, in dem die Lehrenden eine unterschiedlich große Rolle spielen können. Hinzu kommt, daß bei großen Widerständen gegen bestimmte Reformvorhaben und bei größeren, anonymen Hochschulen Lehrendenbefragungen vermutlich hilfreicher sein können und den Erhebungsaufwand eher lohnen als bei geringen Widerständen und kleineren Hochschulen, wo „ohnehin jeder jeden kennt“ und eine direkte Kommunikation bzw. die Einbeziehung eines größeren Anteils oder aller Lehrenden in anstehende Reformvorhaben leichter möglich ist.

## 6. Bedingungen erfolgreicher Institutionalisierung

*Verfolgt man die Diskussionen und Erfahrungsberichte an den Hochschulen Deutschlands in den vergangenen etwa zehn Jahren, so scheint es, daß neben solchen Voraussetzungen für bzw. Anforderungen an Leistungsbewertungen und Leistungsanreize wie grundsätzliche Angemessenheit der Konzepte und Modelle sowie dem Nachweis einer methodisch und „wissenschaftlich soliden“ Vorgehensweise bei der Erhebung und Auswertung von Kriterien bzw. Indikatoren noch weitere wesentliche Aspekte gibt, die auch als Bedingungen erfolgreicher Institutionalisierung bezeichnet werden können. Als Frage kann man hierzu formulieren: Was sind fördernde, was sind hemmende Bedingungen bei der Einführung und Umsetzung von Leistungsbewertungen und Leistungsanreizen an Hochschulen? Solche Bedingungen können bestimmte Interessen- und Machtkonstellationen innerhalb der Hochschulen sein (wie z.B. Interessenkonflikte zwischen Professoren und Studierenden, die ggf. mit der Gruppenstruktur der Hochschulen zusammenhängen), oder die von außen auf die Hochschulen wirken (z.B. massiver Druck der Politik auf die Hochschulen). Hierzu wird erstens die These diskutiert, ob zu starker Druck von außen eher zu Abwehrreaktionen bzw. Umgehungsstrategien führt und ob daher nicht eher versucht werden sollte, die (mehrheitliche) Akzeptanz der Betroffenen zu erlangen. Als weitere These wird daran anschließend diskutiert, ob dort, wo derzeit keine Akzeptanz herrscht, bestimmte Rahmenbedingungen so verändert werden könnten, daß über die Zeit hinweg gesehen größere Akzeptanz bei den Betroffenen erlangt werden kann. Eine solche Möglichkeit wird im Zusammenhang von Akzeptanz und Arbeitszufriedenheit gesehen. Dies führt wiederum zu der Frage, inwieweit sich aus der empirischen Analyse von Determinanten der Arbeitszufriedenheit Ansatzpunkte ableiten lassen, welche Aspekte eher förderlich und welche eher hemmend sind, Leistungsbewertungen und Leistungsanreize selbst und die mit ihnen verfolgten Ziele zu erreichen.*

Gliederung:

- 6. Bedingungen erfolgreicher Institutionalisierung
  - 6.1 Machtkonstellationen, Interessen und Akzeptanz
  - 6.2 Lehrendenbefragungen als Instrument der Akzeptanzuntersuchung
  - 6.3 Die Eignung vorhandener Lehrendenbefragungen
    - 6.3.1 Entstehungshintergrund der bundesweiten Lehrendenbefragungen
    - 6.3.2 Zur Anlage und Repräsentativität der bundesweiten Lehrendenbefragungen
    - 6.3.3 Ergebnisse der Lehrendenbefragungen zur Arbeitssituation
    - 6.3.4 Wunsch und Wirklichkeit beim Zeitbudget der Lehrenden an deutschen Hochschulen
    - 6.3.4 Einschätzung hochschulpolitischer Reformvorhaben an deutschen Hochschulen
  - 6.4 Akzeptanz von Reformvorhaben und Arbeitszufriedenheit bei Lehrenden in Dresden, Frankfurt/ Oder und Zittau-Görlitz
    - 6.4.1 Anlage und Durchführung der eigenen Lehrendenbefragungen
    - 6.4.2 Betrachtung von Rücklauf und Repräsentativität der Befragungen
    - 6.4.3 Akzeptanz hochschulpolitischer Reformvorhaben
    - 6.4.4 Berufliche Zufriedenheit der Lehrenden am Beispiel der TU Dresden
    - 6.4.5 Das Zeitbudget und sein Zusammenhang mit der Arbeitszufriedenheit
  - 6.5 Determinanten der beruflichen Zufriedenheit
    - 6.5.1 Konzepte zu Determinanten der beruflichen Zufriedenheit
    - 6.5.2 Empirische Überprüfung der Konzepte
    - 6.5.3 Subgruppenanalysen
    - 6.5.4 Bivariate und multivariate Zusammenhangsanalysen
    - 6.5.5 Fazit zu Determinanten der beruflichen Zufriedenheit
  - 6.6 Fazit zu Bedingungen erfolgreicher Institutionalisierung

## 6.1 Machtkonstellationen, Interessen und Akzeptanz

Grundsätzlich kann man die Überlegungen zur Betrachtung des Zusammenhanges von Machtkonstellationen und Akzeptanz zu den derzeit diskutierten Reformvorhaben zur Einführung von Leistungsbewertungen und Leistungsanreizsystemen in zwei Ansätze einteilen. Während beim ersten gedanklichen Ansatz von der Gruppenstruktur an den Hochschulen ausgegangen wird, sehen andere Autoren den Systemangehörigen als einzelnes Individuum (unabhängig von seiner Gruppenzugehörigkeit) und dessen Akzeptanz als Ausgangspunkt für ihre Überlegungen. Dies wird hier als zweiter Ansatz betrachtet. Sicherlich gibt es weitere Ansätze, die Bedingungen erfolgreicher Institutionalisierung von Reformvorhaben zu untersuchen. So wurde im Zusammenhang mit Leistungsbewertungen und Anreizmodellen sowie Experteninterviews bereits darauf hingewiesen, daß u.a. die Transparenz von Kriterien und Verfahren, die Gewichtung des Leistungsanteils und nicht zuletzt die Information und Partizipation der Betroffenen ebenfalls wichtige Rahmenbedingungen sind. Nachfolgend soll jedoch von den beiden o.g. Grundüberlegungen zur Akzeptanz und deren Zusammenhang mit Rahmenbedingungen wie z.B. Machtkonstellationen innerhalb der Hochschulen ausgegangen werden, da sie Gelegenheit bieten, diese mit Ergebnissen empirischer Untersuchungen zu konfrontieren und dabei die Perspektive der Betroffenen einzubeziehen.

Als Beispiel für den ersten Ansatz, der von der Gruppenstruktur an den Hochschulen ausgeht, kann Rindermann (2001: 306ff.) eingeordnet werden. Er sieht die Machtkonstellation der Gruppe der Professoren innerhalb der Hochschulen als wesentlichen zu beachtenden Punkt bei Bewertungen zur Qualität der Hochschullehre und der Leistungen von Hochschullehrenden (vgl. auch Keller 2000: 155ff.).<sup>228</sup> Folgt man dieser Argumentation, könnte vermutet werden, daß das Vorhandensein oder Nichtvorhandensein bestimmter Interessen- oder Machtverhältnisse innerhalb der Hochschulen eine Bedingung erfolgreicher Institutionalisierung ist.<sup>229</sup>

---

<sup>228</sup> Keller (2000) diskutiert ausführlich den Einfluß von Machtstrukturen auf Veränderungen an den deutschen Hochschulen in den vergangenen Jahrzehnten (wenngleich mit deutlich erkennbarer Sympathie für Studentenbewegungen). Bei Keller finden sich auch zahlreiche weitere Literaturhinweise.

<sup>229</sup> Anhaltspunkte für den Einfluß von Machtkonstellationen bestimmter Gruppen bietet auch der Band von Goedegebuure u.a. (1993), in dem im internationalen Kontext untersucht wurde, wer welche Rolle bei der Gesetzgebung, der Kontrolle über Studienprogramme, Forschungsprogramme, Hochschulmanagement und Mittelverteilung spielte. Hierbei wurde dies allerdings weit über hochschulinterne Machtkonstellationen hinaus betrachtet.

Beispielhaft für den zweiten Ansatz, der eher vom einzelnen Individuum (unabhängig von seiner Gruppenzugehörigkeit) ausgeht, können Autoren genannt werden, die sich seit längerer Zeit auch in verschiedenen Projektzusammenhängen mit der Einführung von Leistungsbewertungen und Leistungsanreizen beschäftigen. In diesen Zusammenhängen wurden Aspekte wie ein starkes öffentliches Interesse an der Thematik, ein massiver Druck der Politik auf eine öffentliche Rechenschaftslegung der Hochschulen<sup>230</sup> oder gar das Infragestellen des Weiterbestandes einer Einrichtung genannt<sup>231</sup>, die die einzelnen Individuen dazu bringen könnten, Reformmaßnahmen zu akzeptieren (vgl. z.B. Pellert 1999, Schenker-Wicki 1996, 2002). Andere Autoren gehen zwar ebenfalls vom Individuum aus, betonen aber, es sei „keineswegs so, daß ‘viel Druck hilft viel’ als Prinzip für Organisationsentwicklungsmaßnahmen gilt. (...) Statt dessen erreicht man mehr, wenn man Ziele fordert, die sich eben noch innerhalb des ‘Akzeptanzteppichs’ der Systemangehörigen befinden“ (vgl. Wottawa 2001: 154). Als gemeinsames Element dieser Diskussion kann man formulieren, daß ein gewisser Druck von außen als Bedingung erfolgreicher Institutionalisierung<sup>232</sup> gesehen wird. Darauf soll nachfolgend noch ausführlicher eingegangen und dabei auch diskutiert werden, welches Maß hier ggf. als angemessen gesehen werden kann und wovon dies möglicherweise abhängt. Zunächst soll jedoch der erste Ansatz zur Machtkonstellation der Gruppe der Professoren und ihrer Interessenlage innerhalb der Hochschulen als wesentlichen zu beachtenden Punkt betrachtet werden.

---

<sup>230</sup> Schenker-Wicki (1996: 46) berichtete, daß in allen bis dato erfolgreich umgesetzten Fällen die Veränderungen nicht freiwillig eingeführt wurden, sondern „auf massiven Druck von außen“. Nach anfänglichem Widerstand wurden sie letztlich akzeptiert (vgl. Schenker-Wicki 2002).

<sup>231</sup> So führt auch Pellert (1999: 298) als Hintergrund (warum dies überhaupt möglich war) für einen erfolgreichen Fall von Maßnahmen der Organisationsentwicklung an, daß es eine vorangehende heftige Diskussion über den Weiterbestand der Einrichtung gab. Dadurch sei die „Einsicht in den Nutzen von organisationsentwicklerischen Maßnahmen erhöht“ worden. Im Zuge von Überlegungen zur Übertragbarkeit des Projektes fügt sie hinzu, daß das Interesse an Veränderung v.a. von innen kommen müsse, aber auch „leichter Außendruck hilfreich“ wäre (vgl. Pellert 1999: 304ff.).

<sup>232</sup> Der Begriff Institutionalisierung wird in Anlehnung an Shils (1975: 69ff.) verwendet, der diesen Begriff für die „Eingliederung einer Wissenschaft in ein akademisches System und nicht nur die Einführung in eine einzige Universität“ prägte (vgl. Shils 1975: 141 – wobei er mit Wissenschaft die Soziologie als Fachdisziplin bezeichnete). Hier geht es allerdings nicht darum, das Themenfeld Leistungsbewertung und Leistungsanreize als Fachdisziplin zu begründen (obwohl es für Leistungsbewertungen i.S.v. Evaluation mit der Einrichtung von entsprechenden Studiengängen in Saarbrücken und Kassel durchaus Bestrebungen in diese Richtung gibt). Vielmehr geht es darum, einen hohen Grad (institutioneller) Verankerung zu beschreiben, der sich nicht – und hier liegt die Gemeinsamkeit – in der Einführung von organisatorisch wenig verankerten Pilotprojekten an einzelnen Hochschulen erschöpft, sondern eine starke Eingliederung im Hochschulsystem meint. Eine *erfolgreiche* Institutionalisierung im Zusammenhang mit Leistungsbewertungen und Leistungsanreizen sollte darüber hinaus die intendierten Wirkungen in höherem Maße erreichen als nicht intendierte Wirkungen. Beispielsweise wäre eine Institutionalisierung dann nicht erfolgreich, wenn zwar nach außen hin – um ggf. Sanktionen zu entgehen – Zielvereinbarungen, die Einführung von leistungsorientierter Mittelvergabe o.ä. Maßnahmen beschlossen und offiziell verfolgt, letztlich aber unterlaufen oder umgangen werden.

### **6.1.1 Machtkonstellationen und Interessen als Institutionalisierungsbedingungen**

Beispielhaft soll hier die Argumentation von Rindermann aufgegriffen und hinsichtlich möglicher Ansatzpunkte für Schlußfolgerungen zu Bedingungen erfolgreicher Institutionalisierung diskutiert werden. Rindermann (2001: 306ff.) betrachtet in Anlehnung an Bourdieu (1988: 274ff.) die Mitgliedergruppen Professoren, Mittelbau und Studierende als ordinale Statusordnung, in welcher die Gruppen unterschiedliche Interessenlagen hätten und daher unterschiedliche hochschulpolitische Positionen vertreten. Rindermann verweist dabei auch auf die Zeit der Studentenbewegung in Frankreich und Deutschland, in der der Mittelbau sich aufgrund seiner Unzufriedenheit mit den Karriereaussichten eher mit studentischen Positionen solidarisiert hätte, während Professoren meist eine konträre Haltung eingenommen hatten (vgl. auch Luhmann 1992). Als Belege dafür, daß es nach wie vor einen Zusammenhang zwischen Gruppenstatus und Akzeptanz hochschulpolitischer Stellungnahmen gibt, führt Rindermann die Lehrendenbefragung von Enders und Teichler (1995) an, die in mehreren Aspekten deutliche Unterschiede zwischen Professoren und Mittelbau aufzeigte. Dies zeigte sich auch in einer Imagestudie der TU Dresden (vgl. Lenz u.a. 2000) und kann in nachfolgenden Abschnitten noch anhand eigener Lehrendenbefragungen untersucht werden. Nach Rindermann (2001: 307) spiegeln diese Einschätzungen „die unterschiedlichen Privilegien verschiedener Statuspositionen wider“. Auch Enders und Teichler (1995) verwiesen darauf, daß die Statusunterschiede in Deutschland im Vergleich zu anderen Ländern besonders ausgeprägt seien.

Sicherlich ist die heutige Situation aufgrund der sehr viel pragmatischeren Haltung vieler Studierender seit den 1990er Jahren (vgl. z.B. Bargel 1996) nicht mit der Studentenbewegung Ende der 1960er Jahre vergleichbar und es kann wohl nicht von einer generellen Solidarisierung des akademischen Mittelbaues mit Studierenden gesprochen werden. Dennoch liegt es aufgrund der genannten Ergebnisse nahe, davon auszugehen, daß die Akzeptanz hochschulpolitischer Reformvorhaben nicht unabhängig vom Status der Betroffenen ist.

Rindermann (2001: 309) ging noch einen Schritt weiter und formulierte, daß bei Lehrenden (und dabei insbesondere Professoren) bei der Einschätzung der Zuverlässigkeit von Lehrbewertungen durch Studierende „Erkenntnisinteressen und statusbezogene Interessen (...) miteinander konfliktieren können“. Weiter heißt es dort: „Wissenschaftler, die zumeist

auch Hochschullehrer sind, der dominierenden Fraktion angehören, untersuchen hierbei, ob Studenten – an der Universität der dominierten Fraktion zuordenbar – sie selbst und ihre Lehre adäquat (objektiv, meßgenau, valide, ...) beurteilen können.“ Zurückgehend auf die Durkheimsche Forderung von 1895 nach einer Distanzierung der Wissenschaftler von Alltagserfahrungen thematisiere auch Bourdieu (1988: 31) die Problematik des Spezialfalls hochschulbezogener Forschung: „Eine soziale Welt zu thematisieren, in die man persönlich verstrickt ist...“.<sup>233</sup> Rindermann verwies allerdings auch darauf, daß ein Problem der Distanzierung nicht nur bezogen auf die Hochschulforschung, sondern generell bestehen kann. Dies sei aber grundsätzlich keine neue Erkenntnis. Er führte hierzu als Beispiele u.a. amerikanische Forschungsergebnisse z.B. zum Zusammenhang von sozialer Herkunft der Forscher und ihrer Position im Streit um die Rolle von Vererbung und Sozialisation an. Aus diesen Überlegungen leitete er die Hypothese ab, daß es einen Zusammenhang zwischen Hierarchieposition und Methodenwahl, Resultaten und Interpretationen auch in der Lehrevaluationsforschung geben könne. Die Gruppe der Professoren sehe demzufolge Bewertungen der Lehre eher skeptisch und komme tendenziell häufiger zu dem Ergebnis, daß solche Lehrbewertungen nicht geeignet sind, die Qualität der Lehre zu beurteilen. Mittelbauangehörige würden dagegen Bewertungen der Lehre durch Studenten eher befürworten und sich mit diesen tendenziell solidarisieren, so Rindermann weiter. Die Diskussion um die Einschätzung der Validität und Zuverlässigkeit studentischer Lehrbewertungen könne man daher auch als ein „Nebenkriegsschauplatz“ oder zumindest als einen Ausdruck eines Machtgefälles zwischen verschiedenen Mitgliedergruppen an den Hochschulen sehen. Zur Überprüfung dieser Hypothese untersuchte Rindermann sämtliche verfügbare Veröffentlichungen deutschsprachiger Wissenschaftler zu Fragestellungen der Bewertung der Lehre durch Studenten an Hochschulen (n = 206) im Zeitraum von 1967 bis 2001. Diese wurden einer inhaltsanalytischen Auswertung unterzogen, inwieweit die Autoren studentische Lehrbewertungen grundsätzlich als valide oder nicht valide einschätzten (oder ob sie sich nicht hierzu positionierten). Hierbei wurde der Begriff der Validität relativ weit

---

<sup>233</sup> Nach Bourdieu (1988: 49) geht „jede Stellungnahme zur sozialen Welt (...) aus vom Standpunkt der Wahrung und Vergrößerung der mit dieser Stellung verbundenen Macht. Universitäre Macht beruht nach Bourdieu (1988: 142) „im wesentlichen auf der Herrschaft über die Instrumente zur Reproduktion der Körperschaft“. Dabei ist zu betonen (vgl. Bourdieu 1988: 243), daß dies implizit, also meist nicht bewusst geschehe: „Was als eine Art organisierte kollektive Verteidigung der Professorenschaft angesehen werden kann, ist nichts weiter als das Resultat von Tausenden selbständiger und gleichwohl aufeinander abgestimmter Reproduktionsstrategien, von Tausenden von Handlungen, die wirksam zur Erhaltung der Körperschaft beitragen, weil sie aus einem gleichsam sozialen Erhaltungstrieb hervorgehen.“

gefaßt und auch Aussagen zur Beeinflußbarkeit des studentischen Urteils durch Biasvariablen und zur Datenqualität allgemein in die Analyse einbezogen. Eine zweite wichtige Kategorie war für Rindermann, inwieweit die Autoren die Einsetzbarkeit studentischer Lehrbewertungen über eine reine Feedback-Funktion hinaus (z.B. für Veranstaltungsvergleiche, Rankings, Berufungsentscheidungen) als geeignet einschätzen. Anschließend wurde dies getrennt nach Statusgruppen ausgewertet.

Im Ergebnis zeigte sich, daß studentische Lehrbewertungen zwar überwiegend (78 zu 38 Fälle) als valide eingeschätzt werden. Professoren beurteilten sie aber in der Tat häufiger als nicht valide (25 zu 14), während Mittelbauangehörige sie deutlich häufiger als valide einschätzten (43 zu 8).<sup>234</sup> Ein ähnliches Bild ergab sich bei der Bewertung der Einsetzbarkeit. Die Korrelationen zwischen Status und Beurteilung studentischer Lehrbewertungen betragen zwischen .33 und .46.

Diese Zusammenhänge können wiederum durch bestimmte Rahmenbedingungen eher verstärkt oder abgeschwächt sein. Als Beispiel für einen der von Rindermann (2001: 323ff.) genannten Aspekte, welche die Beurteilung studentischer Lehrbewertungen moderieren können, soll der unterschiedliche Methodisierungsgrad<sup>235</sup> der Publikationen dargestellt werden. So geht er davon aus, daß in Arbeiten mit einem hohen Methodisierungsgrad der Statureffekt weniger stark auftritt. Dies sei insbesondere bei quantitativen empirischen Arbeiten unter Hinzuziehung von standardisierten Korrelationskoeffizienten und/ oder Signifikanztests der Fall, für die wissenschaftliche Konventionen existieren und in denen der Interpretationsspielraum im Gegensatz z.B. zu reinen Stellungnahmen (ohne Bezug auf wissenschaftliche Konventionen) aufgrund der Überprüfbarkeit der Interpretationen an den

---

<sup>234</sup> Auf die Darstellung der anderen Statusgruppen Studierende (z.B. Abschlussarbeiten, Stellungnahmen von Fachschaftsvertretern), Externe (z.B. Journalisten) und derjenigen Gruppen, bei denen Status bzw. Bewertung unklar ist, wird hier verzichtet.

<sup>235</sup> Mit Methodisierungsgrad ist hier v.a. die Orientierung an Konventionen gemeint, wie sie in empirischen Methodenlehrbüchern genannt werden. Rindermann (2001: 311ff.) nahm außerdem noch weitere, detailliertere Auswertungen zu Einflußfaktoren vor, die die Akzeptanz moderieren konnten, so z.B. den unterschiedlichen Publikationsmedien, in denen je nach Herausgebern überwiegend pro- oder contra Positionen vertreten wurden. Da das Grundmuster des Gruppenbezugs der Akzeptanz dasselbe ist, wird hier jedoch auf eine umfassende Darstellung verzichtet und auf die Originalpublikation verwiesen.

<sup>237</sup> Wie bereits im Kapitel 3 erläutert, war die Rücklaufquote bei den Professoren mit 12 Prozent sehr gering und zudem nach Befragtengruppen sehr unterschiedlich (vgl. Minssen/ Wilkesmann 2003: 111). Zudem erfolgte kein Hinweis, ob die Repräsentativität der Befragung bezüglich der Grundgesamtheit anhand wichtiger, potentiell die Ergebnisse beeinflussender Merkmale der Befragten überprüft wurde (z.B. Fächerzusammensetzung). Da davon auszugehen ist, daß dies nicht geschah und die Ergebnisse bei diesem geringen Rücklauf stark verzerrt sein können und diese Befragung zudem eine andere Zielsetzung hatte (neben der Akzeptanz v.a. die Wirkungen der indikatorisierten Mittelverteilung an Nordrhein-westfälischen Hochschulen zu untersuchen) sowie aufgrund des Zeitpunktes ihrer Veröffentlichung bei der Konzeption eigener Befragungen keine Berücksichtigung finden konnte, wird sie hier nicht weiter diskutiert.

Ergebnissen eingeschränkt ist. Tatsächlich zeigt sich für den akademischen Mittelbau, daß bei von Rindermann als streng „wissenschaftlich“ eingestuften Arbeiten ein ausgewogeneres Verhältnis herrscht (18 zu 7 Fälle) als in „Stellungnahmen“, in denen alle für Validität plädierten. Bei den Professoren, von denen allerdings nur sehr wenige als streng „wissenschaftlich“ eingestufte Arbeiten vorlagen, zeigt sich kein solcher Effekt.

Hier gab es bei beiden Arten von Publikationen hohe Anteile derjenigen, die gegen die Validität plädierten. Inwieweit die Einschätzungen beim Mittelbau durch einen zusätzlichen Einfluß der Befristung moderiert werden, hatte Rindermann nicht thematisiert. So erscheint es vorstellbar, daß unbefristet beschäftigte Mittelbauangehörige tendenziell anders urteilen als befristet Beschäftigte. Dies wäre allerdings nur mit großem zusätzlichem Aufwand erhebbar gewesen, da die Befristung üblicherweise in Publikationen nicht angegeben wird. Da insgesamt der akademische Mittelbau meist als wissenschaftlich eingestufte Arbeiten und Professoren überwiegend Stellungnahmen veröffentlichten und dabei der Mittelbau stärker für Validität plädiert, ist der Anteil der Beurteilungen als „valide“ in Arbeiten mit hohem Methodisierungsgrad insgesamt höher.

Zusammenfassend kann zum Zusammenhang zwischen Hierarchieposition und Beurteilung bzw. Akzeptanz von Bewertungen der Lehre durch Studierende formuliert werden, daß diese den Ergebnissen der empirischen Untersuchung Rindermanns zufolge nicht unabhängig vom Status der Urteilenden sind und dies sich stärker in solchen Publikationen bemerkbar macht, in denen weniger streng wissenschaftliche Maßstäbe angelegt werden und größerer Interpretationsspielraum besteht.

Doch was folgt daraus für eine Betrachtung als Bedingung erfolgreicher Institutionalisierung als Indikator für Leistungsbewertungen bis hin zu Leistungsanreizen? Rindermann (2001: 330) appelliert hier v.a. an die Einhaltung allgemein anerkannter Qualitätssicherungsmechanismen der wissenschaftlichen Methodik, wie z.B. aktive Rezeption des Forschungsstandes, Anwendung angemessener Methodik, stringente Interpretationen und Schlußfolgerungen. Er spricht sich eher gegen einen politisierenden Ansatz aus und sieht im Aufzeigen des o.g. Zusammenhanges sogar eine mögliche Gefährdung wissenschaftlicher Standards, „da Standpunkte nicht mehr fachwissenschaftlich, sondern womöglich nur noch auf der Ebene der Wissenschaftssoziologie oder hochschulpolitischer Auseinandersetzungen diskutiert werden“ (vgl. Rindermann 2001: 330). Weiter heißt es, eine Lösung liege sicherlich nicht darin, „daß Perspektivität wissenschaftlicher Arbeit nicht minimierbar und deshalb zu akzeptieren oder gar gutzuheißen sei“ (Rindermann 2001: 334). Allerdings nennt

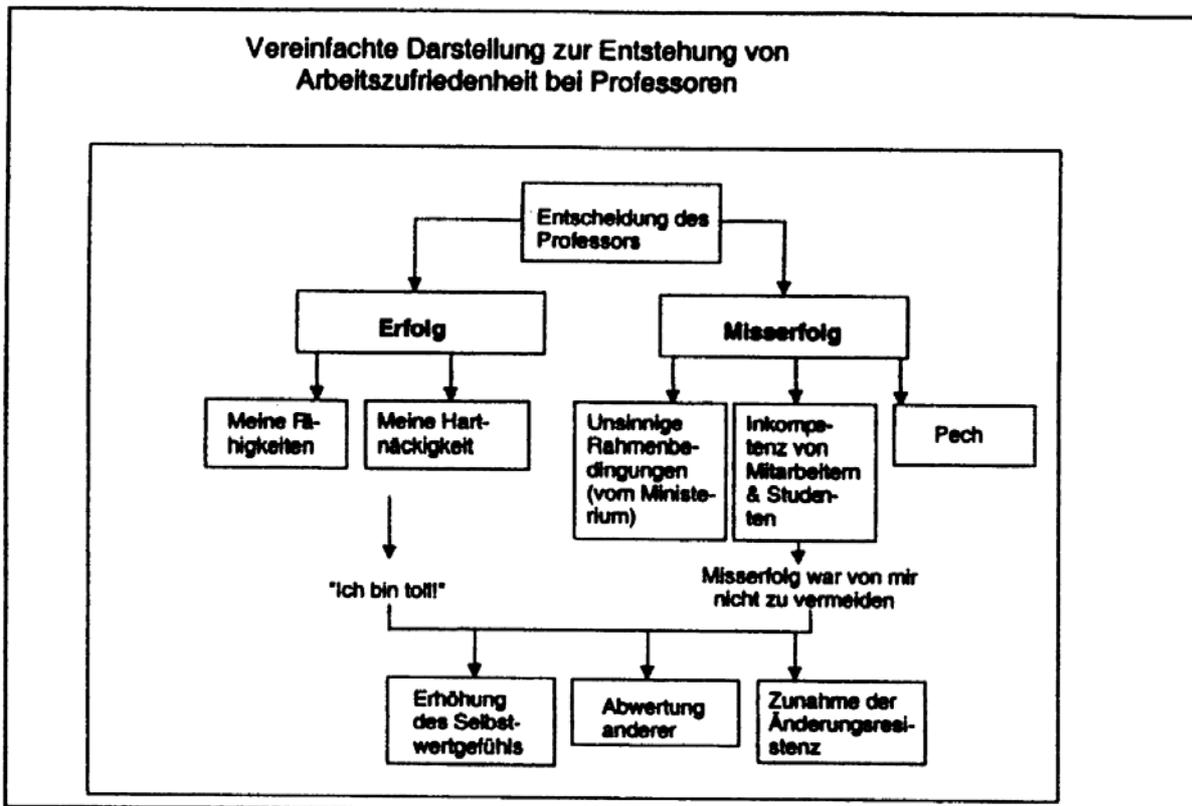
Rindermann selbst keine (andere) Lösungsmöglichkeit und hofft offenbar, daß sich im Laufe der Zeit die besseren, d.h. hier die empirisch fundierteren Argumente durchsetzen werden. Dies stünde jedoch im Widerspruch zur These, daß die Mitgliedergruppe der Professoren tendenziell zu vermeiden trachtet, was ihre Machtposition gefährden könnte und demzufolge Bewertungen der Lehre eher ablehnen sowie häufiger solche Methoden wählen bzw. Resultate veröffentlichen und diese dahingehend interpretieren, daß sie nicht geeignet sind, die Qualität der Lehre zu beurteilen. Und genau dies sollten die Auswertungen von Rindermann ja gerade zeigen, daß die Nutzung wissenschaftlicher Methoden und Einhaltung der entsprechenden Standards aufgrund der vorhandenen Machtstrukturen bei der Klärung der Frage der Datenqualität studentischer Lehrbewertungen nur bedingt geschieht und z.B. durch Veröffentlichung von Stellungnahmen (anstelle von stärker an methodischen Konventionen orientierten Publikationsformen) umgangen werden kann. Bleibt man in dem gedanklichen Ansatz Rindermanns und geht nach dessen Logik vom Gruppenstatus als wesentliche erklärende Variable der Einschätzung der Validität bzw. der Akzeptanz studentischer Lehrbewertungen zur Leistungsbewertung aus, erscheint es wenig aussichtsreich, allein über Argumentationen (und ohne Veränderung von Rahmenbedingungen) die Akzeptanz derjenigen Statusgruppe für Reformmaßnahmen erlangen zu wollen, deren Interessen dadurch am meisten gefährdet wären. Daher erscheint es zur Auflösung des Problems hilfreich, den Ausgangspunkt der Überlegungen, die Statuszugehörigkeit als erklärende Variable anzusehen, zu verlassen oder um andere Ansätze zu erweitern.

Als mögliche Alternative böte sich nun – würde man den eingangs dieses Kapitels genannten Einschätzungen von Schenker-Wicki (1996, 2002) und Pellert (1999) folgen – ein erhöhter externer Druck an, z.B. seitens der Politik (bis hin zum Infragestellen des Weiterbestandes einer Einrichtung – vgl. Pellert 1999). Offenbar war es möglich, in bestimmten Fällen auch auf diese Weise die Akzeptanz der Betroffenen zu erlangen. Es geht aber in Deutschland in den nächsten Jahren um einen breiten, wenn nicht gar flächendeckenden Einsatz und es erscheint kaum möglich, flächendeckend den Weiterbestand von Hochschulen in Frage zu stellen. Zudem ist gerade bei Leistungsbewertungen und Leistungsanreizen – dies zeigen die Erfahrungen der vergangenen Jahre in Deutschland und auch in anderen Ländern, die Gefahr des Unterlaufens und Umgehens besonders groß, so daß die angestrebten Ziele wahrscheinlich nicht erreicht würden.

### 6.1.2 Akzeptanz als Bedingung erfolgreicher Institutionalisierung

Diese Überlegungen führen zum bereits weiter oben erwähnten zweiten Ansatz, der den Systemangehörigen als einzelnes Individuum (unabhängig von Status und Machtposition) und dessen Akzeptanz als Ausgangspunkt für die Überlegungen zugrundelegt. Als Beispiel für diesen Ansatz soll Wottawa (2001) herangezogen werden. Wottawa (2001: 153ff.) geht davon aus, daß bei Angehörigen einer Organisation, die seit langem nach dem Prinzip der Input- oder Verhaltenskontrolle geführt werden und von denen man verlangt, daß sie sich auf eine andere Art von Systemsteuerung einzustellen haben, dies aus psychologischen Gründen massive Widerstände auslösen muß. Kurz zusammengefaßt formuliert Wottawa (2001: 154): „Wie bei jedem Versuch der Verhaltensänderung ist es für einen solchen Reformansatz unverzichtbar, daß überhaupt eine Verursachung eines suboptimalen Zustandes durch die eigene Person gesehen wird (Prinzip der Hedonistischen Verzerrung).“

#### Internale Attribuierung von Erfolg und Mißerfolg bei Professoren (nach Wottawa)



Quelle: Wottawa (2001: 154)

Als wesentliches Problem hierbei kann angesehen werden, daß der Wunsch von außenstehenden Instanzen nach einer Verhaltensänderung mit dem Prinzip des Selbstwertschutzes kollidieren könne, was zu einer emotionalen Abwehr des Veränderungswunsches führen müsse (und der Personen oder Institutionen, die einen solchen repräsentieren). Massiver Druck von außen erhöht dann nur die Widerstände, so daß das Prinzip „Viel Druck hilft viel“ nicht zum Erfolg führen kann. Wottawa (2001: 154) fügt hinzu: „... schon gar nicht bei einer relativ starken Rechtsstellung und Existenzsicherung von Systemangehörigen“. Er plädiert deshalb dafür, gerade jene Ziele zu fordern, die sich eben noch innerhalb des „Akzeptanzteppichs“ der Systemangehörigen befinden. Auf diese Weise könnte nach Wottawa eine schrittweise Verhaltensänderung erreicht werden, ohne das diese mit dem Prinzip des Selbstwertschutzes kollidiert.

Wottawas Ansatz setzt allerdings das Wissen voraus, welche Ziele gerade noch akzeptiert werden oder – um mit Wottawas Worten zu sprechen – wo der „Akzeptanzteppich“ endet und wo durch Verletzung des Prinzips des Selbstwertschutzes auch die Arbeitszufriedenheit beeinträchtigt werden könnte. Dieses Wissen kann möglicherweise bei in Umstrukturierungen und Reformen erfahrenen Hochschul- oder Fakultätsleitungen vorhanden sein.

Wottawa (2001: 156) ist hier allerdings selbst skeptisch und formuliert: „Es erfordert daher ein geschicktes (am besten professionelles) Vorgehen von Organisationsentwicklung, um ein optimales Ergebnis mit möglichst geringen Reibungsverlusten zu erzielen.“ Außerdem verweist Wottawa auf Qualitätsmanagementmodelle, auf denen man bei der Organisationsentwicklung aufbauen könne, die allerdings bisher in den Hochschulen nur vereinzelt angewendet werden (vgl. auch Abschnitt zu Qualitätsmanagementmodellen in dieser Arbeit).

Daß aber auch professionelle Hochschulentwickler und -Berater nicht immer genau wissen, wo der „Akzeptanzteppich“ endet, zeigen Einschätzungen von Experten, die nicht in jedem Fall mit Ergebnissen von Lehrendenbefragungen an Hochschulen übereinstimmen müssen (vgl. z.B. Unterschiede in den Einschätzungen zwischen Professoren, Dekanen und Hochschulleitungen in Minssen/ Wilkesmann 2003 sowie Äußerungen von Experten in den eigenen Experteninterviews im vierten Kapitel dieser Arbeit und den nachfolgend dargestellten Ergebnissen eigener Lehrendenbefragungen). Ähnliche Erfahrungen wurden aber auch bereits bei Mitarbeiterbefragungen in Unternehmen gemacht (vgl. z.B. Bartel-Lingg 1996). Fehleinschätzungen der Akzeptanz können aber gravierende Folgen haben. Laut Wottawa (2001: 157) „zeigt sich immer wieder, daß Verhaltensvorschriften bei fehlender Akzeptanz durch die Universitätsangehörigen keinen Nutzen bringen, manchmal

sogar gegenläufige als die eigentlich angestrebten Resultate zur Folge haben“. Eine korrekte Einschätzung der Akzeptanz hätte also nach dem Ansatz von Wottawa eine hohe Bedeutung. Vergleicht man die Ansätze von Rindermann (2001) und Wottawa (2001), so zeigt sich neben den unterschiedlichen Ausgangspunkten (Gruppe vs. Individuum) auch eine unterschiedliche Akzentuierung bezüglich der Akzeptanz. Rindermann formulierte hierzu, daß die Perspektivität wissenschaftlicher Arbeit nicht akzeptiert oder gar gutgeheißen werden sollte. Daraus läßt sich ableiten, daß in letzter Konsequenz die Akzeptanz der Reformvorhaben bei den betroffenen Professoren für ihn nicht der Schwerpunkt seiner Überlegungen ist, sondern eher die wissenschaftliche Begründbarkeit der Eignung von studentischen Lehrbewertungen. Wottawa hingegen sieht die Akzeptanz bei den Betroffenen (Professoren und Mitarbeitern) als wesentliche Bedingung erfolgreicher Institutionalisierung von hochschulpolitischen Reformvorhaben.

Da sich Wottawas Ansatz mit den Überlegungen und Erfahrungsberichten zur Einführung von Leistungsanreizen deckt (allerdings bisher v.a. im Bereich der Privatwirtschaft), in denen von mehreren Autoren die Wichtigkeit der Akzeptanz bei den Betroffenen bzw. Mitarbeitern betont wird (vgl. Abschnitt zu Leistungsanreizmodellen in dieser Arbeit) und dies auch mit Ergebnissen empirischer Untersuchungen übereinstimmt (vgl. Grüning 2002, Speckbacher/ Bischof 2000) sollen nachfolgend Ansätze, die die einzelnen Systemangehörigen bzw. Individuen und ihre Akzeptanz als Bedingung erfolgreicher Institutionalisierung untersuchen, als Ausgangspunkt genutzt werden. Auf andere Ansätze wird aus den genannten Gründen nicht weiter eingegangen.

Für eine möglichst repräsentative und vergleichbare Informationsgewinnung zur Einschätzung der aktuellen Akzeptanz bei den Systemangehörigen erscheinen ergänzend zu Experteninterviews schriftliche standardisierte Befragungen von Lehrenden geeignet. An dieser Stelle soll noch einmal darauf hingewiesen werden, daß diese Methode der Befragung trotz relativ gesehen geringerem Aufwand (als z.B. bei einer höheren Anzahl von Experteninterviews) und insgesamt höherer Zuverlässigkeit generell die Gefahr einer geringeren Tiefenschärfe in sich birgt. Sie kann und soll also die Einschätzungen von Experten nicht ersetzen, sondern mit vergleichbaren, standardisierten Informationen aus der Perspektive der Betroffenen ergänzen.

Neben der eingeschränkten Tiefenschärfe liegen die Grenzen einer Akzeptanzuntersuchung mit Hilfe von Lehrendenbefragungen aber auch darin, daß eine wirkliche Prognose künftiger Akzeptanz unter sich im Zeitverlauf verändernden Bedingungen nicht möglich ist. Auch

wenn Lehrendenbefragungen eine hohe Zuverlässigkeit im Sinne von Retestreliaibilität bescheinigt werden kann, so gilt die „Reproduzierbarkeit“ der Ergebnisse unter der Annahme konstanter Bedingungen, die durch geplante Veränderungen gerade nicht gegeben sind.

Die Veränderlichkeit der Akzeptanz impliziert aber, daß sie sich sowohl verbessern als auch verschlechtern könnte und beeinflubar ist. Auch Wottawas Ansatz mit der schrittweisen Umsetzung von Reformvorhaben unter Beachtung des „Akzeptanzteppichs“ geht von einer im Zeitverlauf veränderbaren Akzeptanz aus. Betrachtet man Akzeptanz als Bedingung erfolgreicher Institutionalisierung eines Reformvorhabens, so sollte ein negatives Ergebnis einer Akzeptanzuntersuchung sicherlich allein kein Grund sein, einen begonnenen Reformprozeß aus diesem Grunde komplett abzubrechen. Es könnte vielmehr der Entscheidungsunterstützung in der Hinsicht dienen, ein Reformvorhaben mit Aspekten zu beginnen oder den Schwerpunkt zunächst auf solche Aspekte zu legen, die zum Befragungszeitpunkt bereits mehrheitlich (oder noch am ehesten) Akzeptanz finden. Andererseits sollte ein positives Ergebnis einer Akzeptanzuntersuchung sicherlich kein Anlaß sein, die Betroffenen („Sie haben ja ihre grundsätzliche Akzeptanz geäußert“) nicht weiter in die Umsetzung von Reformvorhaben einzubeziehen, da ansonsten die Akzeptanz vermutlich unter solchen Umständen auch relativ schnell deutlich geringer werden kann. Zusätzlich zu den Ergebnissen direkter Befragung zur Akzeptanz können weiterführende Analysen des Zusammenhanges der Akzeptanz mit intervenierenden Variablen der Suche nach Ansatzpunkten dienen, wie die Akzeptanz evtl. mittel- und langfristig zu verbessern wäre.

Bezieht man diese Überlegungen zurück auf den Zusammenhang von Machtkonstellationen, Interessen und Akzeptanz, so eröffnen Ansätze wie der von Wottawa gedankliche Lösungsmöglichkeiten, wie hochschul- und gesellschaftspolitisch gewollte bzw. gesetzlich als notwendig erachtete Reformvorhaben an den Hochschulen unter Beachtung der Akzeptanz bei den Betroffenen schrittweise umgesetzt werden könnten. Folgt man diesem Ansatz, kann die Akzeptanz bei den Betroffenen als wesentlicher Ansatzpunkt für eine Diskussion als Bedingung erfolgreicher Institutionalisierung betrachtet und Lehrendenbefragungen als Instrument der Akzeptanzuntersuchung diskutiert werden. Akzeptanz soll hierbei breit verstanden werden und nicht nur als bloße Zustimmung zu einem Reformvorhaben gesehen werden, sondern auch Aspekte einbeziehen, die von mit dieser Thematik befassten Experten als (für die langfristige Akzeptanz) bedeutsam eingeschätzt wurden (vgl. Abschnitt zu

Experteninterviews in dieser Arbeit), wie die Einschätzung der Fairneß bzw. Leistungsgerechtigkeit. Da solche Aspekte in Befragungen zur Arbeitszufriedenheit erfasst werden und die Arbeitszufriedenheit in den Überlegungen von Wottawa (2001:154) ebenfalls eine Rolle spielt, sollen neben der Akzeptanz im engeren Sinne auch Aspekte der Arbeitszufriedenheit in die Diskussion einbezogen werden. Die Einbeziehung der Arbeitszufriedenheit erscheint außerdem von Bedeutung, da – wie bereits einleitend kurz dargestellt – davon ausgegangen wird, daß eine langfristige Sicherung bzw. Förderung von Akzeptanz nur möglich ist, wenn bei Reformvorhaben die Auswirkungen auf die Arbeitszufriedenheit insgesamt berücksichtigt werden. Dies soll jedoch Gegenstand der Diskussion nachfolgender Abschnitte sein.

## **6.2 Lehrendenbefragungen als Instrument der Akzeptanzuntersuchung**

Wenn neben der Einschätzung der Situation von Lehre und Studium sowie den Rahmenbedingungen der Lehre bzw. Arbeitsbedingungen durch Lehrende auch die Akzeptanz von geplanten Veränderungen und die Arbeitszufriedenheit von Lehrenden bedeutsam ist, bietet es sich an, Lehrendenbefragungen auch als Instrument der Akzeptanzuntersuchung zu diskutieren.

Lehrendenbefragungen können einerseits für eine direkte Befragung zur Feststellung der aktuellen Akzeptanz genutzt werden, ohne allerdings daraus auf eine künftige Akzeptanz unter veränderten Bedingungen schließen zu können. Andererseits kann versucht werden, mit Hilfe von Lehrendenbefragungen genauere Informationen über Ursachen von Akzeptanz oder Nicht-Akzeptanz zu erlangen und Ansatzpunkte zu ihrer Beeinflussung zu gewinnen.

Beispielsweise könnte, sofern die Orientierungen der von den geplanten Veränderungen Betroffenen als relativ stabil eingeschätzt werden, aus bestimmten Orientierungen möglicherweise zumindest die Grundrichtung der Auswirkungen eines Reformvorhabens auf die Akzeptanz abgeleitet werden. Wenn außerdem in Anlehnung an Wottawas Vorstellungen ein moderierender Effekt der Arbeitszufriedenheit auf die Akzeptanz gesehen wird und gleichzeitig die Erhöhung der Arbeitszufriedenheit auch als ein Ziel von Qualitätsmanagement- und Anreizmodellen genannt wird, sollten ebenso die möglichen Auswirkungen auf die Arbeitszufriedenheit der Lehrenden in die Überlegungen einbezogen werden. So könnte aus dem Vorhandensein einer geringeren Arbeitszufriedenheit bei eher lehrorientierten Professoren und Mitarbeitern sowie gleichzeitigem Vorliegen starker Zusammenhänge der Arbeitszufriedenheit mit der beruflichen Anerkennung daraus in gewissen Grenzen geschlußfolgert werden, daß die Arbeitszufriedenheit der Lehrorientierten bei einer höheren Anerkennung ihrer Lehrleistungen wahrscheinlich höher ausfiele.

Der Zusammenhang von Arbeitszufriedenheit und Akzeptanz wird hierbei darin gesehen, daß Reformmaßnahmen, die dazu führen, daß Lehrende mit ihrer beruflichen Situation insgesamt unzufriedener sind, auch schwerer Akzeptanz finden dürften. Umgekehrt könnten Reformmaßnahmen, die vielleicht zunächst keine große Akzeptanz finden, aber auf längere Sicht zu einer höheren Arbeitszufriedenheit führen, langfristig auch eher Akzeptanz finden.

Der Zusammenhang zwischen Arbeitszufriedenheit und Akzeptanz ist hierbei als ein über die Zeit vermittelter zu sehen.

So könnte z.B. ein Zusammenhang der Lehrorientierung von Lehrenden mit ihrer Arbeitszufriedenheit dann auch ein Ansatzpunkt zur Erhöhung der Akzeptanz sein. Weitere Ansatzpunkte wären Einzelaspekte der beruflichen Situation, die besonders stark mit der Arbeitszufriedenheit insgesamt in Zusammenhang stehen. Bevor jedoch eine direkte Akzeptanzuntersuchung mit Hilfe von Bewertungen von hochschulpolitischen Reformvorhaben und eine indirekte Untersuchung möglicher Auswirkungen von Reformvorhaben über die Arbeitszufriedenheit mittels Zusammenhangsanalysen erfolgen kann, soll die Eignung vorhandener Lehrendenbefragungen hierfür diskutiert werden.

### **6.3 Die Eignung vorhandener Lehrendenbefragungen**

In Deutschland wurden im letzten Jahrzehnt bereits mehrere Lehrendenbefragungen durchgeführt. Inhalte waren neben der Einschätzung der Lehr- und Studiensituation und der Zufriedenheit mit verschiedenen Aspekten der beruflichen Situation auch Fragen zur Akzeptanz von Reformvorhaben der Hochschulpolitik (vgl. Kopp/ Weiß 1993, Enders/ Teichler 1995, Schaeper 1995, Winter 1999, Minssen/ Wilkesmann 2003<sup>237</sup>) sowie zur Akzeptanz geplanter Fortbildungsangebote (vgl. Voigt 2000, Pötschke 2004). Aufgrund ihres Inhaltes konnten diese Befragungen bereits eine Grundlage sein für die Diskussion der Eignung von Lehrendenbefragungen an Hochschulen in Deutschland.

Bislang lagen jedoch – angesichts der hohen Dynamik der Veränderungen im Hochschulbereich – kaum aktuelle Ergebnisse zur Bewertung von Reformvorhaben der Hochschulpolitik vor und insbesondere an ostdeutschen Hochschulen existierten generell nur vereinzelt Lehrendenbefragungen, die außerdem in wichtigen Teilen nicht mit anderen vergleichbar waren (vgl. z.B. Winter 1999, Voigt 2000).

Um einen besseren Überblick über die Situation an ostdeutschen Hochschulen zu gewinnen und im Zusammenhang mit dem Thema dieser Arbeit zusätzlich spezielle Forschungsfragen<sup>238</sup> zu klären, wurden daher an drei ostdeutschen Hochschulen (TU Dresden, Europa-

---

<sup>238</sup> Eine Forschungsfrage hierzu lautete zunächst einmal entsprechend der Auswahl der Hochschulen, worin bzw. wie stark sich Hochschulen bezüglich der Einschätzung ihrer Arbeitsbedingungen in der Lehre, im Arbeitsklima, in der Einschätzung des Fortbildungsbedarfs im Bereich der Lehre und letztlich in der Bewertung hochschulpolitischer Reformvorhaben unterscheiden (können). Entsprechend dem speziellen Interesse der Arbeit am Thema Leistungsbewertung und Leistungsanreizsysteme wurden weiterhin zusätzlich zu den bereits bundesweit verwendeten Reformvorhaben zur Akzeptanz von Peer review und Studenten-

Universität Viadrina, Fachhochschule Zittau-Görlitz) eigene Lehrendenbefragungen durchgeführt. Diese bilden ebenfalls eine Diskussionsgrundlage.

Zunächst sollen im ersten Abschnitt kurz die Anlage der Studien und ausgewählte Ergebnisse der bundesweiten Lehrendenbefragungen vorgestellt werden, damit eine Einschätzung ihrer Aussagekraft möglich wird.<sup>239</sup> Eine umfassendere Darstellung ist bereits in einschlägigen Publikationen erfolgt.

Bisher nicht außerhalb der jeweiligen Hochschulen publiziert wurden die eigenen Lehrendenbefragungen an ostdeutschen Hochschulen. Daher sollen die Anlage der Studien und ausgewählte Ergebnisse anschließend im zweiten Abschnitt dieses Kapitels ausführlicher dargestellt werden. Sie stellen eine Grundlage auch für die Diskussion der Akzeptanz von Leistungsbewertungen und Leistungsanreizen und Möglichkeiten ihrer Beeinflussung dar, die im vierten Abschnitt dieses Kapitels erfolgt.

### **6.3.1 Entstehungshintergrund der bundesweiten Lehrendenbefragungen**

In Deutschland wurden im letzten Jahrzehnt zwei bundesweite Lehrendenbefragungen durchgeführt, eine von Kopp und Weiß, die andere von Enders und Teichler. Entstehungshintergrund der Studie von Kopp und Weiß (1993) war zunächst eine Befragung von Hochschullehrern der kanadischen Provinz Quebec zu den Folgen der in den 1980er Jahren einsetzenden Sparpolitik an Universitäten und damit in Zusammenhang stehenden Managementmaßnahmen (vgl. Haché/ Crespo 1989). Die Existenz einer ähnlichen Problemlage in den Hochschulsystemen anderer westlicher Industrienationen begründete das Interesse an einer vergleichenden Untersuchung der Frage, welche Anpassungsmuster sich auf Seiten der Universitäten bei verschärfter Mittelknappheit unter differenten strukturellen und organisatorischen Bedingungen herausbilden (vgl. Kopp/ Weiß in Enders/ Teichler 1995: 105). Unter besonderer Berücksichtigung der Situation in der Bundesrepublik

---

befragungen (vgl. Enders/ Teichler 1995 sowie Kopp/ Weiß 1993) noch Absolventenbefragungen sowie explizite Fragestellungen zur Kopplung von Leistungsanreizsystemen mit (subjektiven) Bewertungen ergänzt. Außerdem wurden Fragestellungen zur Einschätzung der Lehr- und Studiensituation verwendet, die bereits in Studenten- und Absolventenbefragungen Anwendung fanden, um somit auch eine Gegenüberstellung der verschiedenen Perspektiven zu ermöglichen. Weitere oder noch konkretere Fragestellungen zum Themenbereich Leistungsbewertung und Leistungsanreizsysteme wären in den Hochschulleitungen nicht "konsensfähig" und die Lehrendenbefragungen in dieser Form dann nicht möglich gewesen.

<sup>239</sup> Da sich der Fragebogen der an den Universitäten in Dresden, Frankfurt/ Oder und Zittau-Görlitz durchgeführten Lehrendenbefragungen in mehreren Fragestellungen an den bundesweiten Lehrendenbefragungen von Kopp/ Weiß (1993) sowie Enders/ Teichler (1995) orientierte, werden lediglich diese beiden Untersuchungen vorgestellt. Andere Hochschullehrer- oder Professorenbefragungen eigneten sich weit weniger als

Deutschland und als Ergebnis vorab durchgeführter Interviews mit Hochschullehrern erfolgte letztlich eine Beschränkung auf ein relativ enges Segment von Fragen, die diese Untersuchung thematisch strukturierten und auch für die eigenen in Dresden, Frankfurt/Oder und Zittau-Görlitz durchgeführten Lehrendenbefragungen im Vordergrund standen:

- Einschätzung der Arbeitssituation (Ausstattung, (arbeits-)klimatische Aspekte, Zufriedenheit)
- Einstellung zu verschiedenen hochschulpolitischen Perspektiven und Reformvorhaben

Auch die Hochschullehrerbefragung von Enders und Teichler (1995) wurde im angelsächsischem Raum initiiert (von der Carnegie Foundation for the Advancement of Teaching (Princeton, N.J.). Das Wissenschaftliche Zentrum für Berufs- und Hochschulforschung der Universität/ Gesamthochschule Kassel führte die Erhebung und Auswertung für Deutschland (alte Bundesländer) durch. Wichtige Themenkomplexe dieser Befragung waren:

- Berufsweg und berufliche Position
- Beschäftigungsbedingungen
- Zeitverteilung und berufliche Prioritäten
- Arbeitsaufgaben
- Leitung und Verwaltung der Hochschule
- Internationale Aspekte des wissenschaftlichen Lebens
- Hochschule und Gesellschaft

Die Themenkomplexe bei Enders und Teichler (1995) sind zwar teilweise identisch mit der Professorenbefragung von Kopp/ Weiß (1993), die einzelnen Fragestellungen waren jedoch oft anders formuliert, so daß Vergleiche der Ergebnisse eher nur zu den Grundaussagen sinnvoll sind. Außerdem wurde bei Enders und Teichler (1995) im Unterschied zu Kopp/ Weiß (1993) auch der akademische Mittelbau einbezogen.

---

Orientierung für die angestrebten Untersuchungsziele, da sie entweder nicht bundesweit durchgeführt wurden oder deutlich andere Ziele verfolgten (vgl. z.B. Schaeper in Enders/Teichler 1995: 130ff.).

### 6.3.2 Zur Anlage und Repräsentativität der bundesweiten Lehrendenbefragungen

Bei Kopp und Weiß (1993) wurden unter Berücksichtigung von Fächergruppen und Regionen 2457 Professoren westdeutscher Universitäten nach dem Zufallsprinzip ausgewählt und im Juni 1992 angeschrieben. 678 von diesen antworteten, 534 Fragebögen (22 Prozent) waren letztlich verwertbar. Nicht in die Untersuchung einbezogen wurden die Fächer Medizin, Sport und Kunstwissenschaften. Der Anteil der Befragten in den einzelnen Fächergruppen entsprach in etwa der zu diesem Zeitpunkt aktuellsten verfügbaren Statistik des Statistischen Bundesamtes von 1991, wobei die Ingenieurwissenschaften zu 8 Prozentpunkten über- und Mathematik/ Naturwissenschaften zu 11 Prozentpunkten unterrepräsentiert waren. Die anderen untersuchten strukturellen Merkmale (Geschlecht, Besoldungsgruppen) stimmten in ähnlichem Maße überein. Obgleich die Ergebnisse der Befragung wegen der relativ geringen Rücklaufquote einige Unsicherheiten in sich bergen könnten, dürften sie aufgrund der Merkmalsverteilung der Befragten für die untersuchten Fächer als weitgehend repräsentativ angesehen werden.<sup>240</sup> Als Vergleichsbasis für die eigenen Befragungen an den Universitäten Dresden und Frankfurt/ Oder sind vor allem die Ergebnisse zu den Arbeitsbedingungen und zur faktischen und als optimal erachteten Arbeitszeit nach Tätigkeiten sowie die Einschätzung hochschulpolitischer Reformvorhaben von Interesse.

Nach Angaben von Enders und Teichler (1995) wurden in ihrer Studie im Sommersemester 1992 insgesamt etwa 10.000 Fragebögen an Lehrende und Forschende aller hauptberuflich wissenschaftlich tätigen Statusgruppen aller Fächergruppen und aller Hochschularten außer Verwaltungshochschulen versandt. Hierbei wurden 60 ausgewählte Fachbereiche bzw. Fakultäten an 18 Hochschulen einbezogen (5 Universitäten, 3 Technische Hochschulen, 2 Gesamthochschulen, 2 Kunst- und Musikhochschulen, 6 Fachhochschulen). Die Analyse stützte sich auf 2801 Fragebögen (28 Prozent Rücklaufquote). 682 Personen (24 Prozent der Befragten) waren Universitätsprofessoren, 1875 Personen (67 Prozent) Angehörige des Mittelbaus an Universitäten. Den Rest bildeten Fachhochschulangehörige. Nach Vergleichen zu vorliegenden Daten der Personalstruktur kommen die Autoren zu dem Schluß, daß die Studie weitgehend repräsentativ ist (vgl. Enders/ Teichler 1995: 15).

Da als Vergleichsbasis für die eigenen Untersuchungen an den Universitäten Dresden und Frankfurt/ Oder ebenfalls vor allem die Ergebnisse zu den Arbeitsbedingungen, zur

---

<sup>240</sup> Entscheidend für die Güte oder allgemeinsprachlich Repräsentativität einer Stichprobe ist nicht allein die Rücklaufquote, sondern insbesondere die Übereinstimmung wichtiger Merkmale der Stichprobe mit der Grundgesamtheit (vgl. z.B. Schneekloth/ Leven 2003: 51, Diekmann 2002: 368ff., Kromrey 1998: 259ff.).

faktischen und als optimal erachteten Arbeitszeit nach Tätigkeiten sowie die Einschätzung hochschulpolitischer Reformvorhaben von Interesse sind, sollen diese nachfolgend kurz vorgestellt werden.

### **6.3.3 Ergebnisse der Lehrendenbefragungen zur Arbeitssituation**

Bei den Ergebnissen der Fragestellungen von Kopp/ Weiß (1993) zu den Arbeitsbedingungen fällt vor allem die Entwicklung der Betreuungsrelation auf. Diese lag 1980 bei 41,4 Studenten je Professor, 1991 waren es knapp 59 (vgl. Kopp/ Weiß 1995: 115). Hohe Anteile negativer Wertungen erhielten die Ausstattung mit wissenschaftlichen Hilfskräften, mit wissenschaftlichen Mitarbeitern und die Ausstattung der Bibliotheken (47 bis 42 Prozent negative Wertungen). Dennoch waren die Professoren mit der überwiegenden Anzahl der erfragten Aspekte zufrieden. Die berufliche Anerkennung in der Gesellschaft (89 Prozent zufrieden und sehr zufrieden), die Beziehungen zu den Studenten (84 Prozent) und das Gehalt (67 Prozent) erhielten hierbei am häufigsten positive Nennungen.

Insgesamt gesehen hatte sich nach Einschätzung der befragten Professoren die Situation in ihrer Organisationseinheit in den letzten 5 Jahren für mehr als die Hälfte nicht verändert. 27 Prozent registrierten eine Verschlechterung, 21 Prozent eine Verbesserung.

An erster Stelle der Ergebnisse von Enders und Teichler (1995) zu Arbeitsbedingungen und -zufriedenheit soll genannt werden, daß die Unterschiede zwischen Universitätsprofessoren und Mittelbauangehörigen so groß ausfielen wie in keinem anderen der über ein Dutzend untersuchten westlichen Industrienationen. Deshalb erfolgte die Auswertung getrennt nach Statusgruppen.

Die Universitätsprofessoren äußerten sich mit 64 Prozent positiven Nennungen insgesamt überwiegend zufrieden zu ihrer beruflichen Situation. Der Mittelwert der Zufriedenheit lag bei 2,4 (auf einer Skala von 1 bis 5). Nur in 4 von 13 Ländern (Hongkong: 2,2/ Australien, Schweden und USA: 2,3) einer international vergleichend angelegten Studie äußerten sich die Professoren geringfügig zufriedener (vgl. Enders/ Teichler 1993: 36f.). Brasilien, Japan und Korea lagen gleichauf, Großbritannien, Chile und Mexiko darunter. Befragt zu einzelnen Aspekten ihrer beruflichen Situation, bestätigte sich dieses Bild: Sowohl bei den ihnen zur Verfügung stehenden materiellen als auch bei den personellen Ressourcen gehören die deutschen Universitätsprofessoren - neben ihren Kollegen aus den USA, Hongkong, den Niederlanden und Schweden - zu denjenigen, die in keinem der

angesprochenen Bereiche überwiegend Mängel konstatieren. Zum Gehalt äußerten sie sich mit 65 Prozent positiven Nennungen nach den Niederlanden und Hongkong am zufriedensten (vgl. Enders/ Teichler 1993: 32). Kopp/ Weiß (1993) kamen mit 65 Prozent auf dasselbe Ergebnis.

Deutlich anders sieht die Arbeitssituation dagegen für den akademischen Mittelbau aus: Dieser ist mit einem Mittelwert von 3,1 zur Einschätzung seiner beruflichen Zufriedenheit insgesamt am unzufriedensten von allen befragten westlichen Industrienationen. Lediglich in Chile ist der akademische Mittelbau ähnlich unzufrieden.

Im Vergleich zu den Professoren äußerten sich vom akademischen Mittelbau mit 32 Prozent nur halb so viele "(sehr) zufrieden". Zwei Problembereiche werden genannt, die dieses hohe Maß an Unzufriedenheit erklären: Erstens ist die Beschäftigungssituation mit 80 Prozent befristet Beschäftigten in keinem anderen der untersuchten Länder so prekär. Zweitens sind die Mittelbauangehörigen in Deutschland besonders abhängig. In keinem anderen der untersuchten Länder hat der Mittelbau so wenig Gestaltungsspielraum in der Lehre und in der Forschung, während es für die Professoren genau umgekehrt ist. In keinem anderen untersuchten Land haben die Universitätsprofessoren mehr Gestaltungsspielraum, was sowohl Lehre als auch Forschung betrifft (vgl. Enders/ Teichler 1995: 25).

### **6.3.4 Wunsch und Wirklichkeit beim Zeitbudget der Lehrenden an deutschen**

#### **Hochschulen**

Befragt zur aufgewandten und als optimal erachteten Arbeitszeit, wandten lt. Kopp und Weiß (1993) die Professoren nach eigenen Angaben durchschnittlich 38 Prozent ihrer Zeit für die Lehre auf (34 Prozent galten als optimal), 27 Prozent für die Forschung (42 Prozent wären optimal), 9 Prozent für Prüfungen (7 Prozent optimal) sowie jeweils 10 Prozent für Hochschulgremien und andere Dienstaufgaben (6 bzw. 5 Prozent optimal). Weitere durchschnittlich 6 Prozent der Arbeitszeit wurden für Gutachten benötigt (5 Prozent optimal). Es wurde auf erhebliche interindividuelle und interfachliche Unterschiede hingewiesen. Nach Vorlesungszeit und vorlesungsfreier Zeit wurde nicht differenziert.

Laut Enders/ Teichler (1995) unterscheiden sich die Statusgruppen Universitätsprofessoren und akademischer Mittelbau beim Zeitbudget erheblich. So gaben die Universitätsprofessoren an, im Jahr durchschnittlich rund ein Drittel ihrer Arbeitszeit für die Lehre aufzuwenden (43 Prozent in der Vorlesungszeit, 20 Prozent in der vorlesungsfreien Zeit).

Etwa zwei Fünftel ihrer Arbeitszeit rechnen sie Forschungstätigkeiten zu (29 Prozent in der Vorlesungszeit, 53 Prozent in der vorlesungsfreien Zeit). Das Zeitbudget der Mittelbauangehörigen sei noch stärker durch die Forschung als durch die Lehre dominiert.

Gegen diese Ergebnisse kann eingewendet werden, daß sie subjektive Einschätzungen darstellen und wenig valide bzw. zuverlässig sind. Wie weiter oben erläutert, wurden jedoch für die Universitätsprofessoren in der Studie von Kopp und Weiß, (die im selben Zeitraum durchgeführt wurde und die nicht nach Vorlesungszeit und vorlesungsfreier Zeit differenzierte,) relativ ähnliche Werte angegeben. (Würde man den Zeitanteil für Lehre und Prüfungen sowie für Forschung und Gutachten addieren, so fielen die Ergebnisse sehr ähnlich aus.) Zu vermuten ist daher, daß dies von den Lehrenden bei Enders und Teichler jeweils mit unter Lehre bzw. Forschung eingeordnet wurde. Auch in einer weiteren, kurze Zeit später durchgeführten Studie von Schaeper, deren Ergebnisse zwar von der Breite her das Fächerspektrum abdecken, aber nicht bundesweit repräsentativ sind, fallen die Ergebnisse zum Zeitbudget ebenfalls von der Grundtendenz her ähnlich aus (vgl. Schaeper in: Enders/ Teichler 1995: 142). Deshalb kann davon ausgegangen werden, daß zumindest die Grundaussagen zum Zeitbudget hinreichend gesichert sind. Hinzuzufügen ist, daß wenn man Studien aus früheren Jahren hinzuzieht, der Anteil der Arbeitszeit, der während der Vorlesungszeit für die Forschung verwandt wird, im Laufe der Jahre scheinbar kontinuierlich immer größer geworden ist (1976/77: 23 Prozent, 1983/84: 27 Prozent, 1990/91: 28 Prozent - vgl. Enders/ Teichler 1995: 20). Angesichts dieser Daten könnte man davon ausgehen, daß die hochschulpolitischen Bemühungen zur Erhöhung des Stellenwertes der Lehre, die in den 1990er Jahren forciert wurden, den Trend zur Erhöhung des Zeitanteils für die Forschung höchstens zu bremsen vermocht haben. Allerdings kann an dieser Stelle nicht geklärt werden, ob sich die subjektiven Einschätzungen der Befragten evtl. auch aufgrund einer anderen Wahrnehmung ihrer Tätigkeit im Zeitverlauf änderten. Vor diesem Hintergrund kann dennoch auch eine andere Interpretation der Angaben der Professoren zu ihren beruflichen Interessen als bei Enders und Teichler vorgenommen werden.<sup>241</sup>

---

<sup>241</sup> Womöglich liegt hier aber auch ein Irrtum der Autoren vor. Denn die in der Tabelle angegebenen Daten sind andere als die im Text formulierten (vgl. Enders/ Teichler 1995: 16 und 17). So präferieren die Professoren, befragt zu ihren Präferenzen in Forschung und Lehre, insgesamt nur zu 35 Prozent eher die Lehre, insgesamt zu 66 Prozent jedoch eher die Forschung.

Im einzelnen gaben 7 Prozent an "mehr in der Forschung", weitere 59 Prozent "in beidem, mehr in der Forschung". 5 Prozent gaben an "mehr in der Lehre" und weitere 30 Prozent gaben an "In beidem, mehr in der Lehre". Stellt man die Angaben gegenüber, präferieren insgesamt nur 35 Prozent eher die Lehre, insgesamt 66 Prozent eher die Forschung.

Im Text formulierten die Autoren, daß die Befunde *dagegen* sprechen, daß "das deutsche Universitätssystem eine einseitige Forschungsorientierung seiner Professoren fördere", da etwa ein Drittel stärker die Forschung und ein Drittel stärker die Lehre betont". Wenn die Angaben in der Ergebnistabelle zugrundegelegt werden, sind die deutschen Universitätsprofessoren aber mit zwei Dritteln Forschung deutlich stärker forschungs- als lehrorientiert. Daß in anderen Ländern des europäischen Raumes die Lehre z. T. noch weniger präferiert wird, ist hierbei von nachrangiger Bedeutung, da die Autoren die Anteile der Lehr- und Forschungsorientierten innerhalb Deutschlands gegenüberstellten.<sup>242</sup>

Schließlich zeigt sich die stärkere Forschungs- als Lehrorientierung auch in den Angaben der deutschen Universitätsprofessoren zum als optimal erachteten Zeitaufwand für Lehre bzw. Forschung im Vergleich zum tatsächlichen Zeitaufwand (vgl. Kopp/ Weiß 1995: 113 Schaeper: 1995: 148). Alle dort diesbezüglich publizierten Ergebnisse zeigen, daß die Forschung in den Präferenzen deutlich vor der Lehre eingeordnet wird.

#### **6.3.4 Einschätzung hochschulpolitischer Reformvorhaben an deutschen Hochschulen**

Bei der Einschätzung hochschulpolitischer Reformvorhaben (nur in der Lehrendenbefragung von Kopp und Weiß 1993) fällt die relativ hohe Zustimmung zu solchen Reformvorhaben auf, die auf die Reformbereitschaft der befragten Hochschullehrer schließen lassen. So sprechen sich knapp zwei Drittel für eine Neuorganisation des Studiums nach angelsächsischem Vorbild aus. Die Einführung von Kurzstudiengängen und differenzierten Abschlüssen wurde hierbei ausdrücklich mit unter solche Maßnahmen gezählt (vgl. Kopp/ Weiß 1995: 123). Bemerkenswert ist außerdem, daß sich drei Viertel der Professoren für "mehr marktwirtschaftliche Elemente im Hochschulwesen" aussprechen. Der Einführung leistungsorientierter Mittelzuweisung stimmen sogar vier Fünftel zu. Mehrheitlich befürwortet wird auch eine regelmäßige Bewertung der Lehrleistungen durch Studierende, des Studienangebotes durch Experten, die Veröffentlichung von Leistungskennzahlen sowie eine stärkere fachlich-inhaltliche Profilbildung der Universitäten. An Grenzen stößt die Reformbereitschaft der Professoren bei der Hochschulfinanzierung, Bildungsgutscheinen und einer Erhöhung der Anzahl privater Hochschulen. Diese sind –obwohl mit immerhin 45

---

<sup>242</sup> Ein weiterer Widerspruch findet sich in den Angaben zum für die Lehrtätigkeit aufgewandten Zeitanteil in der vorlesungsfreien Zeit: Hier ist im Text von 26 Prozent die Rede, während in der Tabelle nur 20 Prozent angegeben wurden. Da dies leicht ein Übertragungsfehler sein kann und die Angabe 26 Prozent rechnerisch mit weiteren Zahlen übereinstimmt, wird nachfolgend diese Zahl verwendet.

bzw. 47 Prozent Zustimmung nur knapp – nicht mehrheitsfähig. Lediglich private Stiftungslehrstühle finden mit 60 Prozent Zustimmung eine Mehrheit.

Wenngleich die Befragung von Kopp/ Weiß als weitgehend repräsentativ eingeschätzt werden kann, sind die Ergebnisse sowohl aufgrund des großen zeitlichen Abstandes als auch aufgrund der Tatsache, daß an ostdeutschen Hochschulen andere Orientierungen vorherrschen könnten, nicht auf die heutige Situation übertragbar. Wie stark sich ggf. die Akzeptanz unterscheiden kann, zeigen in nachfolgendem Abschnitt dargestellte Ergebnisse eigener Lehrendenbefragungen.

#### **6.4 Akzeptanz von Reformvorhaben und Arbeitszufriedenheit bei Lehrenden in Dresden, Frankfurt/ Oder und Zittau-Görlitz**

An dieser Stelle soll noch einmal in Erinnerung gerufen werden, daß die Zielstellungen der Lehrendenbefragungen an die Professoren- bzw. Hochschullehrerbefragungen von Kopp/ Weiß (1993) und Enders/ Teichler (1995) angelehnt waren. D.h., es soll eine Einschätzung der Rahmenbedingungen von Lehre und Studium als Bestandteil der Arbeitsbedingungen, (arbeits)klimatischer Aspekte sowie der Arbeitszufriedenheit erhoben werden - und zwar möglichst auch im Vergleich zur Einschätzung von Hochschulabsolventen und Studierenden. Außerdem sollten Informationen für angestrebte bzw. anzustrebende Fortbildungen in der Lehre gesammelt sowie die Einstellung der Lehrenden zu verschiedenen hochschulpolitischen Perspektiven und Reformvorhaben erfaßt werden. Während das Fortbildungsinteresse hier nicht unmittelbar von Bedeutung ist und daher hierzu auf die entsprechenden Arbeitberichte verwiesen werden kann (vgl. Krempkow 2001, 2002, 2003), ist die Einstellung der Lehrenden zu verschiedenen hochschulpolitischen Reformvorhaben eine Grundlage, um die Akzeptanz von Vorhaben und ggf. die Notwendigkeit akzeptanz-erhöhender Maßnahmen schon im Vorfeld abschätzen zu können.

Im einzelnen wurden in den Lehrendenbefragungen folgende Themenbereiche angesprochen:

- Situation von Lehre und Studium
- Zufriedenheit mit der beruflichen Situation
- Fortbildung in der Lehre
- Themen der Hochschulpolitik

Die Besonderheit der Lehrendenbefragungen in Dresden, Frankfurt/ Oder und Zittau-Görlitz ist auch darin zu sehen, daß in diesen erstmals vergleichbar zu bundesweit angelegten (westdeutschen) Befragungen Lehrende mehrerer Hochschulen mit ostdeutscher Berufsbiographie einbezogen wurden. Das trifft insbesondere für technische und naturwissenschaftliche Fachrichtungen zu, die nicht erst in den 1990er Jahren neu aufgebaut wurden, die aber von der Anzahl der Lehrenden her noch viele Hochschulen im Osten Deutschlands dominieren. Bei diesen Lehrenden kann aufgrund ihres anderen Erfahrungshintergrundes eine andere Perspektive auf ihre Arbeitssituation und auf hochschulpolitische Reformvorhaben vermutet werden. So hatte die Lehre nach Ansicht vieler Hochschullehrer vor 1990 im Osten Deutschlands an den Hochschulen allgemein einen größeren Stellenwert als heute.

#### **6.4.1 Anlage und Durchführung der eigenen Lehrendenbefragungen**

Die für die eigenen Lehrendenbefragungen ausgewählten Hochschulen sollten sich in Anlehnung an die Vorgehensweise bei den Experteninterviews nach Größe, Fächerspektrum sowie in den rechtlichen Rahmenbedingungen (Bundesland) möglichst stark unterscheiden (vgl. Diskussion zur Auswahl der Hochschulen für Experteninterviews im vierten Kapitel dieser Arbeit). Ein wesentliches Auswahlkriterium stellte aber natürlich auch die Tatsache dar, welche Hochschulen bereit waren, eine Lehrendenbefragung zumindest ideell zu unterstützen.

Während in Dresden eine gesättigte Zufallstichprobe mit 400 nach dem Zufallsprinzip ausgewählten Befragten (von etwa 1500 Lehrenden insgesamt) erfolgte, fanden an den sehr viel kleineren Hochschulen in Frankfurt/ Oder und Zittau-Görlitz die Befragungen nach dem Prinzip der Vollerhebung aller Lehrenden statt. Die repräsentative Auswahl der zu befragenden Lehrenden wurde an der TU Dresden durch das Dezernat Personal in Anlehnung an eine 1999/ 2000 durchgeführte Imageanalyse der TU Dresden (vgl. Lenz/ Wolter/ Winter 2000) durchgeführt. In Frankfurt/ Oder und Zittau-Görlitz erfolgte die Versendung der Fragebögen durch den/die zuständige/n Prorektor/in mit der Hauspost.

Da die Auswertung der Befragung natürlich anonym zu erfolgen hatte, wurde zur Sicherung der Anonymität lediglich die Zugehörigkeit zur Fakultät und zur Statusgruppe (Professor/in bzw. Hochschullehrer/in oder wissenschaftliches Personal) erfragt. Schon bei der Angabe zur Befristung der Tätigkeit wurden die Befragten explizit darauf hingewiesen, daß sie diese

Frage auch freilassen können, wenn deren Beantwortung aus ihrer Sicht die Anonymität beeinträchtigt. Zur Sicherung der Anonymität mußte auf die Erhebung (weiterer) soziodemographischer Daten wie Alter und Geschlechtszugehörigkeit verzichtet werden.

Neben den standardisierten Bewertungsfragen wurden in so gut wie allen Fällen die Fragestellungen als halboffene Fragestellungen mit zusätzlicher freier Antwortmöglichkeit konzipiert, um den Lehrenden die Möglichkeit zur Ergänzung und ggf. Korrektur der Itemlisten zu geben. Außerdem wurden die Lehrenden explizit aufgefordert, weitere Anregungen zu Qualitätsentwicklungsmaßnahmen bzw. ihre Meinung zu einem weiteren Thema zu formulieren und dafür ggf. auch ein zusätzliches Blatt zu verwenden.

#### **6.4.2 Betrachtung von Rücklauf und Repräsentativität der Befragungen**

Bevor ausgewählte Ergebnisse der Lehrendenbefragungen am Beispiel der TU Dresden im einzelnen diskutiert werden, soll die Überprüfung der Repräsentativität<sup>243</sup> der Befragungen vorgestellt werden. Da dies anhand von Strukturmerkmalen des Personals an der TU Dresden geschehen soll, kann hierbei gleichzeitig ein Überblick über die Struktur der Befragten gegeben werden.

Insgesamt wurden Mitte Juni 2001 an der TU Dresden 406 Fragebögen an vom Dezernat Personal zufällig ausgewählte Lehrende versendet. Bis Mitte Juli hatten 168 Lehrende geantwortet. Eine am 16. Juli versendete Erinnerungs- und gleichzeitig Dankespostkarte an alle Lehrenden brachte bis Mitte August einen Rücklauf von 206 auswertbaren Fragebögen. Da von 3 Lehrenden (die auf die Erinnerungspostkarte mit einem Anruf reagierten) bekannt ist, daß sie a) zwar als wissenschaftliche Mitarbeiter angestellt, aber nicht in der Lehre tätig sind, b) erst seit 1 Semester an der TU Dresden tätig sind und bisher keine Lehrveranstaltungen durchführten, oder c) inzwischen verstorben sind (Anruf von der Sekretärin), ergibt sich eine maximale Erreichbarkeit von 403 Lehrenden. Daraus ergibt sich eine Rücklaufquote von rund 51 Prozent. Auch bei den anderen beiden Hochschulen lagen die Rücklaufquoten bei etwa 50 Prozent.<sup>244</sup> Im Vergleich zu den bundesweiten Hochschul-lehrer- und Professorenbefragungen (28 Prozent bei Enders/ Teichler 1995 sowie 22 Prozent bei Kopp/ Weiß 1993) kann die Rücklaufquote als sehr befriedigend eingeschätzt werden.

---

<sup>243</sup> Zum Begriff vgl. Kromrey (1998: 259) bzw. Fußnote 118 in dieser Arbeit.

<sup>244</sup> Im einzelnen waren es 47 Befragte an der Europa-Universität Viadrina in Frankfurt/ Oder und 78 Befragte an der Hochschule Zittau-Görlitz.

Zur Überprüfung der Repräsentativität der Stichprobe und damit letztlich auch zur Sinnhaftigkeit von Vergleichen mit bundesweit angelegten Befragungen werden folgende Merkmale<sup>245</sup> herangezogen:

- Fakultäten/ Fachbereiche bzw. Fächergruppen
- Status bzw. Gruppenzugehörigkeit
- Befristung

Zur Überprüfung der Verteilung nach Fakultäten bzw. Fächergruppen wurde vom Dezernat Personal der TU Dresden, welches auch die Zufallsauswahl der zu befragenden Lehrenden an dieser Hochschule übernommen hatte, eine Liste mit der Personalstruktur zum Zeitpunkt der Auswahl der Lehrenden angefordert. Eine Gegenüberstellung der Verteilungen nach Fakultäten ergab, daß die Anteile bei der überwiegenden Anzahl der Fakultäten maximal 1, in wenigen Fällen 2 und nur in einem Fall (Mathematik/ Naturwissenschaften) 5 Prozentpunkte voneinander abwichen.

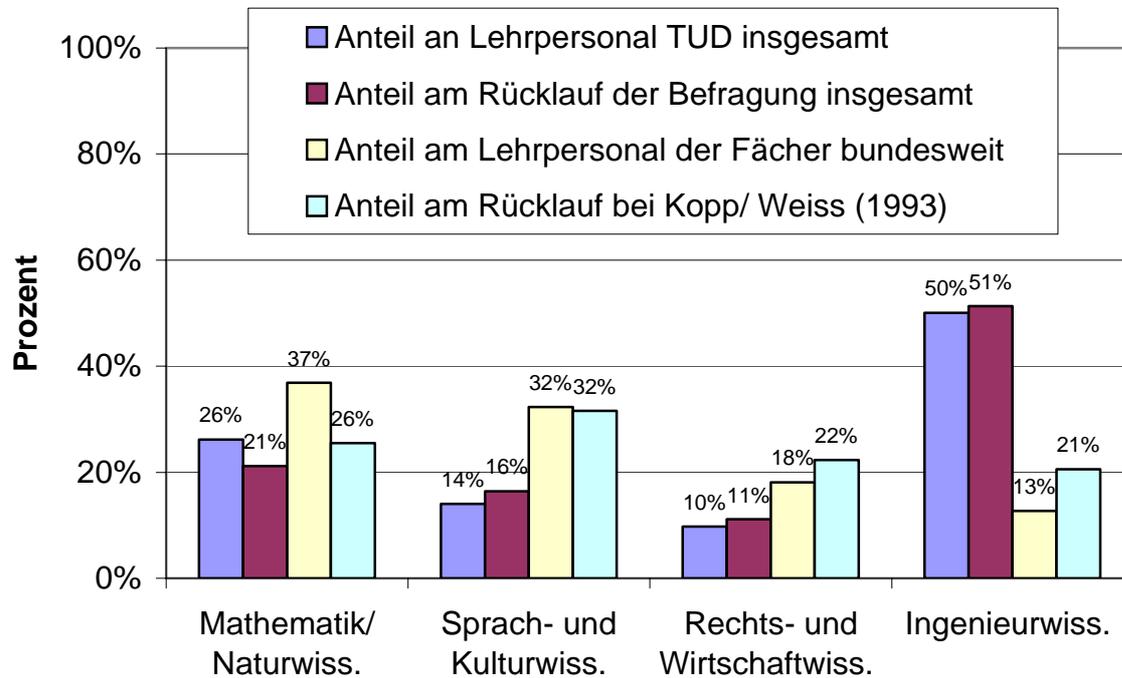
Will man die Verteilung nach Fakultäten/ Fachbereichen mit der in bundesweiten Studien vergleichen, so ist eine Zusammenfassung nach Fächergruppen notwendig. Bei Kopp/ Weiß (1993) wurde die Zusammenfassung des Statistischen Bundesamtes zugrundegelegt. Dieses Raster soll zu Vergleichszwecken übernommen werden.

---

<sup>245</sup> Wie bereits weiter vorn erwähnt, wurden Angaben zu Geschlecht, Alter oder andere soziodemographische Merkmale aus Rücksicht auf die Anonymität der Befragten nicht erfasst.

<sup>247</sup> Bei Betrachtung über alle drei Hochschulen sind die Hochschullehrer mit 44% im Vergleich zur Studie von Enders/ Teichler (1995) durch die Einbeziehung der Fachhochschule überrepräsentiert, da es an dieser kaum wissenschaftliche Mitarbeiter gibt. Insgesamt 4% machten keine Angaben.

## Rücklauf Lehrendenbefragung TU Dresden/ bundesweit nach Fächergruppen



Quellen: Krempkow 2001, Kopp/ Weiß 1993

Die Graphik zeigt, daß der Rücklauf bei der Lehrendenbefragung an der TU Dresden nach Fächergruppen dem Anteil der Lehrenden der TU Dresden in den jeweiligen Fächergruppen besser entspricht als bei der bundesweiten Befragung von Kopp/ Weiß (1993). Allerdings gibt es eine geringe Unterrepräsentierung der Naturwissenschaften. Außerdem weicht die Verteilung der befragten Lehrenden nach Fächergruppen deutlich vom Anteil der Fächer bundesweit ab. Deshalb sind Vergleiche mit bundesweiten Ergebnissen ohne Berücksichtigung der evtl. die Ergebnisse „verzerrenden“ Fächerkultur eher vorsichtig zu interpretieren. Betrachtet man die Anteile der Fächergruppen über alle drei befragten Hochschulen und vergleicht diese mit der bundesweiten Fächerstruktur, so sind die Ingenieurwissenschaften zwar immer noch etwas überrepräsentiert, die Anteile stimmen aber deutlich besser überein.

Die zweite Möglichkeit, die Repräsentativität der Stichprobe zu überprüfen, bestand im Vergleich der Anteile der Statusgruppen. Dies wurde wiederum anhand der vom Dezernat Personal zur Verfügung gestellten Strukturmerkmale vorgenommen, wobei 30 Prozent Hochschullehrer und 70 Prozent wissenschaftliche Mitarbeiter angegeben wurden.

Der Vergleich zeigte, daß die Befragten mit 34 Prozent Hochschullehreranteil die Struktur des Personals an der TU Dresden ebenfalls mit nur geringen Abweichungen widerspiegeln. Allenfalls kann von einer geringen Überrepräsentierung der Hochschullehrer im Vergleich zum unter wissenschaftliche Mitarbeiter zusammengefaßten Personal gesprochen werden. Ein kleiner Teil der Befragten (5 Prozent) machte von der auf Wunsch des Datenschutzbeauftragten ausdrücklich im Anschreiben freigestellten Möglichkeit Gebrauch, bei dieser Frage keine Angabe vorzunehmen. Da aber die Ergebnisse ohnehin meist getrennt nach Statusgruppen präsentiert werden, bleibt dies ohne Einfluß auf die Ergebnisse. Die Antworten derjenigen, die keine Angabe zum Status machten, flossen lediglich dort ein, wo Ergebnisse über alle Lehrenden der TU Dresden dargestellt werden.

Gemessen an der Struktur der Befragten bei Enders/ Teichler (1995)<sup>247</sup>, die ihre Studie als repräsentativ für die Bundesrepublik Deutschland (alte Bundesländer) einschätzten, entspricht die Verteilung der Befragten an der TU Dresden mit einem Drittel Professoren und etwa zwei Dritteln wissenschaftlichem Mittelbau gut den bundesweiten Vergleichswerten. Zwar ist zu beachten, daß bei der Lehrendenbefragung in Dresden nur haushaltsfinanzierte hauptberuflich beschäftigte Universitätsangehörige (ohne ruhendes Beschäftigungsverhältnis) befragt und damit alle diejenigen Beschäftigten nicht erfaßt wurden, die z.B. über Drittmittel angestellt wurden. Da diese Befragung aber als Lehrendenbefragung konzipiert wurde und über Drittmittel Beschäftigte überwiegend in der Forschung beschäftigt sind, entspricht dies der Zielstellung der Befragung. In anderen Lehrendenbefragungen wurde aus diesem Grunde ähnlich vorgegangen (vgl. Schaeper in Enders/ Teichler 1995).

Die dritte Möglichkeit zur Überprüfung der Repräsentativität waren die Angaben zur Befristung. Keine Angaben hierzu machten insgesamt 7 Prozent der Befragten. Auch dies war, wie bereits erwähnt, bei der Befragung freigestellt. Von allen Befragten insgesamt sind zwei Fünftel befristet und drei Fünftel unbefristet beschäftigt. Diese Anteile entsprechen in etwa den vom Dezernat Personal angegebenen 39,1 Prozent befristet Beschäftigten. Während jedoch die Hochschullehrer mit gut vier Fünfteln unbefristet beschäftigt sind, so ist bei den wissenschaftlichen Mitarbeitern mit zwei Dritteln der überwiegende Anteil

befristet.<sup>249</sup> Dies entspricht in etwa der Verteilung der Befragten in vorangegangenen bundesweiten Studien.

Da es nachfolgend nicht nur um rein deskriptive Auswertungen geht, sondern auch weiterführende Analysen vorgenommen werden sollen und die Fallzahl an der Universität Frankfurt/ Oder und der Hochschule Zittau-Görlitz hierfür nicht immer ausreicht, sollen die Ergebnisse entweder exemplarisch nur für die TU Dresden und/ oder für alle drei Hochschulen insgesamt vorgenommen werden. Eine umfassende Darstellung der Ergebnisse findet sich in den entsprechenden Arbeitsberichten (vgl. Krempkow 2002, 2003).

### **6.4.3 Akzeptanz hochschulpolitischer Reformvorhaben**

Wie weiter oben bereits dargestellt, war der Entstehungshintergrund dieser Lehrendenbefragung wie auch der Studie von Kopp und Weiß (1993), an die sich dieser Fragenkomplex anlehnt, eine Hochschullehrerbefragung zu den Folgen der in den 1980er Jahren einsetzenden Sparpolitik und damit in Zusammenhang stehender Managementmaßnahmen. Da sich die Hochschulen in Deutschland derzeit in einer ähnlichen Situation befinden, wurde ein Großteil der Fragen übernommen und lediglich die Fragestellungen zur Akzeptanz der Einbeziehung subjektiver Bewertungen in Systeme leistungsorientierter Mittelvergabe noch konkretisiert. Nachfolgend wird zunächst ein Überblick über die Bewertung hochschulpolitischer Reformvorhaben am Beispiel der TU Dresden gegeben, wobei Ergebnisse nach Statusgruppen differenziert werden.

Zunächst sollen jedoch die Ergebnisse anhand der arithmetischen Mittelwerte dargestellt werden, anschließend nach Häufigkeiten. Beide Darstellungsformen haben ihre Vor- und Nachteile. Eine Darstellung der Häufigkeiten mittels sogenannter „gestapelter Balkendiagramme“ ermöglicht zwar eine vollständige Wiedergabe aller Informationen zu den Antworten, aber keinen Vergleich zwischen den Statusgruppen in übersichtlicher Form. Eine zusammengefaßte Darstellung nur der zustimmenden Antwortanteile ermöglicht zwar übersichtliche Vergleiche, enthält aber keine Informationen mehr zu „Extremantworten“ an den Skalenenden. In die arithmetischen Mittelwerte fließen dagegen solche „Extremantworten“ durch die Berechnungsformel mit ein und sie ermöglichen einen übersichtlichen Vergleich der Statusgruppen. Voraussetzung ist das Vorliegen einer zentralen Tendenz der Antworten.

---

<sup>249</sup> Fasst man den Anteil der befristet Beschäftigten über alle drei Hochschulen zusammen, ändert sich der

## Bewertung hochschulpolitischer Reformvorhaben

Mittelwert von 1 (stimme zu) bis 4 (lehne ich ab)

1,0 1,5 2,0 2,5 3,0 3,5 4,0



Quelle: Krempkow 2001<sup>250</sup>

Die Unterschiede zwischen Hochschullehrern und wissenschaftlichen Mitarbeitern sind zum Teil beträchtlich. Die Hochschullehrer antworteten zu den meisten Fragen tendenziell zurückhaltender als die Mitarbeiter.

Anteil nur geringfügig auf 39%. 6% machten keine Angaben.

<sup>250</sup> Die Antworten wurden in dieser Graphik wie auch bei ähnlichen Graphiken in der Reihenfolge der Einschätzung durch alle Lehrenden insgesamt geordnet.

Signifikante Unterschiede finden sich besonders bei Fragen zur leistungsorientierten Mittelvergabe. Klar überwiegende Zustimmung erfährt bei Hochschullehrern wie Mitarbeitern nur die fachlich-inhaltliche, dicht gefolgt von der qualitativen Profilbildung. (Dies wird wie die Veröffentlichung von Leistungskennzahlen von den einzelnen Fächergruppen aber sehr unterschiedlich bewertet.)

Ebenfalls noch mehrheitlich, aber nicht mehr mit deutlicher Mehrheit zugestimmt wird der hochschulinternen Veröffentlichung von Leistungskennzahlen. Die hochschulinterne wird damit der sachsenweiten Veröffentlichung deutlich vorgezogen. Der Einbeziehung von Experten in bzw. leistungsorientierte Mittelvergabe durch Experten sowie der Einbeziehung von Absolventen in leistungsorientierte Mittelvergabe wird nur noch von den wissenschaftlichen Mitarbeitern mehrheitlich zugestimmt. Von den Hochschullehrern werden diese drei letztgenannten Statements bereits mehrheitlich abgelehnt. Dies würde der These von Rindermann (2001) zur Statusabhängigkeit der Akzeptanz von (studentischen) Bewertungen entsprechen, die hier sinngemäß auch für die Akzeptanz von Absolventenbewertungen gelten könnte. Möglicherweise gibt es aber auch andere, inhaltliche Gründe für die stärkere Akzeptanz von Absolventenbefragungen. Die Anmerkungen der befragten Lehrenden in den Fragebögen lieferten hierfür keine Hinweise. Allerdings finden sich solche Gründe in den Experteninterviews, die im Rahmen dieser Arbeit durchgeführt wurden (vgl. Abschnitt 4.4.2). So wird von den Experten häufig erwähnt, daß Absolventen im Gegensatz zu Studenten bereits einige Zeit Abstand zum Studium hätten und durch das Messen an der Praxis eine andere Sicht bekommen könnten.

Ungefähr ausgewogen sind die Einschätzungen der wissenschaftlichen Mitarbeiter zur sachsenweiten Veröffentlichung von Leistungskennzahlen, zu privaten Stiftungslehrstühlen sowie (schon mit leicht ablehnender Tendenz) zu mehr marktwirtschaftlichen Elementen im Hochschulsystem und zur Einbeziehung studentischer Bewertungen in leistungsorientierte Mittelvergabe. Die Einschätzung der Hochschullehrer ist nur zu den privaten Stiftungslehrstühlen etwa ausgewogen, die anderen Statements lehnen sie wiederum mehrheitlich ab.

Ein deutlich umgekehrtes Antwortmuster, daß die Hochschullehrer stärker zustimmen als die wissenschaftlichen Mitarbeiter, findet sich lediglich für Bildungsgutscheine zur teilweisen Finanzierung des Hochschuletats durch Studenten, was von den Hochschullehrern fast noch ausgewogen bewertet, von den Mitarbeitern jedoch deutlich abgelehnt wird. Die Neuorganisation des Studiums nach angelsächsischem Vorbild und mehr private Hochschulen werden von beiden Statusgruppen gleichermaßen überwiegend abgelehnt.

Ebenfalls Ablehnung erfahren die Formulierungen „leistungsorientierte Mittelzuweisung durch Absolventen“ sowie Studenten, welche von den Hochschullehrern wiederum deutlich stärker abgelehnt werden als von den Mitarbeitern.

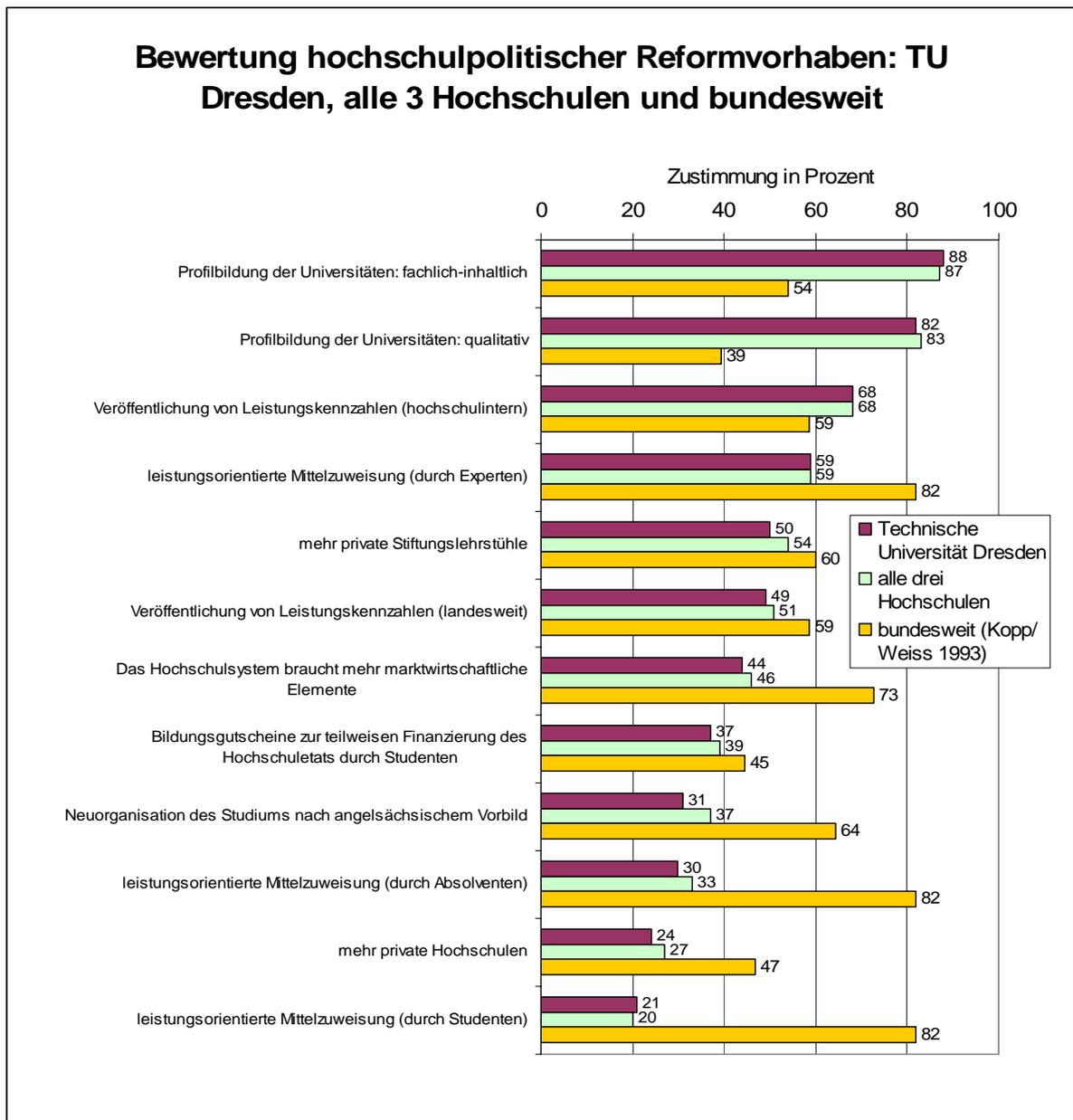
Bei den Einschätzungen zur leistungsorientierten Mittelzuweisung ist allerdings zu beachten, daß hier bewußt unterschiedliche Formulierungen in den Antwortvorgaben gewählt wurden, um ggf. den Einfluß der Frageformulierung deutlich machen zu können. Während Kopp/ Weiß allgemein nach der leistungsorientierten Mittelzuweisung fragten, wurde dies in den eigenen Lehrendenbefragungen spezifiziert nach verschiedenen Gruppen, durch die eine Bewertung stattfinden könnte. Hierbei gab es einerseits eine sehr stark verkürzte Formulierung, wobei in der Antwortvorgabe von Kopp/ Weiß zu leistungsorientierter Mittelzuweisung nur ergänzt wurde: „durch Experten“, „durch Studenten“ und „durch Absolventen“. Die Grundüberlegung hierbei war, unterschiedliche Frageformulierungen zu testen und die Antworten bei geringer Frageumformulierung noch bedingt vergleichbar zu belassen. Andererseits gab es komplett neu formulierte Antwortvorgaben, bei denen ausführlicher von „Einbeziehung der Bewertung von Experten/ Studenten/ Absolventen in die leistungsorientierte Mittelvergabe“ die Rede war. Wie sich zeigt, hat die Frageformulierung hier enormen Einfluß auf die Bewertung der hochschulpolitischen Reformvorhaben. So findet die ausführlichere Formulierung zur Einbeziehung studentischer Bewertungen, der Bewertungen von Absolventen und Experten deutlich mehr Zustimmung als die verkürzte Variante.

Wird nachfolgend die Darstellung in Prozentwerten betrachtet, die zum Vergleich mit der bundesweiten Studie vorgenommen wurde, so zeigt sich der Einfluß der Frageformulierung noch deutlicher: Gemessen in Prozent, findet die Einbeziehung studentischer Bewertungen zwar ebenfalls keine Mehrheit, aber mit über 40% wesentlich mehr Zustimmung als die verkürzte Formulierung mit 21%. Die Einbeziehung der Bewertungen von Absolventen mit rund 55% und Experten mit rund 70% der Befragten erfährt in der ausführlicheren Form bereits mehrheitlich Zustimmung, während sie bei der verkürzten Formulierung bezüglich der Absolventen mit 30% und bezüglich der Experten mit 59% ebenfalls deutlich geringer ausfiel. Insgesamt kann zur Formulierung der Antwortvorgaben festgehalten werden, daß diese großen Einfluß auf die absolute Zustimmung bzw. Akzeptanz haben kann. Außerdem kann je nach Darstellungsform (Mittelwerte oder Prozentwerte) ein etwas anderes Ausmaß der Zustimmung suggeriert werden, da sich bei den Mittelwerten eine häufige Nutzung der Skalenendpunkte stärker niederschlägt. Ob eine mehrheitliche Zustimmung gesehen werden

kann oder nicht, hängt wesentlich von der Frageformulierung, aber auch von der Darstellungsform ab. Über die verwendeten unterschiedlichen Frageformulierungen und Darstellungsformen stabil ist allerdings die relative Abfolge der Zustimmung zu den einzelnen Gruppen von Bewertenden: Die Gruppe der Experten, zu denen sich die Professoren selbst zählen können, erfährt die stärkste Akzeptanz, während die Gruppe der Absolventen etwas weniger und die der Studenten die geringste Zustimmung erhält.

Nachfolgend werden zur Einordnung der Bewertung hochschulpolitischer Statements die Ergebnisse der Befragung der Hochschullehrer an der TU Dresden den Ergebnissen der bundesweiten Studie von Kopp und Weiß (1993) gegenübergestellt, wofür anstelle der Mittelwerte in diesem Fall generell die Prozentangaben zur Zustimmung zu nutzen waren. Hierbei zeigt sich, daß die Hochschullehrer an der TU Dresden 2001 merklich andere Bewertungen vornahmen als die Hochschullehrer bundesweit 1993. Zu den meisten Statements (außer Profilbildung und hochschulinterne Veröffentlichung von Leistungskennzahlen) fallen die Bewertungen weit weniger zustimmend aus. Aufgrund des großen zeitlichen Abstandes wäre es jedoch eine unzulässige Interpretation, hieraus eine geringere Reformbereitschaft der Lehrenden an der TU Dresden im Vergleich zu anderen Hochschulen abzuleiten. Vielmehr dürfte die größere Skepsis der Lehrenden im Vergleich zur „Aufbruchstimmung“ der frühen 1990er Jahre auch mit der politischen „Schlechtwetterlage“ Sachsens und den überproportionalen Kürzungen im Bildungsbereich zusammenhängen, die im Zuge einer Haushaltssperre im Zeitraum der Befragung (Juli 2001) bekanntgegeben wurde.

## Bewertung hochschulpolitischer Reformvorhaben: TU Dresden, alle 3 Hochschulen und bundesweit

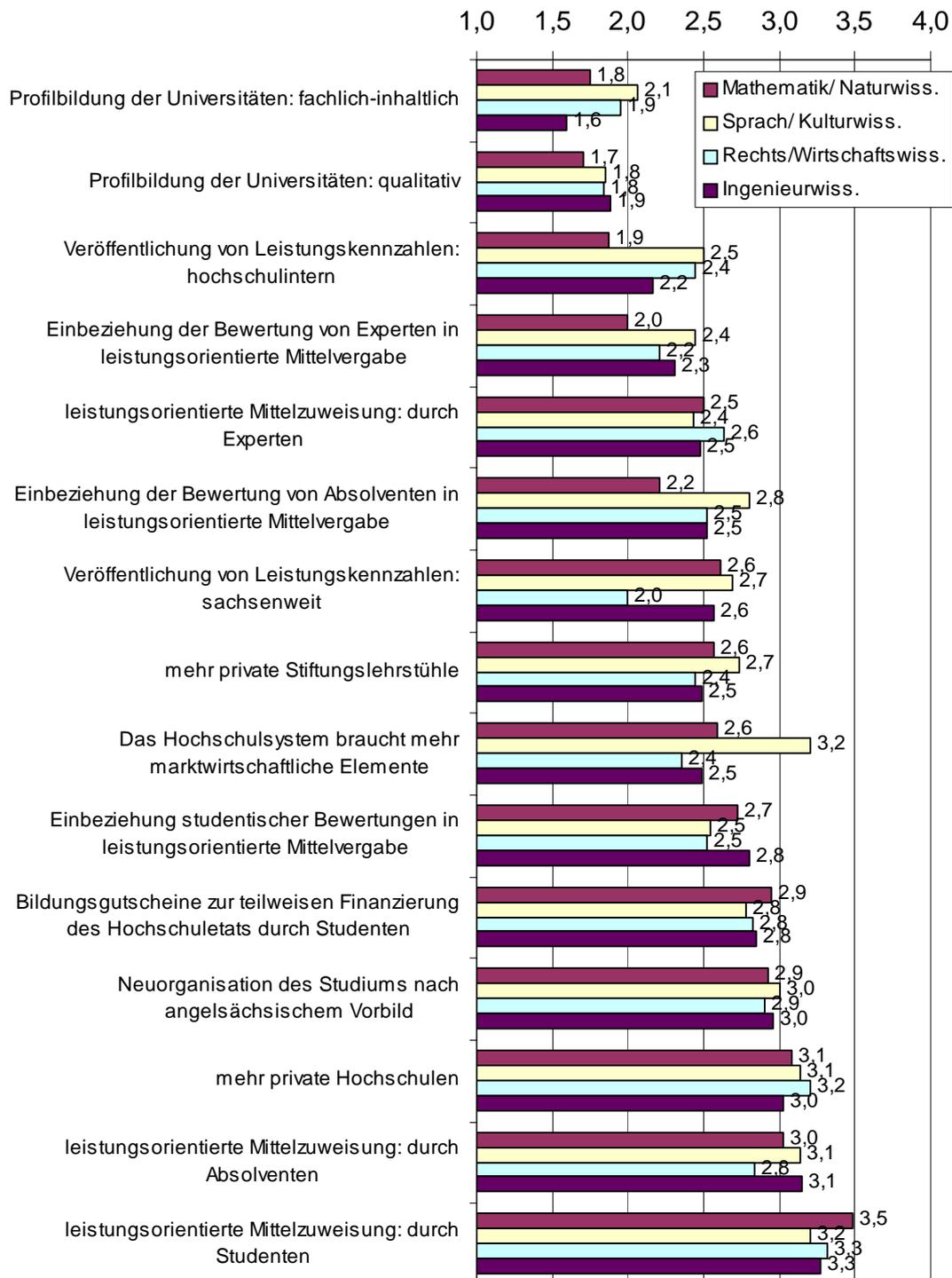


Quellen: Krempkow 2001 und 2003, Kopp/ Weiß 1993

In nachfolgender Graphik wurde für die TU Dresden nach Fächergruppen differenziert.

## Bewertung hochschulpolitischer Reformvorhaben durch Lehrende an der TU Dresden in den Fächergruppen

Mittelwert von 1 (stimme zu) bis 4 (lehne ich ab)



Quelle: Krempkow: 2001

Die Graphik zeigt, daß die Sprach- und Kultur- sowie die Rechts- und Wirtschaftswissenschaften, welche in der Studie von Kopp und Weiß einen deutlich größeren Anteil der Lehrenden stellen, die Statements zur fachlich- inhaltlichen Profilbildung und zur hochschulinternen Veröffentlichung von Leistungskennzahlen deutlich ablehnender bewerten als die Natur- und die Ingenieurwissenschaften. Die Rechts- und Wirtschaftswissenschaften bewerten jedoch eine sachsenweite Veröffentlichung von Leistungskennzahlen zustimmender als die anderen Fächergruppen. Etwas positiver werden auch mehr marktwirtschaftliche Elemente im Hochschulsystem eingeschätzt, welche wiederum von den Sprach- und Kulturwissenschaften<sup>251</sup> sehr viel stärker abgelehnt werden als von den anderen Fächern. Rechts-/ Wirtschafts-/ Sprach- und Kulturwissenschaften stimmen jedoch der Einbeziehung studentischer Bewertungen in leistungsorientierte Mittelvergabe etwas mehr zu als andere Fächergruppen.

#### **6.4.4 Berufliche Zufriedenheit der Lehrenden am Beispiel der TU Dresden**

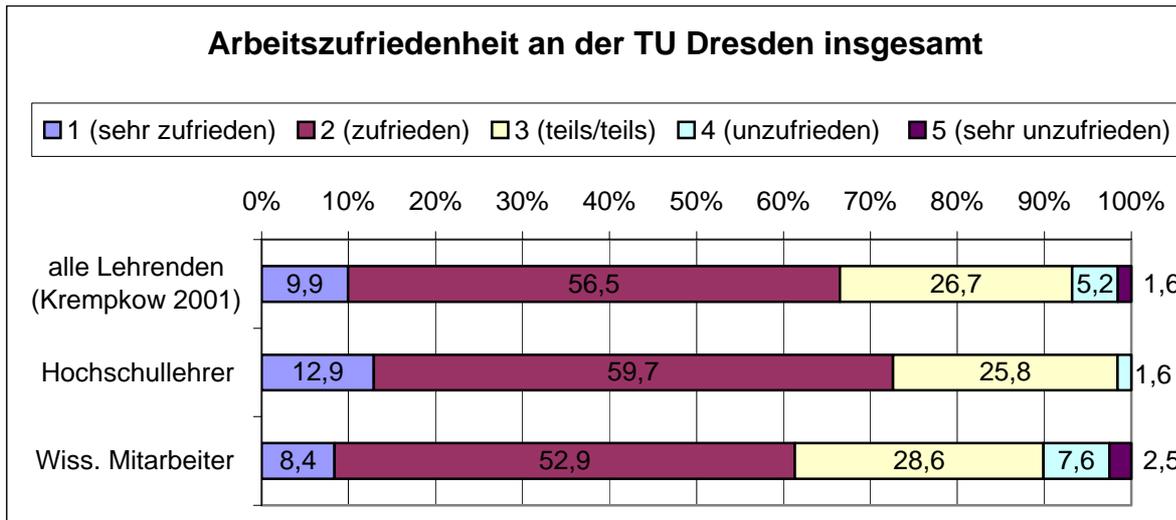
Von großem Interesse für die Funktion von Lehrendenbefragungen zur Erfassung der „Mitarbeiterperspektive“, aber auch wegen ihres langfristigen Zusammenhanges mit der Akzeptanz von Reformvorhaben ist die berufliche Zufriedenheit der Lehrenden, die häufig auch kurz als Arbeitszufriedenheit bezeichnet wird. Diese erfolgte global bzw. als allgemeine Gesamtschätzung und/ oder als Einschätzung von konkreten Einzelaspekten. Zunächst soll die allgemeine Gesamtschätzung vorgestellt werden.

Betrachtet man die Arbeitszufriedenheit an der TU Dresden über alle Lehrenden, so liegt der Mittelwert mit 2,3 deutlich im positiven Bereich. Bei einer Betrachtung nach Statusgruppen ist allerdings festzustellen, daß die wissenschaftlichen Mitarbeiter bei einem Mittelwert von 2,5 deutlich unzufriedener sind als die Hochschullehrer bei einem Mittelwert von 2,1.

Welche Angaben die Lehrenden an der TU Dresden zur Arbeitszufriedenheit im einzelnen machten, zeigt nachfolgende Graphik:

---

<sup>251</sup> Die Lehrenden der Philosophischen Fakultät der TU Dresden wurden hier zur Fächergruppe der Sprach- und Kulturwissenschaften gerechnet, da die in dieser Fakultät vertretenen Fächer nach der Fächersystematik der Statistikämter überwiegend den Sprach- und Kulturwissenschaften zugerechnet werden.



Quelle: Krempkow: 2001

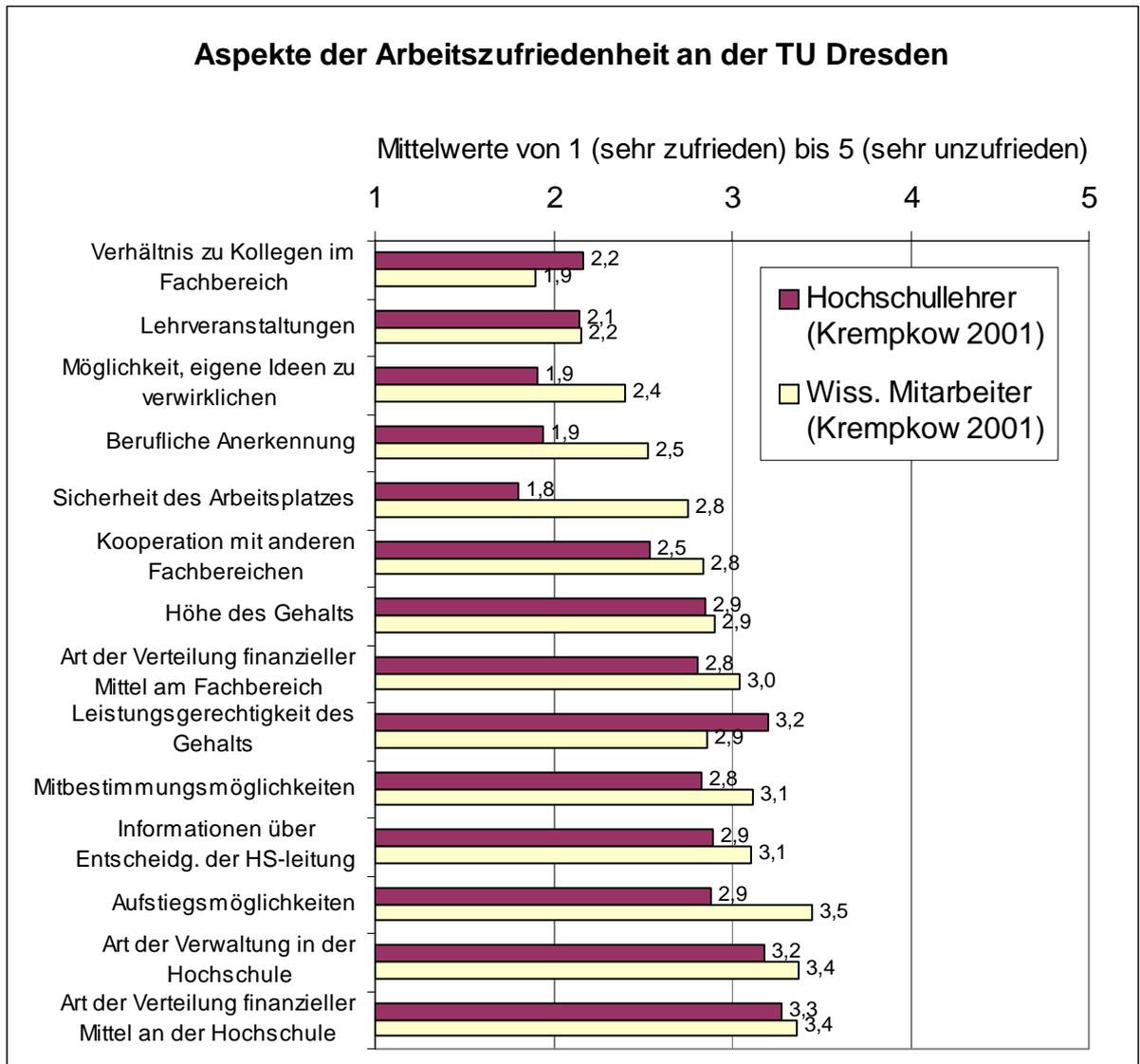
Die Mittelwerte für beide Statusgruppen fallen trotz ihres Unterschiedes beide deutlich positiver aus als bei der bundesweiten Hochschullehrerbefragung von Enders und Teichler,<sup>252</sup> wo der Mittelwert für den akademischen Mittelbau mit 3,1 und für die Professoren mit 2,4 angegeben wurde. Ähnlich positive Mittelwerte wie von den Professoren an der TU Dresden wurden bei Enders und Teichler sonst nur in Hongkong (internationaler Spitzenwert) angegeben, die Mitarbeiter sind im Mittel ähnlich zufrieden wie in Japan (vorderstes Drittel im internationalen Vergleich).

Daß dies keineswegs in allen ostdeutschen Hochschulen so sein muß, zeigt eine zusätzliche Vergleichsmöglichkeit zur Arbeitszufriedenheit innerhalb der Bundesrepublik Deutschland, wozu die Ergebnisse der Befragung von 231 Hochschullehrern an der Universität Halle (vgl. Winter 1999) herangezogen werden können. Hier fielen die Rücklaufquoten an den einzelnen Fachbereichen zwar sehr unterschiedlich aus (70 Prozent bei Jura und Sozialwissenschaften bis unter 40 Prozent bei Pharmazie, Sprach-/ Literatur- sowie Erziehungswissenschaften), und es ist aus der Veröffentlichung leider nicht ersichtlich, welchen Einfluß dies auf die Ergebnisse hatte und ob sie hierdurch ggf. „verzerrt“ sind. Bei dieser Befragung wurde jedoch ein hochschulweiter Mittelwert zur Arbeitszufriedenheit von 2,6 berechnet (auf einer Skala von 1 = „sehr gut“ bis 5 = „mangelhaft“). Da der Unterschied jedoch deutlich ausfällt und es bei größeren Unterschieden unwahrscheinlicher ist, daß diese

<sup>252</sup> Bei Kontrolle eines evtl. Fächereinflusses wurden nur geringe und nicht signifikante Differenzen zwischen den einzelnen Fächergruppen festgestellt, so daß hier die Vergleichbarkeit gegeben ist.

allein auf verzerrende Biaseffekte zurückzuführen sind, wird hier davon ausgegangen, daß es tatsächlich einen Unterschied gibt.

Die globale Einschätzung der Arbeitszufriedenheit kann jedoch stärker von Erwartungshaltungen beeinflußt sein als die Bewertung einzelner konkreter Aspekte. Außerdem bietet sie weniger Ansatzpunkte für Verbesserungen. Deshalb werden nachfolgend zusätzlich die detailliert abgefragten Aspekte der Arbeitszufriedenheit dargestellt.



Quelle: Krempkow: 2001

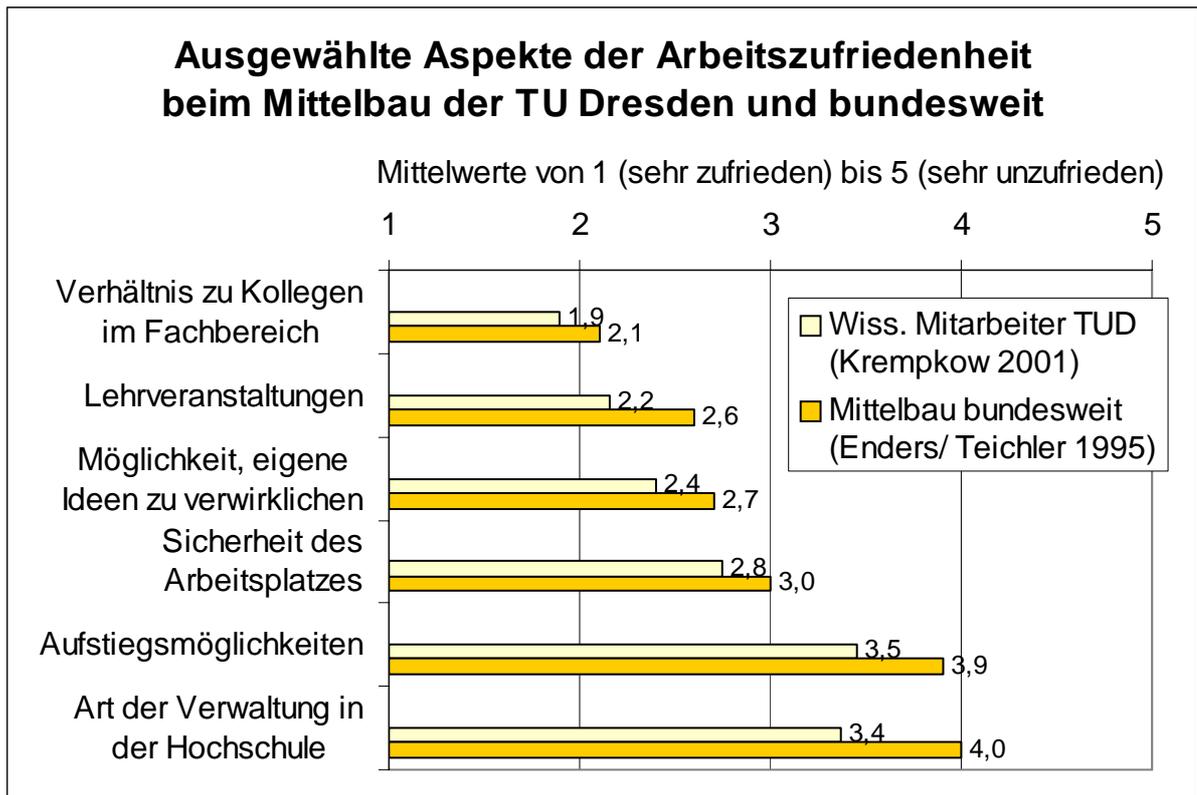
Betrachtet man die abgefragten Aspekte der Arbeitszufriedenheit im einzelnen, so fallen zunächst die enormen Unterschiede zwischen Hochschullehrern und wissenschaftlichen Mitarbeitern auf, die sich insbesondere bei der Sicherheit des Arbeitsplatzes, den

Aufstiegsmöglichkeiten, der Möglichkeit, eigene Ideen zu verwirklichen, sowie bei der beruflichen Anerkennung zeigen. Hier sind die Hochschullehrer sehr viel häufiger zufrieden als die wissenschaftlichen Mitarbeiter. Die Tendenz, daß die Hochschullehrer zufriedener sind, zeigt sich – wenngleich in nicht so großem Ausmaß – in allen Aspekten mit Ausnahme des Verhältnisses zu Kollegen am Fachbereich und der Leistungsgerechtigkeit des Gehalts.

Von allen erfragten Aspekten am zufriedensten waren die Hochschullehrer mit der Sicherheit des Arbeitsplatzes, der Möglichkeit, eigene Ideen zu verwirklichen und der beruflichen Anerkennung. Die Mitarbeiter waren mit dem Verhältnis zu Kollegen am Fachbereich und den Lehrveranstaltungen am zufriedensten. Am unzufriedensten waren die Mitarbeiter mit den Aufstiegsmöglichkeiten, der Art der Verwaltung und der Verteilung finanzieller Mittel in der Hochschule. Mit den Mitbestimmungsmöglichkeiten und Informationen über Entscheidungen der Hochschulleitung waren sie jedoch fast ebenso unzufrieden. Die Hochschullehrer sind ebenfalls mit der Art der Verteilung finanzieller Mittel und der Verwaltung am unzufriedensten, wobei letzteres gleichauf liegt mit der Unzufriedenheit bei der Leistungsgerechtigkeit des Gehalts.

Werden die Ergebnisse zu Aspekten der Arbeitszufriedenheit nicht nur absolut betrachtet, sondern auch mit denjenigen im Vergleich, die aus der bundesweiten Untersuchung von Enders und Teichler (1995: 24) zur Verfügung stehen, so zeigt sich, daß die wissenschaftlichen Mitarbeiter an der TU Dresden in allen dort einzeln aufgeführten Aspekten der Arbeitszufriedenheit zufriedener sind als der Mittelbau bundesweit.

Da Einschätzungen konkreter Aspekte erfahrungsgemäß weniger von allgemeinen Erwartungshaltungen beeinflußt werden als allgemeine Einschätzungen, kann davon ausgegangen werden, daß die (relativ) höhere Arbeitszufriedenheit der wissenschaftlichen Mitarbeiter an der TU Dresden im Vergleich zum Bundesschnitt nicht allein eine Folge niedrigerer Erwartungen ist. Einschränkend muß aber hinzugefügt werden, daß für den Vergleich nicht alle Aspekte zur Verfügung standen und daß die Unzufriedenheit mit mehreren Aspekten (Aufstiegchancen, Art der Verwaltung in der Hochschule) immer noch sehr hoch ist – auch im Vergleich zu den Hochschullehrern.



Quelle: Krempkow: 2001

Vor allem sind die wissenschaftlichen Mitarbeiter der TU Dresden mit der Art der Verwaltung in der Hochschule, den Lehrveranstaltungen, aber auch mit den Aufstiegsmöglichkeiten deutlich zufriedener als bundesweit. Nur wenig zufriedener bzw. in etwa genauso zufrieden oder unzufrieden wie im Bundesvergleich sind sie mit der Sicherheit des Arbeitsplatzes und dem Verhältnis zu Kollegen im Fachbereich. Außerdem fällt auf, daß trotz der Unterschiede bei der Bewertung die Rangfolge weitgehend dieselbe ist. Dies läßt vermuten, daß die Bewertungsmuster an der TU Dresden auch im Vergleich zu bundesweiten Studien weitgehend stabil sind und ein zufälliges Zustandekommen nicht nur statistisch – was bereits durch Signifikanztests überprüft wurde und ausgeschlossen werden konnte – sondern auch faktisch gering ist.

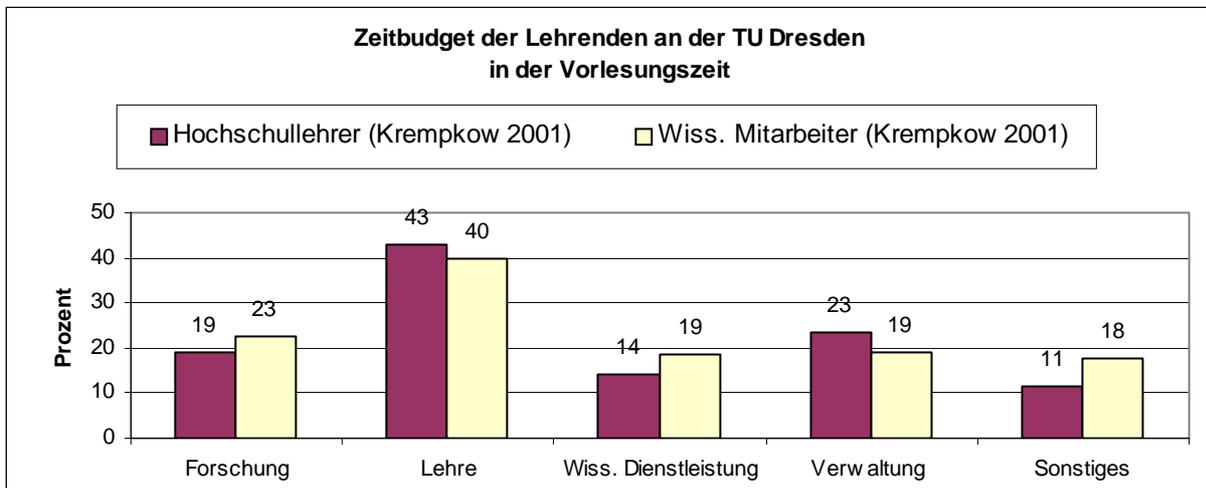
#### 6.4.5 Das Zeitbudget und sein Zusammenhang mit der Arbeitszufriedenheit

Das Zeitbudget ist neben einer Betrachtung seines Zusammenhanges mit der Arbeitszufriedenheit einer der interessantesten Untersuchungsbereiche, da es zeigen könnte, inwieweit hochschulpolitische Bemühungen zur Erhöhung des Stellenwertes der Lehre über die Jahre hinweg praktisch wirksam waren. Zudem könnte die zeitliche Entwicklung von Differenzen

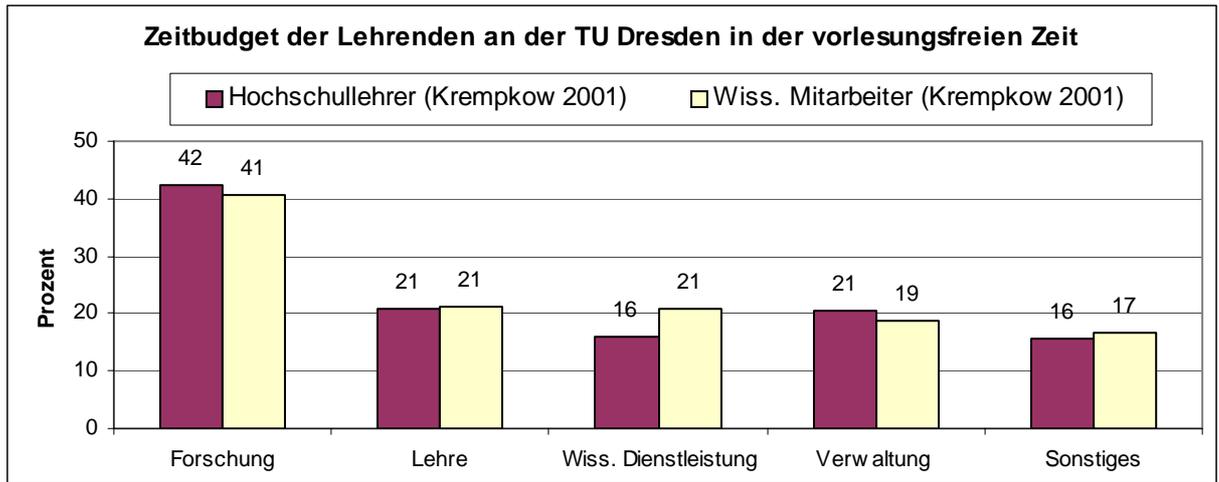
zwischen tatsächlicher und als optimal erachteter Arbeitszeit auch die Arbeitszufriedenheit mit beeinflussen. Dies soll nach der Vorstellung und Einordnung der Ergebnisse zum Zeitbudget in Ergebnisse bundesweiter Studien untersucht werden.

Die Befragung an der TU Dresden zum prozentualen Anteil der Arbeitszeit, der für die 5 genannten Aufgabenbereiche (Forschung, Lehre, wiss. Dienstleistung, Verwaltung, Sonstiges) investiert wird, ergab erwartungsgemäß in der Vorlesungszeit ein klares Übergewicht für die Lehre, in der vorlesungsfreien Zeit für die Forschung. Ähnliche Ergebnisse wurden zuvor in mehreren bundesweiten Hochschullehrerbefragungen (Kopp/Weiß 1993, Enders/ Teichler 1995), aber auch in regional durchgeführten Lehrendenbefragungen (vgl. Schaeper 1995, Winter 1999) publiziert.

Auffällig waren jedoch die relativ hohen Zeitanteile von jeweils etwa einem Fünftel der Arbeitszeit, die für wiss. Dienstleistungen und Verwaltungsaufgaben aufgewandt wurden. Die Angaben der Hochschullehrer und der wissenschaftlichen Mitarbeiter unterschieden sich voneinander vor allem darin, daß wissenschaftliche Mitarbeiter etwas mehr Zeit für wiss. Dienstleistungen und Sonstiges sowie weniger Zeit für Verwaltungsaufgaben aufwenden.



Quelle: Krempkow: 2001



Quelle: Krempkow: 2001

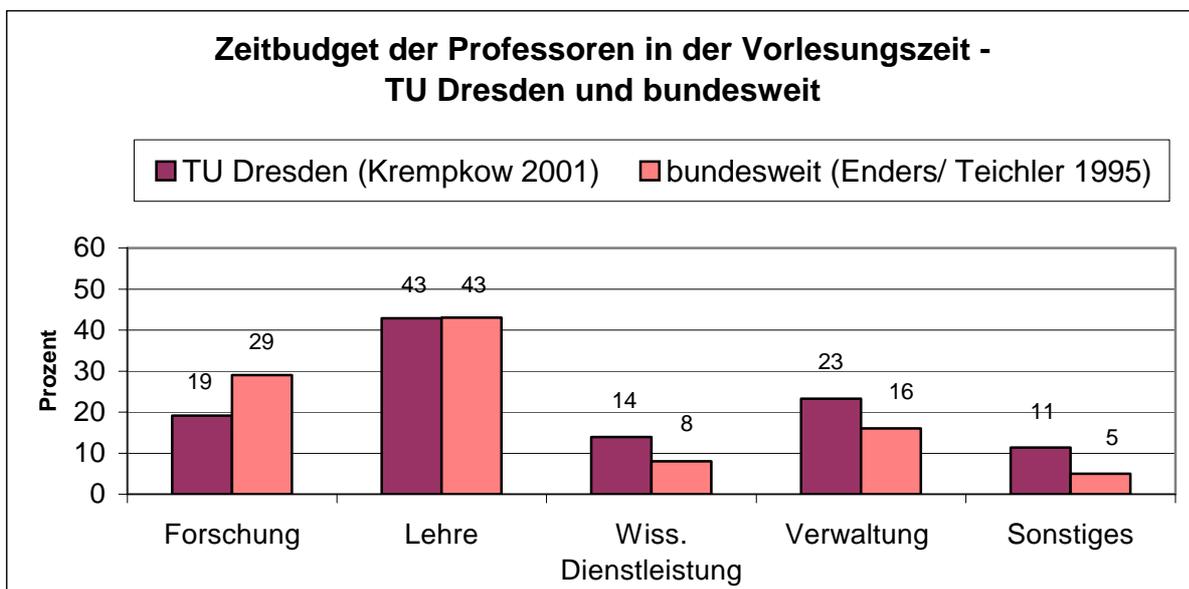
Zur Angabe „Sonstiges“ wurden die Lehrenden gebeten, in einer offenen Frage hierzu zu nennen, was genau sie dort einordneten. Hiervon machten 19 Prozent der Lehrenden Gebrauch.

Im einzelnen wurden unter Sonstiges folgende Aufgabenbereiche genannt:

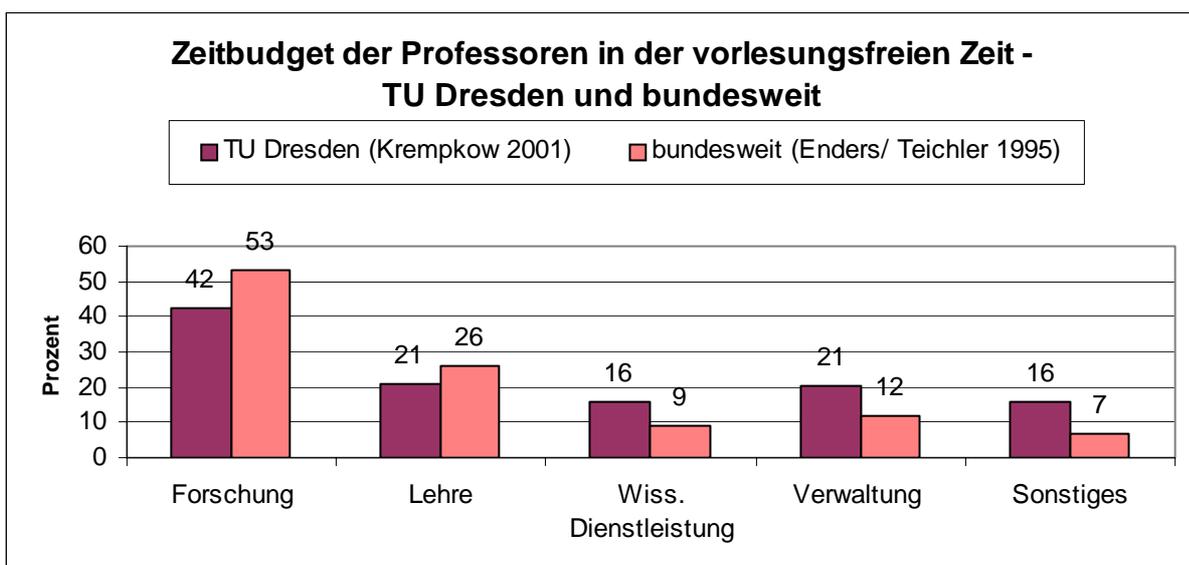
(innerhalb der Statusgruppen alphabetisch geordnet)

Hochschullehrer	Wissenschaftliche Mitarbeiter
Diplomatie	Anträge schreiben
Direktor im Nebenamt der MPA	Beratung
Drittmittel-Akquisition	Beratung, Gremien
Fachgesellschaft an Institut	Betreuung
Gremien, Organisationen	Drittmittelerfordernis
Gremienarbeit	Einkauf und Organisation
Gutachten	externe Dienstleistungen (Kommission, Gremien)
Konferenzorganisation	Gleichstellung
neue Vorlesung	Konzeption/Lehrmaterial entwickeln
Territorium	Lektüre
Vorträge, Publikationen	Netzadministration
Vorträge, Tagungen	Pendeln Campus, Informatik
Weiterbildung	Projektarbeit
	Projektbeantragung, Akquisition
	Projekte
	Rechneradministration und dazugehörige Dienstleistungen
	Seminare, Kolloquien
	Studienorganisation u. -beratung
	Technikbetreuung
	Telefonate, kopieren, tausend Kleinigkeiten
	Wegezeit zum Campus
	Weiterbildung (Fortbildung)

Die Zeitanteile der Hochschullehrer für die Aufgabenbereiche wiss. Dienstleistungen und Verwaltung und der Anteil für sonstige Aufgaben fielen deutlich höher aus als im Bundesschnitt, während der Zeitanteil für die Forschung etwa 10 Prozentpunkte geringer ausfiel als bundesweit. Dieses Phänomen zeigte sich sowohl in der Vorlesungszeit als auch in der vorlesungsfreien Zeit. Für die Lehre ist der Zeitaufwand in der Vorlesungszeit gleich hoch, in der vorlesungsfreien Zeit geringer als bundesweit. Dies bedeutet, daß der geringere Zeitaufwand für die Forschung an der TU Dresden nicht der Lehre zugute kommt, sondern für Verwaltungs-, Dienstleistungs- oder sonstige Aufgaben verwendet wird.



Quellen: Krempkow: 2001, Enders/ Teichler 1995

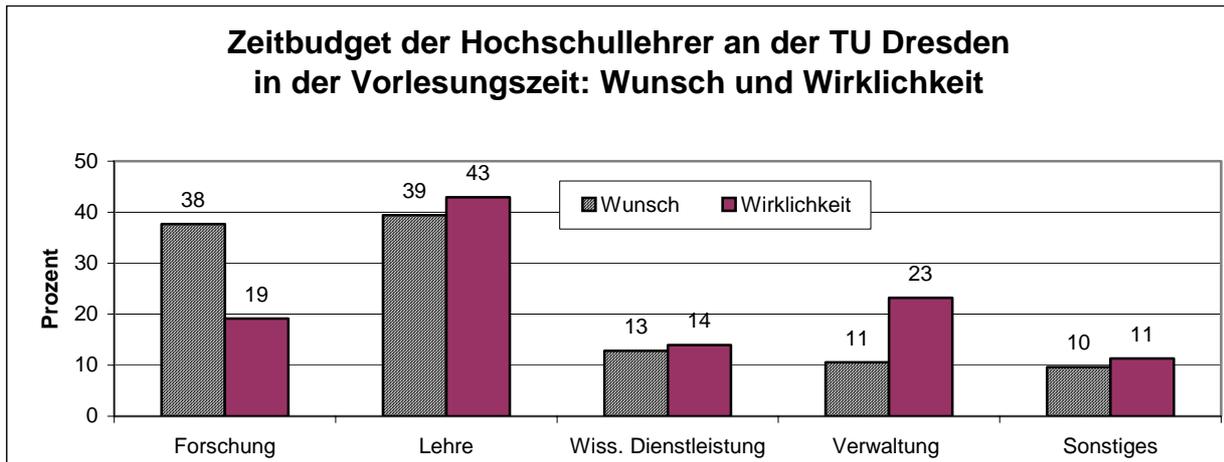


Quellen: Krempkow: 2001, Enders/ Teichler 1995

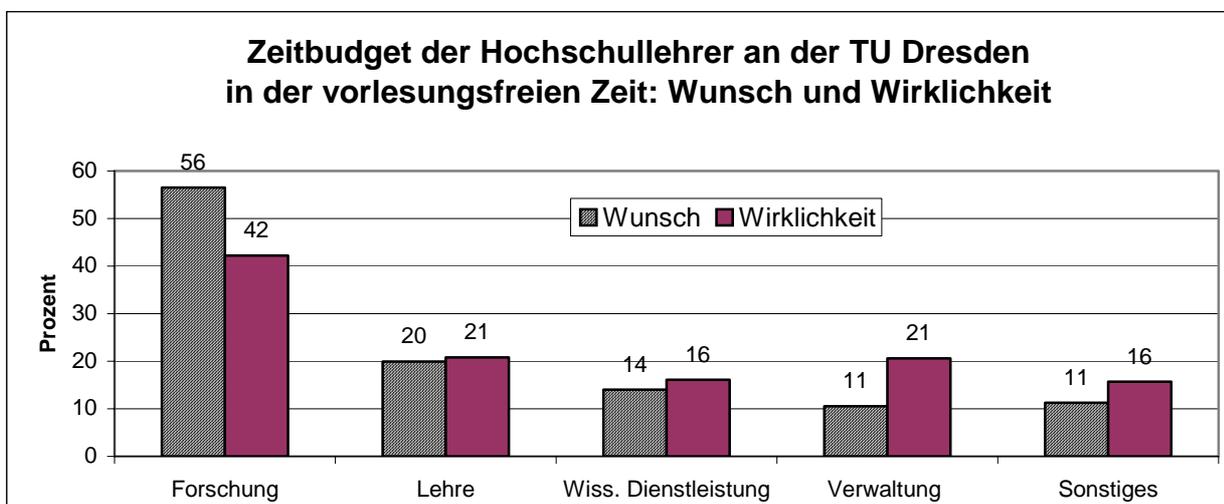
Der Vergleich mit anderen Hochschullehrer- und Professorenbefragungen zeigt, daß die Hochschullehrer an der TU Dresden vergleichsweise weniger Zeit für Forschung aufwenden, aber nicht mehr als im Bundesschnitt für die Lehre (vgl. Enders/ Teichler 1995: 20 sowie Schaeper 1995: 142). Hinzuzufügen ist, daß der Anteil der Arbeitszeit, der während der Vorlesungszeit für die Forschung verwandt wird, im Laufe der Jahre kontinuierlich größer geworden zu sein scheint (1976/77: 23 Prozent, 1983/84: 27 Prozent, 1990/91: 28 Prozent, vgl. Enders/ Teichler 1995: 20). Angesichts dieser Daten könnte es sein, daß die hochschulpolitischen Bemühungen zur Erhöhung des Stellenwertes der Lehre, die in den 1990er Jahren forciert wurden, den Trend zur Erhöhung des Zeitanteils für die Forschung zwar vielleicht zu bremsen vermochten. Der Lehre verhalfen sie jedoch möglicherweise faktisch gegenüber der Forschung nicht zu einem höheren Stellenwert. Der Anstieg des Zeitanteils für die Forschung ist jedoch nicht sehr groß und es können auch andere Faktoren hierbei eine Rolle gespielt haben, so z.B. zunehmender Konkurrenzdruck bei der Akquise von Forschungsgeldern.

Ein direkter Vergleich des Zeitbudgets der wissenschaftlichen Mitarbeiter an der TU Dresden mit Ergebnissen bundesweiter Befragungen von wissenschaftlichen Mitarbeitern kann an dieser Stelle leider nicht erfolgen, da hierzu in den entsprechenden Veröffentlichungen kaum Ergebnisse publiziert wurden. Die verbalen Umschreibungen der Ergebnisse (vgl. Enders/ Teichler 1995: 22) lassen aber darauf schließen, daß die Ergebnisse an der TU Dresden mit den Ergebnissen dieser Studien weitgehend übereinstimmen.

Die als optimal erachteten Zeitbudgets unterscheiden sich wie in o. g. Studien deutlich von den tatsächlichen Zeitbudgets. Dies sollen nachfolgende Gegenüberstellungen zunächst für die Hochschullehrer, dann für die wissenschaftlichen Mitarbeiter veranschaulichen.



Quelle: Krempkow: 2001



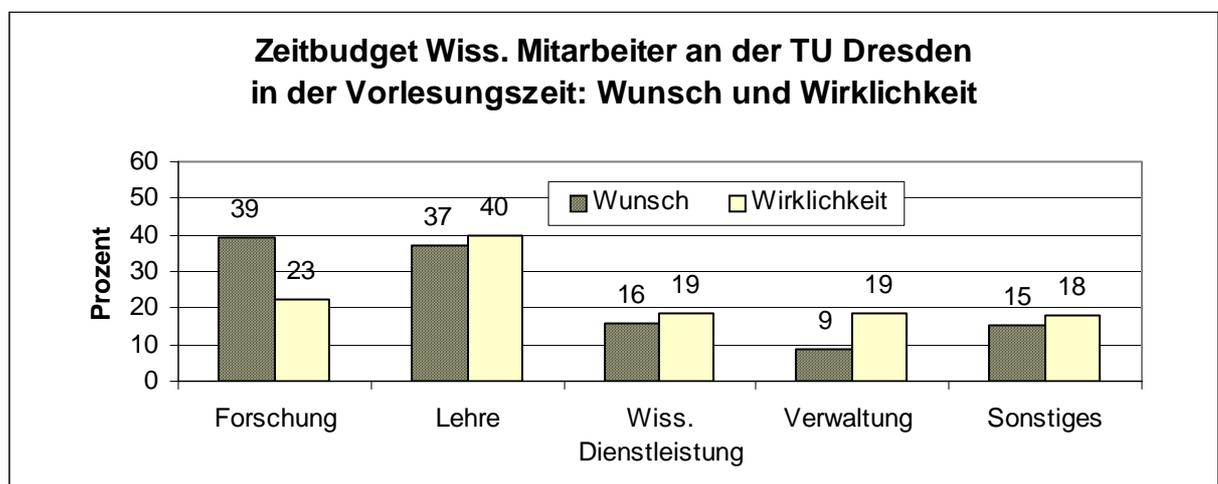
Quelle: Krempkow: 2001

Die Graphik zeigt, daß sich die Hochschullehrer an der TU Dresden insgesamt deutlich stärker der Forschung widmen möchten, als sie es derzeit tun (können). In der Vorlesungszeit würden sie mit durchschnittlich 38 Prozent gern doppelt soviel Zeit wie bisher und genausoviel Zeit wie für die Lehre (39 Prozent) aufwenden. In der vorlesungsfreien Zeit würden sie gern rund 60 Prozent ihrer Arbeitszeit für die Forschung investieren. Nur ein kleiner Teil der Hochschullehrer würde jedoch gern weniger Zeit in die Lehre einbringen. Deutlich weniger Zeit investieren wollen die meisten Hochschullehrer jedoch in Verwaltungsaufgaben (11 Prozent statt 23 Prozent bzw. 20 Prozent). Auch in wiss. Dienstleistungen und in sonstige Aufgaben würden die Hochschullehrer zugunsten der Forschung gern etwas weniger Zeit investieren.

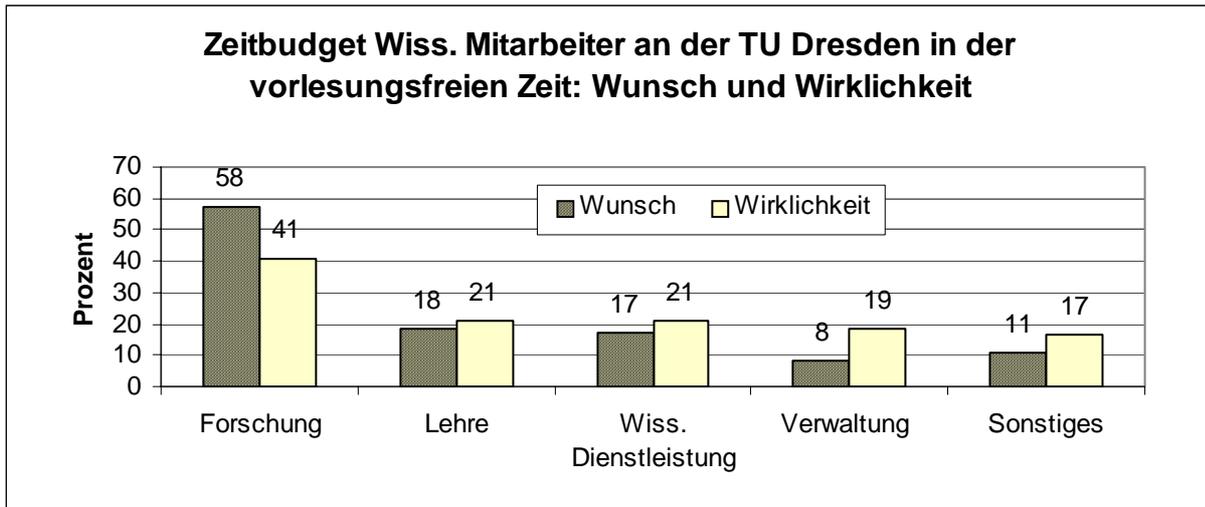
Dies entspricht in der Grundtendenz den Ergebnissen der bundesweiten Hochschullehrerbefragungen, wobei dort allerdings – wie weiter oben bereits beschrieben – schon jetzt mehr Zeit für die Forschung aufgebracht wird (vgl. Kopp/ Weiß 1993: 113).

In anderen Studien wurden die Lehrenden analog befragt, ob ihre Interessen eher in der Lehre oder eher in der Forschung liegen (vgl. Enders/ Teichler 1995: 16 sowie Schaeper 1995: 142). Da diese Fragestellung aber relativ allgemein ist und die Befragten nicht wie bei der Frage nach der konkreten prozentualen Aufteilung der Arbeitszeit aufgefordert werden, ihren Wunsch an der realen Situation zu messen, erscheint die konkretere Fragestellung im Zusammenhang mit der Arbeitszufriedenheit geeigneter, wenngleich hierdurch keine direkte Vergleichbarkeit der Ergebnisse gegeben ist. Aber auch ohne direkte Vergleichbarkeit kann festgestellt werden, daß Hochschullehrer tendenziell mehr Interesse an der Forschung als an der Lehre zeigen, obwohl die Hochschullehrer mehrheitlich angaben, ihre Interessen lägen in beidem (Forschung und Lehre). Die Grundtendenz einer etwas stärkeren Forschungs- als Lehrorientierung kann damit in allen genannten Untersuchungen aufgezeigt werden.

Ähnlich große Unterschiede zwischen Wunsch und Wirklichkeit ergab auch die Untersuchung der Befragungsergebnisse zum Zeitbudget für die wissenschaftlichen Mitarbeiter an der TU Dresden.



Quelle: Krempkow: 2001



Quelle: Krempkow: 2001

Wie die beiden Graphiken zeigen, wünschen sich auch die wissenschaftlichen Mitarbeiter in der Vorlesungszeit deutlich mehr Zeit für die Forschung (+17 Prozentpunkte) und deutlich weniger für Verwaltungsaufgaben (-10 Prozentpunkte). Außerdem würden sie gern jeweils etwas weniger Zeit (je 3 Prozentpunkte) für wiss. Dienstleistungen, die Lehre und Sonstiges aufwenden. In der vorlesungsfreien Zeit zeigt sich dieselbe Tendenz, wobei auch hier ein weitaus höherer Zeitanteil auf die Forschung entfällt.

Der Einfluß des Zeitbudgets auf die Arbeitssituation und -zufriedenheit wurde bei Enders/Teichler (1995) mittels einer Unterteilung der Lehrenden in eher Lehrorientierte und eher Forschungsorientierte untersucht. Diese Untersuchung soll ebenfalls vorgenommen werden. Hierzu wurden die Lehrenden anhand der auf die Vorlesungszeit bezogenen Antworten zum Zeitbudget in zwei Gruppen unterteilt: Die erste Gruppe der Lehrorientierten bilden mit 21 Prozent aller Befragten diejenigen, die gern mehr als die bisher von ihnen aufgewandte Arbeitszeit für die Lehre investieren würden. Die zweite Gruppe bilden mit 86 Prozent aller Befragten die Forschungsorientierten, die gern mehr als die bisher von ihnen aufgewandte Arbeitszeit für die Forschung nutzen würden. Diese Gruppeneinteilung ist nicht trennscharf, da ein Teil der Lehrenden (18 Prozent aller Befragten) gern mehr Zeit sowohl für die Lehre als auch für die Forschung aufwenden würde. Dies ist zwar der überwiegende Anteil der Lehrorientierten. Angesichts des in hochschulpolitischen Diskussionen immer wieder beschworenen Ideals der Einheit von Forschung und Lehre erscheint es jedoch als erstaunlich wenig. Der Anteil nicht klar einer Gruppe zuzuordnender Lehrender wurde daher in den jeweiligen Gruppen belassen. Nachfolgende Kreuztabelle zeigt jeweils die Anteile

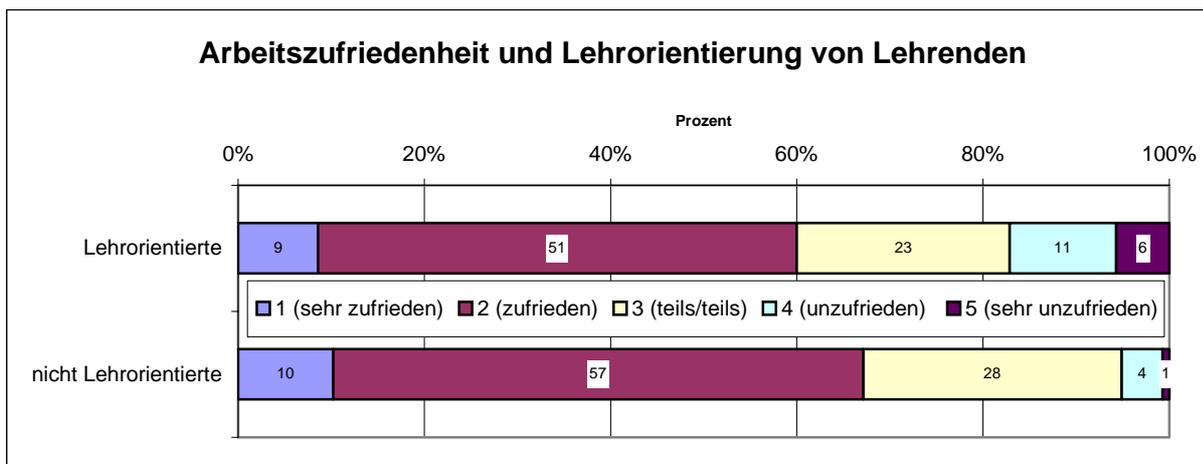
der Forschungsorientierten und Lehrorientierten sowie den Anteil der Forschungs- und Lehrorientierten an allen Befragten insgesamt. (Ein Teil der Befragten machte keine Angaben zu ihrem Status, daher sind hier insgesamt nur 173 Befragte ausgewiesen.) Die Anteile der Forschungsorientierten und der Lehrorientierten innerhalb der Hochschullehrer und der wissenschaftlichen Mitarbeiter unterscheiden sich kaum. Allerdings ist der Anteil der Nur-Lehrorientierten (also derjenigen, die gern mehr Lehre, aber nicht mehr Forschung machen würden), vollständig auf wissenschaftliche Mitarbeiter zurückzuführen.

**Kreuztabelle Forschungsorientierte und Lehrorientierte**

			würden gern mehr Zeit in Lehre investieren als sie bisher tun		Gesamt
			nein	ja	
würden gern mehr Zeit in Forschung investieren als sie bisher tun	nein	Anzahl	19	5	24
		% der Gesamtzahl	11,0%	2,9%	13,9%
	ja	Anzahl	118	31	149
		% der Gesamtzahl	68,2%	17,9%	86,1%
Gesamt		Anzahl	137	36	173
		% der Gesamtzahl	79,2%	20,8%	100,0%

Quelle: Krempkow: 2001

Für die beiden Gruppen der Forschungsorientierten und der Lehrorientierten soll untersucht werden, wie sie sich in ihrer Arbeitszufriedenheit insgesamt sowie in einzelnen Aspekten der Arbeitszufriedenheit unterscheiden (Lehrveranstaltungen, berufliche Anerkennung, Leistungsgerechtigkeit des Gehalts usw.). Es erscheint naheliegend, hierzu als Hypothese zu formulieren, daß aufgrund des relativ geringen Stellenwertes der Lehre die Lehrorientierten unzufriedener sind. Hierbei müsste aber auch ein evtl. vorhandener Einfluß der Zugehörigkeit zu Statusgruppen untersucht werden (vgl. Diskussion am Anfang dieses Kapitels). Wie dies und weitere Aspekte (z.B. Befristung, Rahmenbedingungen der Lehre usw.) die Arbeitszufriedenheit beeinflussen, kann hier nicht abschließend erörtert werden. Wie dies geschehen könnte, wird jedoch nachfolgend in einer umfassenderen Analyse von „Determinanten der Arbeitszufriedenheit“ erläutert (vgl. auch Schaeper 1995: 139ff.). An dieser Stelle soll zunächst nur die nach Forschungs- bzw. Lehrorientierten getrennt vorgenommene Untersuchung der Arbeitszufriedenheit dargestellt werden, um exemplarisch die Bedeutung weiterer Untersuchungen aufzuzeigen. Nachfolgende Graphik zeigt die genaue Antwortverteilung aller Lehrenden für Lehrorientierte und nicht Lehrorientierte:



Quelle: Krempkow 2001

Die Berechnung der arithmetischen Mittel für alle Lehrenden zeigt, daß Lehrorientierte mit einem Mittelwert von 2,5 signifikant unzufriedener sind als nicht Lehrorientierte mit einem von 2,3. Werden nur die wissenschaftlichen Mitarbeiter betrachtet, so wird der Unterschied in der Arbeitszufriedenheit noch größer (Lehrorientierte 2,8/ nicht Lehrorientierte 2,3) und sehr signifikant. Bei den Hochschullehrern gibt es keinen signifikanten Unterschied. Die Unzufriedenheit der Lehrorientierten kann daher vor allem auf die Unzufriedenheit der lehrorientierten wiss. Mitarbeiter zurückgeführt werden.

Auch wenn man die einzelnen Aspekte der Arbeitszufriedenheit getrennt nach Lehrorientierten und nicht Lehrorientierten untersucht, so zeigt sich in der überwiegenden Anzahl der Aspekte die Tendenz, daß die Lehrorientierten unzufriedener sind. Auch dies gilt wieder vor allem für die Mitarbeiter. Die größten und signifikanten Unterschiede zwischen lehrorientierten und nicht lehrorientierten Mitarbeitern finden sich bei den Aspekten Arbeitsplatzsicherheit und berufliche Anerkennung. Auffällig ist, daß dies auch exakt die beiden Aspekte der Arbeitszufriedenheit sind, in denen sich die befragten Mitarbeiter von den befragten Hochschullehrern am stärksten unterscheiden. Es besteht also vermutlich ein Zusammenhang zwischen der Lehrorientierung der wiss. Mitarbeiter und ihrer Arbeitszufriedenheit.

Wenn Lehrorientierte also tendenziell unzufriedener sind und dies mit dem geringen Stellenwert der Lehre zusammenhängt, könnte man annehmen, daß sie Reformvorhaben aufgeschlossener gegenüberstehen, die den Stellenwert der Lehre erhöhen könnten, so z.B. der Einbeziehung studentischer Bewertungen oder der von Absolventen in leistungs-

orientierte Mittelvergabe. Hierzu lassen sich keine signifikanten Differenzen zwischen den Gruppen feststellen. Das kann einerseits daran liegen, daß dies möglicherweise durch einen Einfluß der Zugehörigkeit zu Statusgruppen o.ä. verdeckt wurde. Andererseits könnte tatsächlich kein solcher Zusammenhang existieren.

Probehalber wurden anhand theoretischer Überlegungen zu möglichen Determinanten der Akzeptanz solcher Reformvorhaben Regressionsmodelle gerechnet (unter Einbeziehung von Status und Befristung neben der Lehrorientierung). Sie ergaben aber keine verwertbaren Resultate. Werden jedoch in Anlehnung an Wottawa (2001) weitere Aspekte, z.B. solche der beruflichen Zufriedenheit, in die Überlegungen mit einbezogen und probehalber Modelle unter Einbeziehung dieser berechnet, so lassen sich z.T. signifikante Ergebnisse erzielen, bei denen sich die berufliche Anerkennung und Autonomie sowie die Zufriedenheit mit der Art der Verwaltung in der Hochschule als Determinanten der Akzeptanz auffassen ließen. Da gegenüber solchen probeweisen Berechnungen jedoch häufig der Vorwurf mangelnder Theoriegeleitetheit erhoben wird, soll dies hier nicht weiter verfolgt werden, sondern ein systematischer Zugang anhand der vorhandenen Modelle der Arbeitszufriedenheit bevorzugt werden. Hierbei wird davon ausgegangen, daß es – wie auch bereits weiter vorn erläutert – auch ohne eine direkte empirische Untersuchung einleuchtend ist, daß die Akzeptanz von Reformvorhaben längerfristig nur zu erzielen ist, wenn sie keine deutlichen negativen Effekte auf die Arbeitszufriedenheit haben. Umgekehrt könnte aber die Akzeptanz längerfristig durchaus steigen, wenn sich zeigt, daß Reformmaßnahmen positive Effekte auf die Arbeitszufriedenheit haben. Ein solcher Zusammenhang von Arbeitszufriedenheit und Akzeptanz tritt aber zeitverzögert ein. Deshalb wäre er nur mittels Längsschnittstudien adäquat überprüfbar, die bisher nicht durchgeführt wurden und leider auch im Rahmen dieser Arbeit aus Ressourcengründen nicht möglich waren.

Da hier jedoch – dem speziellen Forschungsinteresse mit dem Fokus auf die Hochschullehre folgend – die Lehrorientierung als einzelner Aspekt herausgegriffen wurde und nicht in das Gesamtkonstrukt Arbeitszufriedenheit eingeordnet wurde, soll die Betrachtung von Einzelaspekten an dieser Stelle beendet werden. Zunächst werden Modelle der Arbeitszufriedenheit vorgestellt und dann die Analysen zu „Determinanten der Arbeitszufriedenheit“ mit dem Gesamtkonstrukt fortgesetzt.

---

<sup>254</sup> Die größere Unzufriedenheit der lehrorientierten Mitarbeiter läßt sich – wie weitere Analysen zeigen –

## **6.5 Determinanten der beruflichen Zufriedenheit**

### **6.5.1 Konzepte zu Determinanten der beruflichen Zufriedenheit**

Wie bereits weiter vorn bei der Darstellung der Ergebnisse der Lehrendenbefragungen kurz angesprochen, können neben den offensichtlichen Einschätzungen der Lehrenden zu ihrer Arbeitssituation und daraus möglicherweise abzuleitenden Schlußfolgerungen weiterführend Zusammenhangsanalysen zu Determinanten der beruflichen Zufriedenheit durchgeführt werden. Diese ermöglichen auch Schlüsse über Zusammenhänge mit Aspekten der beruflichen Situation, die Lehrenden nicht immer bewußt sind und damit potentielle Ansatzpunkte zur Akzeptanzerhöhung. Die zentrale Frage dabei ist: Wovon hängt Zufriedenheit ab?

Für die Zwecke einer Analyse der Determinanten der Arbeitszufriedenheit und der Ableitung möglicher Schlußfolgerungen eignen sich die Lehrendenbefragungen an der TU Dresden, der Europa-Universität Frankfurt/ Oder und der Hochschule Zittau-Görlitz gut, da sie viele Items enthalten, die mit einschlägigen Untersuchungen übereinstimmen bzw. diesen entsprechen. Nach Fischer (1991) existiert zwar kein einheitliches Theoriekonzept für das Konstrukt Arbeitszufriedenheit, sondern vielmehr ein Konglomerat verschiedener theoretischer Modelle. Entsprechend der Übersicht von Westermann u.a. (1996) über Merkmale und Kategorien der Arbeits- und Studienzufriedenheit, die sowohl auf theoretischen Überlegungen als auch auf empirischen Analysen beruhen, sind die Kategorien jedoch insgesamt zu den hier gewählten relativ ähnlich. Von annähernd denselben Kategorien gehen auch Borg (2000: 3) und Weinert (1998: 214) aus.

Den höchsten Grad an Übereinstimmung mit eigenen theoretischen Überlegungen und den darausfolgenden Fragestellungen erzielt jedoch die Aufteilung in der „Skala zur Messung der Arbeitszufriedenheit“ (SAZ) nach Fischer und Lück (1972), ergänzt um das Merkmal „Anerkennung“ nach Locke (1976). Die Kategorien wurden in neueren Studien z.B. von Weinert (1987) sowie Neuberger und Allerbeck (1987) noch zusammengefaßt, jedoch nicht grundsätzlich weiterentwickelt oder in Frage gestellt. Interessant vor dem Hintergrund eines der Ziele dieser Arbeit, mit Studenten- und Lehrendenbefragungen Ansätze zur Erhöhung der Studien- und Arbeitszufriedenheit zu gewinnen, ist auch die Übertragung der Kategorien der Arbeitszufriedenheit auf das Studium von Apenburg (1980). Die Übertragung der

Kategorien von Apenburg (1980) konnte in einer jüngeren Studie von Westermann (1996) empirisch weitgehend bestätigt werden.

Versucht man, Determinanten der Arbeitszufriedenheit nach Schaeper (1995) gleichen Kategorien zuzuordnen, findet sich für viele mindestens eine Entsprechung.

### Übersicht zu Determinanten der beruflichen Zufriedenheit

nach Fischer und Lück (SAZ)	nach Locke	nach Weinert	nach Schaeper
<b>Arbeit</b>			
<ul style="list-style-type: none"> <li>- Bedingungen am Arbeitsplatz</li> <li>- Möglichkeiten zu persönlicher Entwicklung</li> <li>- Verhältnis zu Vorgesetzten</li> </ul>	<ul style="list-style-type: none"> <li>- Arbeitsbedingungen</li> <li>- Arbeitsinhalt</li> <li>- Vorgesetzte</li> </ul>	<ul style="list-style-type: none"> <li>- Arbeit selbst</li> <li>- Vorgesetzte</li> </ul>	<ul style="list-style-type: none"> <li>- Arbeitsbedingungen wie z.B. Arbeitszeitaufteilung</li> <li>- Autonomie</li> </ul>
<b>Erfolg</b>			
<ul style="list-style-type: none"> <li>- Bezahlung</li> <li>- Aufstiegsmöglichkeiten</li> </ul>	<ul style="list-style-type: none"> <li>- Bezahlung</li> <li>- Beförderung</li> <li>- Anerkennung</li> <li>- zusätzliche Vergünstigungen</li> </ul>	<ul style="list-style-type: none"> <li>- Bezahlung</li> <li>- Beförderungsmöglichkeiten</li> </ul>	<ul style="list-style-type: none"> <li>- Status (inkl. Bezahlung)</li> <li>- Befristung (unbefristete Stelle ist Voraussetzung für Beförderung)</li> </ul>
<b>Kontext</b>			
<ul style="list-style-type: none"> <li>- Verhältnis zu Kollegen</li> <li>- Verhaltensweisen von Management und Führung</li> </ul>	<ul style="list-style-type: none"> <li>- Kollegen</li> <li>- Firma / Institution und Management</li> </ul>	<ul style="list-style-type: none"> <li>- Mitarbeiter/Kollegen</li> </ul>	<ul style="list-style-type: none"> <li>- Sozialklima (u.a. „Disziplineffekt“ o.a. „Fächerkultur“)</li> </ul>

Quelle: überarbeitete Darstellung von Westermann u.a. (1996), ergänzt mit Schaeper 1995

Da in den für die vorliegende Arbeit durchgeführten Lehrendenbefragungen zu allen Kategorien der Arbeitszufriedenheit analog der SAZ Items vorhanden und die Frage zur

(allgemeinen) beruflichen Zufriedenheit analog zur SAZ gestellt wurden, erfolgt die Analyse von Determinanten beruflicher Zufriedenheit auf dieselbe bzw. ähnliche Weise.

### Gegenüberstellung von Determinanten beruflicher Zufriedenheit und Fragestellungen

Determinanten der beruflichen Zufriedenheit nach Fischer und Lück sowie nach Locke	Fragestellungen in Lehrendenbefragungen an der TU Dresden und der Europa-Universität Frankfurt/ Oder
<p><b>Arbeit/ Arbeitsbedingungen</b></p> <ul style="list-style-type: none"> <li>- Arbeitsbedingungen</li> <li>- Möglichkeiten zu persönlicher Entwicklung</li> <li>- Verhältnis zu Vorgesetzten</li> </ul>	<ul style="list-style-type: none"> <li>- Rahmenbedingungen der Lehrveranst., im Fachbereich, Arbeitszeitaufteilung und die Lehrveranst. selbst - Fragen 1, 2, 3, 4, 6a/d)</li> <li>- Möglichkeit, eigene Ideen zu verwirklichen - Frage 6f)</li> </ul>
<p><b>Erfolg</b></p> <ul style="list-style-type: none"> <li>- Bezahlung</li> <li>- Aufstiegsmöglichkeiten</li> <li>- Anerkennung</li> </ul>	<ul style="list-style-type: none"> <li>- Höhe des Gehalts - Frage 6h), Leistungsgerechtigkeit des Gehalts - F6i)</li> <li>- Aufstiegsmöglichkeiten - F6e)</li> <li>- berufliche Anerkennung - F6g)</li> </ul>
<p><b>Kontext</b></p> <ul style="list-style-type: none"> <li>- Verhältnis zu Kollegen</li> <li>- Verhaltensweisen von Management und Führung</li> </ul>	<ul style="list-style-type: none"> <li>- Verhältnis zu Kollegen im Fachbereich - f6b, Kooperation mit anderen Fachber. - F6c)</li> <li>- Informationen über Entscheidungen der Hochschulleitung - F6n), Art der Verwaltung in der Hochschule - F6j), Art der Verteilung finanzieller Mittel innerhalb der Hochschule/ der Fakultät - F6k/ F6l), Mitbestimmungsmöglichkeiten - F6m)</li> </ul>

Quelle: Eigene Darstellung, Krempkow (2001)

Das „Verhältnis zu Vorgesetzten“ wurde in den eigenen Lehrendenbefragungen nicht erfragt, da Hochschullehrer keine Vorgesetzten wie in Unternehmen der Privatwirtschaft haben. Schließlich ist ihre Tätigkeit durch einen Grundgesetzartikel zur Freiheit der Forschung und Lehre geschützt. Beamtenrechtlich „Dienstvorgesetzte“ wie z.B. Dekane

können lediglich als „primus inter pares“ gelten. Für wissenschaftliche Mitarbeiter sind zwar die Professoren Vorgesetzte, jedoch ist auch hier die Situation nur schwer mit der in der Privatwirtschaft vergleichbar.

### **6.5.2 Empirische Überprüfung der Konzepte**

Die Konzepte zu Determinanten beruflicher Zufriedenheit lassen sich auf verschiedenen, jeweils mehr oder weniger aussagekräftigen und leicht nachvollziehbaren Wegen empirisch überprüfen.

Am einfachsten nachvollziehbar sind wohl Subgruppenanalysen, da hierbei lediglich verschiedene Untergruppen der Befragten gegenübergestellt werden. Solche Subgruppenanalysen müssen – zumal wenn sie mit Signifikanzanalysen gekoppelt werden – nicht weniger zuverlässig sein als andere, kompliziertere und daher schwerer nachvollziehbare Analysen. Daher sollen hier mögliche Determinanten beruflicher Zufriedenheit zunächst durch Vergleich sich signifikant voneinander unterscheidender Befragtengruppen veranschaulicht werden.

### **6.5.3 Subgruppenanalysen**

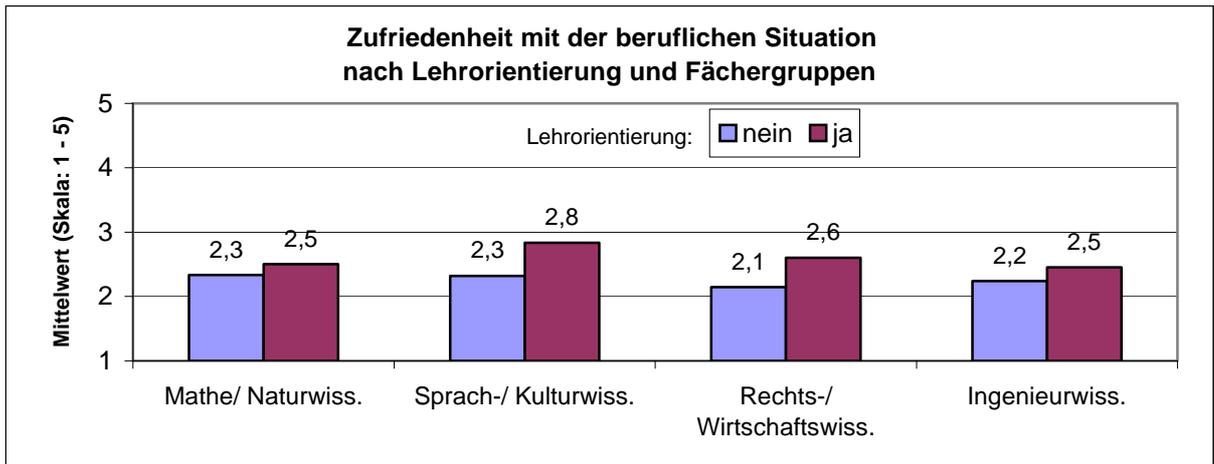
Für die Subgruppenanalysen (vgl. Schaeper 1995) soll exemplarisch die Lehrorientierung an der TU Dresden herangezogen werden. Hierzu wurden die Lehrenden anhand der auf die Vorlesungszeit bezogenen Antworten zum Zeitbudget in zwei Gruppen unterteilt, wobei die erste Gruppe der Lehrorientierten mit 21 Prozent aller Befragten diejenigen bilden, die gern mehr als die bisher von ihnen aufgewandte Arbeitszeit für die Lehre investieren würden. Die zweite Gruppe bilden mit 86 Prozent aller Befragten die Forschungsorientierten, die gern mehr Zeit (als bisher von ihnen aufgewandt) für die Forschung nutzen würden. 18 Prozent aller Befragten würden gern mehr Zeit sowohl für die Lehre als auch für die Forschung aufwenden. Dieser Anteil nicht klar zuzuordnender Lehrender wurde in jeder der beiden Gruppen berücksichtigt.

Für die Gruppe der Lehrorientierten wurde untersucht, inwieweit sie sich in Bezug auf die Arbeitszufriedenheit insgesamt sowie einzelne Aspekte der Arbeitszufriedenheit von den Nicht-Lehrorientierten unterscheidet. Die Untersuchung der Arbeitszufriedenheit insgesamt zeigt, daß Lehrorientierte mit einem Mittelwert von 2,5 signifikant unzufriedener sind als

Nicht-Lehrorientierte mit einem Mittelwert von 2,3. Dies gilt für wissenschaftliche Mitarbeiter noch stärker als für Hochschullehrer.

Die Analyse der einzelnen Aspekte der Arbeitszufriedenheit getrennt nach Lehrorientierten und Nicht-Lehrorientierten zeigte in der überwiegenden Anzahl der Aspekte ebenfalls die Tendenz, daß die Lehrorientierten unzufriedener sind. Auch dies gilt insbesondere für die Mitarbeiter.<sup>254</sup> Die größten signifikanten Unterschiede zwischen lehrorientierten und nicht lehrorientierten Mitarbeitern finden sich bei den Aspekten Arbeitsplatzsicherheit und berufliche Anerkennung. Es bestehen also offenbar Beziehungen zwischen Lehrorientierung und Aspekten der Arbeitszufriedenheit. Lehrorientierte könnten demnach möglicherweise durch eine höhere Arbeitsplatzsicherheit und Anerkennung ihrer Leistungen motiviert und damit die Lehre insgesamt gestärkt werden.

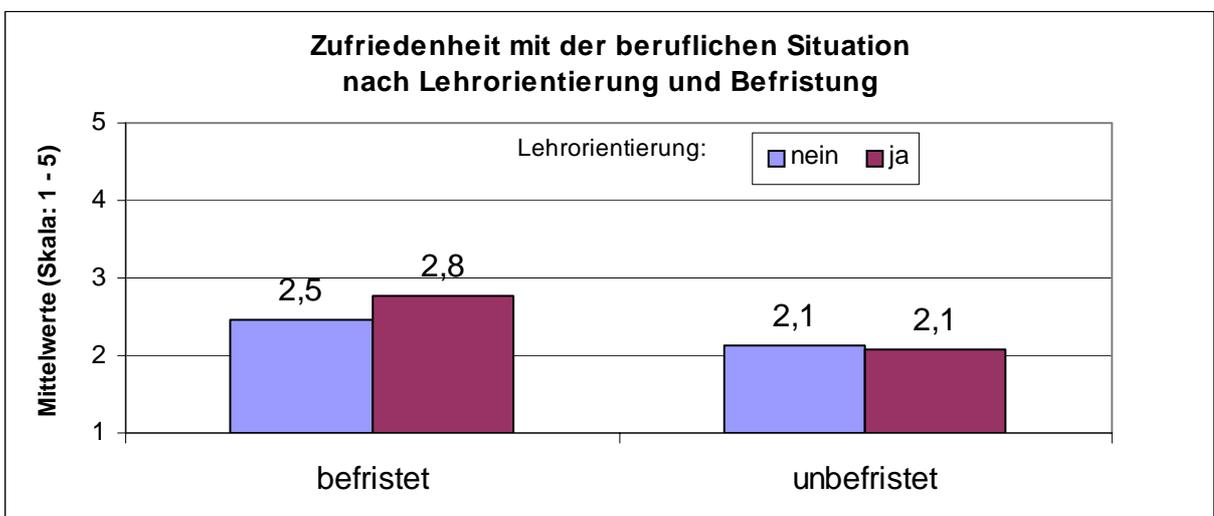
Dies könnten aber evtl. auch sogenannte „Scheinkorrelationen“ sein, die eigentlich stärker mit anderen Merkmal(en) als dem untersuchten zusammenhängen. Auf eine relativ einfache Weise läßt sich dies z.B. dadurch kontrollieren, daß man andere, potentiell „verzerrende“ Faktoren konstant hält, indem man die Befragten nach diesen Merkmalen aufteilt und das Vorhandensein der Zusammenhänge in den kleineren Untergruppen überprüft. Ist der Zusammenhang auch in den Unter- oder Subgruppen feststellbar, so kann man davon ausgehen, daß es sich bei dem anfangs festgestellten um einen – zumindest von den untersuchten Faktoren – unbeeinflussten Zusammenhang handelt. Dies läßt sich natürlich mittels komplexerer Analysen in vielen Fällen besser, aber bei weitem nicht so anschaulich überprüfen. Beispielsweise wäre es vorstellbar, daß die Unterschiede in der Arbeitszufriedenheit weniger ein Effekt der Lehrorientierung sind, sondern vielmehr aus der Fächerkultur („Disziplineffekt“) oder aus anderen Merkmalen der Befragten, wie z.B. Status oder Befristung des Arbeitsverhältnisses resultieren (vgl. Schaeper 1995: 146). Zunächst soll daher überprüft werden, ob die Lehrorientierten auch in allen Fächergruppen unzufriedener sind. Wie nachfolgende Graphik zeigt, ist dies der Fall:



Quelle: Krempkow (2001)

Die Mittelwertunterschiede sind allerdings in den einzelnen Fächergruppen (auch aufgrund der relativ gering werdenden Fallzahl von  $n=93$  bis  $n=20$ ) nicht mehr signifikant.

Da die Untersuchung von Statureffekten bereits weiter vorn erfolgte und keine signifikanten Unterschiede nach wissenschaftlichen Mitarbeitern und Hochschullehrern ergab, wird anhand einer Subgruppenanalyse noch die Wirkung der Befristung untersucht: Die Gegenüberstellung der befristet und unbefristet Beschäftigten nach Lehrorientierung zeigt bei ersteren eine deutlich niedrigere Zufriedenheit. Bei letzteren ist dies nicht der Fall. Der positive Effekt der unbefristeten Beschäftigung hebt also evtl. den negativen Effekt der Lehrorientierung auf. Hieraus kann zunächst geschlossen werden, daß die Befristung neben der Lehrorientierung evtl. eine eigene Determinante beruflicher Zufriedenheit darstellt.



Quelle: Krempkow (2001)

So anschaulich Subgruppenanalysen aber auch sein mögen, sie sind in ihrer Aussagekraft begrenzt. Dies gilt insbesondere dann, wenn nicht nur zwei Gruppen von Befragten miteinander verglichen werden sollen, sondern der Einfluß weiterer Aspekte wie z.B. der Rahmenbedingungen der Lehre auf die Arbeitszufriedenheit untersucht werden soll. Dies geschieht daher nachfolg anhand von Zusammenhangsanalysen. Da – wie sich bei den weiteren Analysen herausstellte – zwar die Stärke der Korrelationen, nicht jedoch deren Richtung zwischen wissenschaftlichen Mitarbeitern und Hochschullehrern variiert und die Fallzahlen für einige Analysen unter dem optimalen Stichprobenumfang liegen würden, wird nachfolgend auf die getrennte Analyse von wissenschaftlichen Mitarbeitern und Hochschullehrern verzichtet. Anderenfalls bestünde die Gefahr, Zusammenhänge systematisch zu „übersehen“.

## 6.5.4 Bivariate und multivariate Zusammenhangsanalysen

Die Analyse der Zusammenhänge wurde in zwei Schritten durchgeführt (vgl. Schaeper 1995). Hierzu wurden erstens Korrelationen<sup>255</sup>, und zweitens multiple Regressionen berechnet (vgl. Backhaus u.a. 2000: 5; Benkhoff u.a. 2001: 20). Die Ergebnisse werden für die TU Dresden und für alle drei Hochschulen dargestellt.

### Übersicht zu signifikanten Korrelationen mit der beruflichen Zufriedenheit insgesamt (p<0,05)

Fragestellungen in Lehrendenbefragungen an der TU Dresden, der Universität Frankfurt/ Oder und der Hochschule Zittau-Görlitz	Korrelationskoeffizienten	
	TU Dresden	alle drei Hochschulen
<b>Arbeit/ Arbeitsbedingungen:</b>		
- Rahmenbedingungen der Lehrveranstaltungen, im Fachbereich, Arbeitszeitaufteilung und die Lehrveranstaltungen selbst:		
Verfügbarkeit von Räumen – Frage 1a).....	.15*	.14*
Akustik der Räume – Frage 1c).....	.18*	.24**
Ausstattung der Räume – Frage 1d).....	-	.15**
Verfügbarkeit der für die LV nötigen Literatur in der Bibliothek – Frage 1e).....	.15*	.16**
Verfügbarkeit von Kommunikationsorten für Lehrende & Studierende – Frage 1f).	.16*	.18**
Aufbau und Struktur des Studiums – Frage 2a).....	.16*	.14*
Umfang des Lehrangebotes – Frage 2b).....	.17*	.16**
Klarheit der Prüfungsanforderungen – Frage 2f).....	.22**	-
Engagement der Lehrenden im Studiengang – Frage 2g).....	.16*	.17**
Fachliche Kompetenz der Lehrenden – Frage 2h).....	.25**	.20**
Beratungs- und Betreuungsqualität – Frage 2i).....	.22**	.18**
Situation der Lehre im Vergleich zu vorhergehenden Semestern – Frage 3).....	.16*	.21**
Forschung, tatsächlicher Anteil (in %), vorlesungsfreie Zeit – Frage 4).....	-.16*	-
Lehrveranstaltungen – Frage 6a).....	.24**	.29**
Sicherheit des Arbeitsplatzes – Frage 6d).....	.33**	.26**
- Möglichkeit, eigene Ideen zu verwirklichen - Frage 6f).....	.43**	.45**
<b>Erfolg:</b>		
- Höhe des Gehalts - Frage 6h).....	.27**	.26**
- Leistungsgerechtigkeit des Gehalts - F6i).....	.22**	.24**
- Aufstiegsmöglichkeiten - F6e).....	.37**	.40**
- berufliche Anerkennung - F6g).....	.53**	.57**
<b>Kontext:</b>		
- Verhältnis zu Kollegen im Fachbereich f6b).....	.22**	.23**
- Kooperation mit anderen Fachbereichen - F6c).....	.19**	.15**
- Informationen über Entscheidungen der Hochschulleitung - F6n).....	.29**	.29**
- Art der Verwaltung in der Hochschule - F6j).....	.26**	.29**
- Art der Verteilung finanzieller Mittel innerhalb der Hochschule - F6k).....	.16*	.25**
- Art der Verteilung finanzieller Mittel innerhalb der Fakultät - F6l).....	.20**	.26**
- Mitbestimmungsmöglichkeiten - F6m).....	.33**	.35**

<sup>255</sup> Es wurde von Intervallskalenniveau ausgegangen. Es ist eine alte Diskussion, ob hier Ordinal- oder Intervallskalenniveau vorliegt. Diese wird z.B. in Bortz/ Döring (2002: 73) erläutert. Demnach sind die Konsequenzen der Annahme eines höheren Skalenniveaus nicht so gravierend, daß die Grundaussagen einer Untersuchung hiervon generell beeinträchtigt wären, da die Bestätigung einer Hypothese durch die Annahme dieses Skalenniveaus eher erschwert wird. Ein hier zusätzlich vorgenommener Vergleich der Korrelationskoeffizienten nach Spearman (für die Annahme von Ordinalskalenniveau) und nach Pearson (Intervallskalenniveau) zeigt nur sehr geringe Unterschiede.

Hierbei werden für die berufliche Zufriedenheit sehr signifikante  $r > .20^{**}$  als bedeutsam eingeschätzt (fettgedruckt dargestellt). Die stärksten Zusammenhänge finden sich mit der beruflichen Anerkennung ( $r = .53^{**}$ ) und den Möglichkeiten, eigene Ideen zu verwirklichen ( $r = .43^{**}$ ). Deutlich schwächer ( $r > .30^{**}$ ) sind die Korrelationen mit den Aufstiegs- und Mitbestimmungsmöglichkeiten, sowie an der TU Dresden mit der Sicherheit des Arbeitsplatzes. Weitere signifikante Zusammenhänge an der TU Dresden und an allen drei Hochschulen insgesamt zeigten sich mit den Informationen über Entscheidungen der Hochschulleitung, der Art der Verwaltung in der Hochschule, der Leistungsgerechtigkeit und Höhe des Gehalts, dem Verhältnis zu Kollegen am Fachbereich und den Lehrveranstaltungen (jeweils  $r = .2^{**}$  bis  $.3^{**}$ ).

Außerdem ergibt sich eine signifikante Beziehung zur beruflichen Zufriedenheit bei bivariaten Analysen der nominalskalierten Variablen für die Befristung (des Arbeitsverhältnisses). Hierzu wurden zunächst Chi-Quadrat-Tests für die Signifikanz durchgeführt und anschließend Eta als Zusammenhangsmaß berechnet. Der Einfluß der Befristung kann mit  $p = 0,00$  und  $\text{Eta} = .29$  empirisch belegt werden.<sup>256</sup>

Im zweiten Schritt wurden die Zusammenhänge mittels multivariater Regressionsanalysen überprüft (vgl. Schaeper 1995: 141). Allerdings muß hierzu angemerkt werden, daß für diese Analysen eine größere Anzahl von Befragten als für bivariate Zusammenhangsanalysen nötig ist (vgl. Bortz/ Döring 2002: 613 ff.). So ist vor vornherein davon auszugehen, daß nur große Effekte bzw. die wesentlichsten Einflußfaktoren überprüft und ggf. bestätigt werden können. Vorteile dieses Analyseverfahrens sind jedoch, daß Zusammenhänge der Einflußfaktoren untereinander berücksichtigt werden und die Einflußgrößen unterschiedlich skaliertes (z.B. auch ordinal und nominal skaliertes) Variablen vergleichbar werden. Nachfolgend sollen daher Vorgehensweise und Ergebnisse der multiplen Regressionsanalysen dokumentiert werden.

Ausgegangen wurde bei dieser Analyse zunächst wieder von einem Modell nach Schaeper (1995). Anschließend wurden die Analysen von Westermann u.a. (1996) hinzugezogen. Wesentlich wären nach Schaeper für die Arbeitszufriedenheit (von Lehrenden) v. a. der Disziplineffekt (Fächerkultur), der Statuseffekt (Gruppenzugehörigkeit) und der Effekt der

---

<sup>256</sup> Es lassen sich – wie ein Überblick über eine bivariate Gesamtanalyse zeigt – weitere sehr signifikante Zusammenhänge finden, die zwar schwach, aber inhaltlich durchaus plausibel sind. Da diese aber m.W. bisher nicht in Modellen der Arbeitszufriedenheit verwendet wurden und hier nicht das Ziel einer Weiterentwicklung solcher Modelle im Vordergrund steht, sollen diese Variablen bei weiteren Analysen unberücksichtigt bleiben. Diese Variablen sind: f9g,h,i,j und f14e.

Befristung des Arbeitsvertrages. Es zeigt sich jedoch, daß wie auch schon bei der separaten Betrachtung der potentiellen Einflußfaktoren von diesen dreien nur die Befristung einen signifikanten Effekt auf die Arbeitszufriedenheit hat<sup>257</sup>.

Da dieses Regressionsmodell<sup>258</sup> mit gerade einmal 7% erklärter Varianz (R-Quadrat) eine „eher bescheidene“ Erklärungskraft hat (vgl. Enders/ Bornmann 2001: 195), wurde unter Hinzuziehung der Überlegungen von Westermann ein neues Modell formuliert. Hierbei wurden zusätzlich alle Aspekte aufgenommen, die nach Westermann u.a. (1996) potentielle Determinanten für die Arbeitszufriedenheit darstellen.<sup>259</sup> Mit diesem Modell können 43% bzw. 37% der Varianz erklärt werden, was bereits über den Werten ähnlicher Untersuchungen liegt.

Signifikante Einflußfaktoren sind nach diesem Modell (mit Einflußgröße Beta):<sup>260</sup>

- f6g: berufliche Anerkennung	(.22*)
- f6f: Möglichkeit, eigene Ideen zu verwirklichen	(.20*)
- f6a: Lehrveranstaltungen	(.16*)

Das erweiterte Regressionsmodell kann für die Ausgangsbedingungen als befriedigend eingeschätzt werden. Dennoch könnte die Erklärungskraft evtl. verbessert werden.

Da das korrigierte R-Quadrat vor allem von der Anzahl der einbezogenen Variablen abhängt (vgl. Backhaus u.a. 2000: 24), könnte durch Ausschluß aller Variablen, die nach den Ergebnissen der Regressionsanalyse nicht oder kaum mit der zu erklärenden abhängigen Variable zusammenhängen, die Erklärungskraft eines hierdurch vereinfachten Modells ggf. noch verbessert werden (vgl. auch Rostampour/ Lemberg 2003). Dies wurde versucht.

<sup>257</sup> Hierbei wurden als Kriterien die voreingestellten Signifikanzwerte für die Wahrscheinlichkeit des F-Wertes für die Aufnahme verwendet (p für Einschluß  $\leq 0,05$ , Ausschluss  $\geq 0,10$ ).

<sup>258</sup> Eine Dokumentation zentraler Ergebnisse der Regressionsanalysen anhand von SPSS-Outputs erfolgt im Anhang D dieser Arbeit.

<sup>259</sup> Da eine Hauptkomponentenanalyse über alle Items zur Arbeitssituation keine verwertbaren Resultate brachte, wurden von diesen alle als bedeutsam eingeschätzten Variablen im Modell einbezogen.

<sup>260</sup> In wissenschaftlichen Publikationen wird z.T. diskutiert, ob der Begriff „Einfluß“ bei Zusammenhangsanalysen gerechtfertigt ist. Für Korrelationen, wird dies im allgemeinen verneint, für Regressionsanalysen bejaht. Außerdem wurde in diesem Fall ergänzend zu den Regressionsanalysen eine Kausalanalyse mit dem Programmpaket AMOS durchgeführt (vgl. Arbuckle/ Wothke 1999). Die für die TU Dresden exemplarisch durchgeführte Kausalanalyse ergab (mit  $R^2=0,43$ ) fast dieselben Determinanten der Arbeitszufriedenheit, wobei ebenfalls die Anerkennung (.24) den stärksten Einfluß hatte, gefolgt von Lehrveranstaltungen (.17) und Befristung (-.17) sowie den Möglichkeiten, eigene Ideen zu verwirklichen (.11). Da auch bei gegebenen Verteilungsvoraussetzungen viele AMOS-Ergebnisse nur approximativ gelten, d.h. bei genügend großen Stichproben (vgl. Baltes-Götz 2002: 21), wurden im folgenden Regressionsanalysen favorisiert.

Jedoch konnte in diesem speziellen Fall mit dem vereinfachten Modell weder die Erklärungskraft verbessert (R-Quadrat = 0,37 bzw. 0,36) noch die Kolinearität verringert werden<sup>261</sup>. Die Beta-Koeffizienten verschieben sich z.T. etwas, so daß hier die Befristung wichtiger wird als die (Zufriedenheit mit) Lehrveranstaltungen. Da der wichtigste Aspekt jedoch auch bei diesem vereinfachten Modell die berufliche Anerkennung bleibt, wird das vereinfachte Modell hier verworfen.

Ein anderes Bild ergab die Auswertung über alle drei beteiligten Hochschulen. Hier war das vereinfachte Modell (korr. R-Quadrat=0,40) gegenüber dem erweiterten Modell (unter Einbeziehung aller bedeutsam korrelierenden Einflußgrößen) für die inhaltliche Interpretation besser geeignet. Es brachte jedoch sehr ähnliche Ergebnisse: Berufliche Anerkennung (.42\*\*), Möglichkeit, eigene Ideen zu verwirklichen (.20\*\*), Lehrveranstaltungen (.17\*\*). Zwar wird auch in ähnlich gelagerten Untersuchungen keine bessere Erklärungskraft erzielt und dies wurde als ausreichend eingeschätzt (vgl. Enders/ Bornmann 2001: 188<sup>263</sup>). Dennoch ist dies möglicherweise auf anderen Wegen noch weiter zu verbessern.

Da nach Schaeper (1995) und nach den Antworten der hier Befragten angenommen werden kann, daß für befristet und unbefristet Beschäftigte jeweils andere Aspekte wichtig sind, soll nachfolgend untersucht werden, ob durch eine getrennte Analyse für diese beiden Befragten-gruppen (Partialmodelle) die Erklärungskraft weiter verbessert werden kann.

Hierbei ist allerdings zu beachten, daß die Fallzahl für jede Analyse entsprechend kleiner wird.<sup>264</sup> Dies muß bei der nachfolgenden Analyse in Kauf genommen werden. In einem ersten Schritt sollen zunächst wieder alle potentiellen Einflußvariablen einbezogen werden.

---

<sup>261</sup> Die Variance Inflation Factors (VIF) liegen alle weit unter 10. Der nach Durbin/ Watson durchgeführte Test auf Autokorrelation ergab einen Koeffizienten sehr nahe dem Optimum 2, so daß er kaum noch verbessert werden konnte (vgl. Brosius 2002: 559 und 564; Backhaus u.a. 2000: 39 ff.).

<sup>263</sup> Enders und Bornmann untersuchten in ihrer Studie ca. 2000 Promovierte zu Ausbildung, Berufsverlauf und Berufserfolg, wobei sie auch umfassende Regressionsanalysen zu Determinanten des beruflichen Erfolgs und der Zufriedenheit vorlegten. Für die Zufriedenheit konnten in verschiedenen Modellen maximal 23% der Varianz erklärt werden. Hierbei (wie auch bei der Lehrendenbefragung) ist zu beachten, daß es sich bei den Befragten im allgemeinen um stark individualisierte Persönlichkeiten handelt, die gerade im Wissenschaftsbetrieb aus den unterschiedlichsten Motiven beruflich tätig sind (vgl. Enders/ Bornmann 2001: 194).

<sup>264</sup> Nach Bortz/ Döring (2002: 613 ff.) können bei Fallzahlen unter ca. 100 Befragten nur noch „große Effekte“ statistisch zuverlässig erkannt werden, kleinere und mittlere hingegen nicht. (80 Befragte gaben in der hier untersuchten Befragung an, befristet beschäftigt, 70 gaben an, Hochschullehrer zu sein. Für die Analyse aller drei Hochschulen liegt die Fallzahl jeweils über 100.)

Im zweiten Schritt werden wie zuvor nur noch Variablen mit einer Signifikanz  $p < 0,2$  einbezogen. Da die grundlegende Vorgehensweise dieselbe ist wie bei den zuvor ausführlicher dargestellten Analysen, werden diese nachfolgend lediglich beschrieben und nicht noch einmal ausführlich dokumentiert.

### **Partialmodell a) befristet Beschäftigte**

Im ersten Schritt können zwar ca. 44 bzw. 31% der Varianz im Partialmodell der Regression für befristet Beschäftigte erklärt werden, es ist jedoch kein Einflußfaktor signifikant. Interessant erscheint hierbei, daß auch bei separater Betrachtung der befristet Beschäftigten die Einschätzung der Sicherheit des Arbeitsplatzes nicht bzw. nicht statistisch belegbar mit der beruflichen Zufriedenheit zusammenhängt (zumindest nicht als „großer Effekt“ nach Bortz/ Döring (2002: 613 ff.)). Im zweiten Schritt wird daher wieder ein vereinfachtes Partialmodell erstellt.

Mit dem vereinfachten Partialmodell können rund 38 bzw. 35% der Varianz für befristet Beschäftigte erklärt werden. Das korrigierte R-Quadrat fällt damit besser aus. Außerdem sind nunmehr drei der Einflußfaktoren signifikant:  $f_{6g}$ : berufliche Anerkennung (.37\*\*),  $f_{6f}$ : Möglichkeit, eigene Ideen zu verwirklichen (.23\*) und  $f_{6i}$ : Leistungsgerechtigkeit des Gehalts (.20\*).

Beide Partialmodelle für befristet Beschäftigte sind dennoch schlechter geeignet als das Gesamtmodell, die Varianz der Zufriedenheit mit der beruflichen Situation zu erklären. Dies liegt aber vielleicht auch an der Heterogenität der Orientierungen wissenschaftlicher Mitarbeiter, die den Großteil der befristet Beschäftigten stellen.

### **Partialmodell b) unbefristet Beschäftigte**

Mit rund 39 % kann im Partialmodell der Regression für unbefristet Beschäftigte ein etwas höherer Anteil der Varianz erklärt werden als im Gesamtmodell. Allerdings ist mit  $f_{6f}$  nur ein Beta-Koeffizient signifikant:  $f_{6f}$ : Möglichkeit, eigene Ideen zu verwirklichen (.31\*). Die Variable  $f_{6j}$  (Art der Verwaltung in der Hochschule) ist mit  $p=0,07$  und  $Beta=.20$  lediglich annähernd signifikant. Nachfolgend wurde mit diesen Variablen wieder ein vereinfachtes Partialmodell erstellt. Zwar sind im vereinfachten Modell diese zwei Einflußfaktoren nunmehr signifikant:  $f_{6f}$ : Möglichkeit, eigene Ideen zu verwirklichen (.36\*\*),  $f_{6j}$ : Art der Verwaltung in der Hochschule (.30\*). Hierbei werden jedoch nur 25 bzw. 24% der Varianz

erklärt, so daß das vereinfachte Partialmodell als weniger gut geeignet einzuschätzen ist, die Varianz der Zufriedenheit mit der beruflichen Situation zu erklären. Somit kann lediglich das Partialmodell für unbefristet Beschäftigte einen etwas besseren Beitrag zur Erklärung der Einflüsse auf die berufliche Zufriedenheit liefern als das Gesamtmodell in Anlehnung an Westermann u.a. (1996). Insgesamt gesehen haben die Partialmodelle zur Befristung keine bessere Erklärungskraft als das Gesamtmodell. Ähnlich fielen die Ergebnisse für die Analyse über alle drei Hochschulen aus, daher wurden diese hier nicht noch einmal dargestellt. Anschließend werden zur Überprüfung eines weiteren potentiell wesentlichen Einflußfaktors außerdem Partialmodelle getrennt nach Status (für Hochschullehrer und wissenschaftliche Mitarbeiter) berechnet.

### **Partialmodell 1) Hochschullehrer**

Für die Hochschullehrer können mit diesem Modell durchaus befriedigende 67 bzw. rund 52% der Varianz erklärt werden, wobei ein Beta-Koeffizient relativ hoch ausfällt.

Vier Beta-Koeffizienten liegen über 0,2 (f2f, f6a, f6g, f6j), jedoch nur ein Einflussfaktor war signifikant: Klarheit der Prüfungsanforderungen (0,36\*). Da ein Einflussfaktor nur wenig Ansatzpunkte für praktische Schlussfolgerungen bietet und die Erklärungskraft evtl. noch zu verbessern ist, wird auch hier ein vereinfachtes Modell gerechnet, welches „nur“ 45% Erklärungskraft hat (was aber deutlich über dem Gesamtmodell liegt). Demnach gibt es drei signifikante Einflussfaktoren:

- f6g: berufliche Anerkennung	(.42**)
- f2f: Klarheit der Prüfungsanforderungen	(.30**)
- f6j: Art der Verwaltung in der Hochschule	(.16*)

Das Modell für einen Gesamtdatensatz aller drei Hochschulen ergibt 38% Erklärungskraft. Die berufliche Anerkennung ist (mit  $\text{Beta}=.52^{**}$ ) auch hier die wichtigste Determinante. Das negative Vorzeichen zur Arbeitsplatzsicherheit ( $\text{Beta}=-.20^{**}$ ) zeigt an, daß bei den Hochschullehrern eine bessere Einschätzung mit insgesamt geringerer beruflicher Zufriedenheit einher geht.

## Partialmodell 2) Wissenschaftliche Mitarbeiter

Für die Wissenschaftlichen Mitarbeiter können mit diesem Partialmodell immerhin 31% der Varianz erklärt werden. Auch hier ist wie beim Partialmodell für Hochschullehrer zunächst nur ein Einflußfaktor signifikant: f6f: Möglichkeit, eigene Ideen zu verwirklichen (.23\*). Für die Variable f6g (berufliche Anerkennung) ist in diesem Modell der Beta-Koeffizient von .21 mit  $p=0,7$  lediglich annähernd signifikant. Die Vereinfachung des Regressionsmodells konnte aber noch eine Verbesserung erzielen. Die Erklärungskraft konnte auf rund 34% gesteigert und die Durban-Watson-Statistik verbessert werden. Zudem sind drei der untersuchten Einflussfaktoren nun signifikant:

- f6g: berufliche Anerkennung	(.38**)
- Befristung	(.22**)
- f6f: Möglichkeit, eigene Ideen zu verwirklichen	(.20**)

Die Einbeziehung aller drei Hochschulen ergibt für die wissenschaftlichen Mitarbeiter ein Partialmodell mit 29% Erklärungskraft und fünf signifikanten Variablen. Den stärksten Einfluß haben hierbei die Möglichkeit, eigene Ideen zu verwirklichen (.28\*\*) und die Arbeitsplatzsicherheit (.23\*\*). Deutlich geringeren Einfluß haben das Verhältnis zu Kollegen im Fachbereich (.20\*\*), die Akustik der Räume (.20\*\*).

Nachfolgend werden die signifikanten Ergebnisse der multiplen Regressionsanalyse zusammenfassend in einer Übersicht dargestellt. Die Werte über alle drei Hochschulen berechneten Werte wurden in eckige Klammern gesetzt. Die signifikanten Ergebnisse, die sich nicht nur für die TU Dresden, sondern über alle drei Hochschulen ergaben, sind durch Fettdruck hervorgehoben, da diese tendenziell verallgemeinerbar sein könnten.

## Ergebnisse der Regressionsanalysen – Gesamtmodell und Partialmodelle nach Status

(TU Dresden/ [alle drei Hochschulen], über TU und 3 Hochschulen hinweg stabile Ergebnisse **fettgedruckt**)

Determinanten der beruflichen Zufriedenheit	Gesamtmodell (alle Lehrenden) korr. R <sup>2</sup> =37%/ [40%]	Partialmodell 1 (Hochschullehrer) korr. R <sup>2</sup> =45%/ [38%]	Partialmodell 2 (wiss. Mitarbeiter) korr. R <sup>2</sup> =34%/ [29%]
<b>Arbeit/</b>			
<b>Arbeitsbedingungen:</b>			
- Arbeitsbedingungen	- <b>Lehrveranstaltungen (Beta=.16*/ [.17**])</b>	- Klarheit der Prüfungsanforderungen (Beta=.30**)/ Arbeitsplatzsicherheit [Beta=	- <b>Befristung (Beta=.22*) / Arbeitsplatzsicherheit [Beta=.23**]</b>
- Möglichkeiten zu persönlicher Entwicklung	- <b>Möglichkeit, eigene Ideen zu verwirklichen (Beta=.20**/ [.20**])</b>	- .20**] - n.s.	- <b>Möglichkeit, eigene Ideen zu verwirklichen (Beta=.20*/ [.28**])</b>
<b>Erfolg:</b>			
- Bezahlung	- nicht signifikant (n.s.)	- n.s.	- n.s.
- Aufstiegsmöglichkeiten	- n.s.	- n.s.	- n.s.
- Anerkennung	- <b>berufliche Anerkennung (Beta=.22*/ [.42**])</b>	- <b>berufl. Anerkennung (Beta=.2**/ [.56**])</b>	- berufliche Anerkennung (Beta=.38**)
<b>Kontext:</b>			
- Verhältnis zu Kollegen im Fachbereich/ Kooperation mit anderen Fachbereichen	- n. s.	- n. s.	- n. s./ Verhältnis zu Kollegen im Fachbereich (Beta=.20**)
- Verhaltensweisen von Management und Führung	- n.s.	- Art der Verwaltung in der Hochschule (Beta=.16*)	- n. s.

### **6.5.5 Fazit zu Determinanten der beruflichen Zufriedenheit**

Mit Hilfe von Korrelations- und multiplen Regressionsanalysen wurde untersucht, welche Faktoren wie stark mit der beruflichen Zufriedenheit zusammenhängen. Mittels Regressionsverfahren wurden auch Zusammenhänge dieser Faktoren untereinander berücksichtigt. Im Ergebnis dieser Analysen konnte am Beispiel der TU Dresden und für alle drei Hochschulen insgesamt dargestellt werden, daß für die berufliche Zufriedenheit von Lehrenden vor allem die berufliche Anerkennung wichtig ist. Daneben sind auch die Möglichkeiten, eigene Ideen zu verwirklichen (Autonomie) sowie die Zufriedenheit mit den Lehrveranstaltungen von großer Bedeutung.

In Partialmodellen für Hochschullehrer bestätigt sich der überragende Einfluß der beruflichen Anerkennung. Für wissenschaftliche Mitarbeiter haben auch beruflicher Freiraum und Befristung bzw. Arbeitsplatzsicherheit Einfluß auf die Zufriedenheit.

Daß Höhe und Leistungsgerechtigkeit des Gehalts auch relevant sein können, ist zwar nicht mit letzter Sicherheit auszuschließen. Aufgrund des hier im Rahmen einer Dissertation ohne weitere finanzielle Mittel maximal zu leistenden Umfangs der Untersuchung konnten nur mittlere bis große Effekte statistisch zuverlässig isoliert werden. Dennoch ist als wahrscheinlich anzunehmen, daß dies gegenüber erstgenannten Faktoren für die berufliche Zufriedenheit der befragten Lehrenden eher nachrangig ist. Ergänzend zu der aus den Experteninterviews abgeleiteten These, daß Fairneß bzw. Leistungsgerechtigkeit eine wesentliche Rolle für die Akzeptanz von Leistungsbewertungen und Anreizmodellen spielen, zeigt sich hier also, daß unter Einbeziehung der Arbeitszufriedenheit als Möglichkeit zur längerfristigen Erzielung von Akzeptanz auch die berufliche Anerkennung und Autonomie wichtig sein können.

## 6.6 Fazit zu Bedingungen erfolgreicher Institutionalisierung

Am Beginn dieses Kapitels wurde herausgearbeitet, daß die Akzeptanz von Reformvorhaben bei den Betroffenen – insbesondere bei einer relativ starken Rechtsstellung und Existenzsicherung – als wesentlicher Ansatzpunkt für die Diskussion von Bedingungen erfolgreicher Institutionalisierung gesehen wird. Weiterhin wurde herausgearbeitet, daß in diese Diskussion die Arbeitszufriedenheit einbezogen werden sollte, da hierin wiederum potentielle Ansatzpunkte zur Sicherung bzw. Erhöhung der Akzeptanz gesehen werden.

Bei einer direkten Akzeptanzuntersuchung mit Hilfe der Frage, inwieweit Lehrende bestimmten hochschulpolitischen Reformvorhaben zustimmen, zeigte sich, daß bei in den Jahren 2001 bis 2003 an drei ostdeutschen Hochschulen durchgeführten Lehrendenbefragungen deutlich weniger Zustimmung zu Reformvorhaben zur leistungsorientierten Mittelzuweisung zu finden war als bei einer von Kopp/ Weiß (1993) durchgeführten bundesweiten Lehrendenbefragung. Allerdings wurden gerade diese Antwortvorgaben konkretisiert, so daß die Ergebnisse nur bedingt vergleichbar sind. Wie andere, zusätzlich zu den in der bundesweiten Lehrendenbefragung verwendeten Antwortvorgaben zeigten, hängt die Zustimmung aber auch wesentlich von der Formulierung ab. Erfragt man die Zustimmung zur „Einbeziehung“ bestimmter subjektiver Bewertungen von Experten, Absolventen oder Studenten, so zeigt sich eine deutlich höhere und für die ersten beiden Gruppen sogar eine mehrheitliche Zustimmung (im Gegensatz zu der in Anlehnung an die bundesweite Befragung verwendeten Formulierung).

Über die verwendeten unterschiedlichen Frageformulierungen stabil ist allerdings die relative Abfolge der Zustimmung zu den einzelnen Gruppen von Bewertenden: Die (Einbeziehung der Bewertungen von) Experten, zu denen sich die Professoren selbst zählen können, erfährt die stärkste Akzeptanz, während die Gruppe der Absolventen (zu der sich die Mitarbeiter zählen können) etwas weniger und die der Studenten deutlich weniger Zustimmung erhält. Die Zustimmung zur Verwendung von subjektiven Bewertungen sinkt also mit dem Status der Gruppe der Bewertenden. Daß Absolventenbefragungen tendenziell stärker akzeptiert werden als Studentenbefragungen, hatten auch die im Rahmen der Experteninterviews befragten Experten ohne Kenntnis der Ergebnisse der Lehrendenbefragungen an ihren Hochschulen so eingeschätzt (vgl. Abschnitte 4.4.2 und 4.4.3 in dieser Arbeit).

Geht man vom derzeitigen Stand der Akzeptanz aus, kann die Einbeziehung subjektiver Bewertungen in Leistungsbewertungen und Leistungsanreize also entweder wie bisher (z.B. bei Evaluationen) durch Experten erfolgen, oder aber zunächst über Absolventenbefragungen etabliert werden. Zwar haben diese den Nachteil, daß Aspekte der Prozeßqualität von Lehre und Studium nur mit großem zeitlichen Abstand erfasst werden können. Eine Durchsetzung studentischer Bewertungen, die bei ähnlicher Datenqualität eine größere Zeitnähe ermöglichen, würde (wenn dies überhaupt möglich ist) aber wahrscheinlich eher zu Unterlaufungs- oder Umgehungsstrategien führen und daher unter den derzeitigen Rahmenbedingungen nicht die gewünschten Ziele erreichen. Auf längere Sicht sollten, wenn die Prozeßqualität in der Qualitätsorientierung eine wichtige Rolle spielt, Studentebefragungen aber nicht aus der Diskussion ausgeschlossen werden. Möglicherweise lassen sich mittel- oder langfristig Ansatzpunkte zur Erhöhung der Akzeptanz finden.

Vermutlich kann die geringe Akzeptanz insbesondere von studentischen Bewertungen zumindest teilweise auch auf konfligierende Interessen bestimmter Statusgruppen zurückgeführt werden. Inwieweit dies der Fall sein kann, versuchte Rindermann anhand einer inhaltsanalytischen Auswertung von deutschsprachigen Publikationen zu belegen und fand hierbei relativ hohe Zusammenhänge zwischen Status und Beurteilung der Eignung studentischer Bewertungen. Aber auch wenn hierdurch die Akzeptanzunterschiede erklärbar sind, bietet diese Erklärung kaum Ansatzpunkte für Maßnahmen zur Verbesserung der Akzeptanz.

Solche Ansatzpunkte ließen sich jedoch eventuell aus der Einbeziehung der Arbeitszufriedenheit in diese Diskussion ableiten. Wenn man – wie am Beginn dieses Kapitels dargelegt – davon ausgeht, daß langfristig nur solche Reformmaßnahmen (höhere) Akzeptanz finden, die die Arbeitszufriedenheit der Lehrenden nicht gefährden, sondern möglichst erhöhen und als Ziel von Leistungsbewertungen auch eine höhere Leistungsgerechtigkeit und Autonomie der Mitarbeiter und damit einhergehend auch eine Erhöhung der Arbeitszufriedenheit formuliert wurde, dann kann neben der akzeptanzfördernden Wirkung einer hohen Arbeitszufriedenheit auch eine Erhöhung der Arbeitszufriedenheit selbst als Bedingung erfolgreicher Institutionalisierung gesehen werden.

Als Fazit einer Analyse der Determinanten der Arbeitszufriedenheit wurde festgestellt, daß v.a. die berufliche Anerkennung Einfluß auf die Arbeitszufriedenheit hat. Dies ist bei Professoren noch stärker ausgeprägt als bei wissenschaftlichen Mitarbeitern. Die Höhe des Gehalts und auch die empfundene Leistungsgerechtigkeit des Gehalts ist dagegen von

nachrangiger Bedeutung für die Arbeitszufriedenheit. Von höherer Bedeutung für die Arbeitszufriedenheit sind für die Lehrenden insgesamt lediglich noch die Autonomie (Möglichkeiten, eigene Ideen zu verwirklichen) sowie die Zufriedenheit mit den Lehrveranstaltungen. Für die wissenschaftlichen Mitarbeiter ist auch die Arbeitsplatzsicherheit/ Befristung von höherer Bedeutung.

Daraus ergibt sich als Schlußfolgerung, daß die von der Bundesregierung beschlossene und derzeit umzusetzende Reform des Dienstrechts und dabei insbesondere die Einführung leistungsorientierter Besoldungszulagen für Professoren als Leistungsanreize (zumindest für die befragten Lehrenden) zu kurz greifen und nicht die angestrebten Effekte erzielen könnte, wenn sie nicht in starkem Maße auch den Aspekt der beruflichen Anerkennung berücksichtigt. Da lediglich die Lehrenden dreier Hochschulen befragt wurden, sind die Ergebnisse jedoch noch nicht ohne weiteres generalisierbar.

Zwei mögliche Ansatzpunkte ergeben sich aber bereits jetzt aus den vorliegenden Resultaten: Erstens wäre es möglicherweise effektiver, stärker direkt bei der beruflichen Anerkennung als bei den Besoldungen (Höhe und Leistungsgerechtigkeit des Gehaltes) anzusetzen. Bisher wird diese v.a. über die Forschungsreputation erlangt. Analog Forschungspreisen könnten als ideelle Leistungsanreize in größerem Umfang z.B. Lehrpreise vergeben werden. Diese dürften aber keine „Preise 2. Klasse“ sein, sondern sollten eine Dotierung in ähnlicher Höhe erhalten wie renommierte Forschungspreise. Parallel zu solch einer in gewisser Weise eher symbolischen Aufwertung sollte aber auch über Mechanismen nachgedacht werden, die für die breite Masse der Lehrenden eine stärkere Anerkennung der Lehrleistungen bewirken. Da Lehrveranstaltungen für die Arbeitszufriedenheit ebenfalls wichtig und Lehrorientierte mit den die Zufriedenheit stark beeinflussenden Faktoren besonders unzufrieden sind, könnten so Unzufriedene besser motiviert und der Stellenwert der Lehre insgesamt gefördert werden. Dies führt zum zweiten Ansatzpunkt.

Zweitens hängt die für Hochschullehrer wie wissenschaftliche Mitarbeiter wichtige berufliche Autonomie mit der Ausstattung bzw. Mittelzuweisung an die Fächer bzw. Fakultäten zusammen. In den letzten Jahren verstärkt eingeführte leistungsorientierte Mittelzuweisungen könnten die berufliche Zufriedenheit daher sowohl in die eine als auch in die andere Richtung stärker beeinflussen als Besoldungszulagen. Sie könnten deshalb (den hochschulpolitischen Zielen entsprechend) genutzt werden und dabei möglicherweise effektiver sein als Besoldungszulagen für Professoren. Um der Lehre (und insbesondere

ihrer Qualität) einen höheren Stellenwert zu verschaffen, wurde von mehreren Autoren die Einbeziehung subjektiver Bewertungen in Leistungsanreizmodelle diskutiert. Da dies derzeit eher abgelehnt wird, aber die Einbeziehung von Ergebnissen aus Absolventenbefragungen bei den Betroffenen Akzeptanz finden könnte, wäre es eine Möglichkeit, den Qualitätsaspekt stärker als bisher in Leistungsanreizmodelle zu integrieren. Auf diese Weise könnte die Lehre und ihre Qualität aufgewertet werden und Leistungen in diesem Bereich breitere Anerkennung finden. Wenn daneben bei der Verwendung der Anteile leistungsorientierter Mittelverteilung, die zur Honorierung hoher Lehrqualität ausgeschüttet werden, ein hoher Handlungsspielraum der Verantwortlichen für die Lehre gesichert wird, kann dies auch zu einem höheren Autonomieerleben der Lehrenden führen. Breitere Anerkennung und ein höheres Autonomieerleben könnten zu einer höheren Arbeitszufriedenheit insgesamt führen. Empirische Studien in der Privatwirtschaft zeigten, daß Selbständigkeit und Autonomie am Arbeitsplatz sowie berufliche Anerkennung neben Qualifizierungsmaßnahmen Motivatoren sind, die zu einer verbesserten Leistung der Betroffenen führen können (vgl. z.B. Bartell-Ling 1996: 277).

Möglicherweise kann es so gelingen, Leistungsanreizsysteme erfolgreich zu institutionalisieren, die sich in möglichst vielen Aspekten positiv und in möglichst wenigen negativ auf die Arbeitszufriedenheit und die Lehrqualität auswirken.

## 7. Zusammenfassung

Diese Arbeit verfolgte das Ziel, sich mit Leistungsbewertungen und Leistungsanreizen für die Lehre an deutschen Hochschulen auseinanderzusetzen. Wo dies möglich erschien, sollten neue Akzente gesetzt werden und bisher unverbundene Diskussionsstränge zusammengeführt werden, um daraus ggf. Möglichkeiten der Weiterentwicklung abzuleiten. Insbesondere dort, wo in Deutschland noch keine oder nur wenige Erfahrungen vorliegen, erschien es sinnvoll, die anderer Länder einzubeziehen. Beispielsweise fällt bei einer solchen Betrachtungsweise auf, daß einige andere Länder mit hoher Hochschulautonomie und relativ weit entwickelten Leistungsanreizmodellen wie der leistungsorientierten Mittelvergabe zugleich auch der Qualitätssicherung und dem Qualitätsmanagement bzw. der Qualitätsorientierung allgemein einen deutlich größeren Stellenwert einräumen. Teilweise enthalten Qualitätsmanagementansätze auch Leistungs- bzw. Qualitätsanreize. Eine Möglichkeit der Weiterentwicklung der für Deutschland vorherrschenden Diskussion könnte daher in der Zusammenführung bislang überwiegend getrennt geführten Debatten zur Qualität und zu Leistungsanreizen gesehen werden. Hierfür wäre jedoch zunächst zu klären, inwieweit eine solche Sichtweise mit den in Deutschland vorherrschenden bzw. vorgeschlagenen Konzepten, Modellen und Instrumenten kompatibel ist. Weiterhin wäre zu klären, welche der verschiedenen Kriterien bzw. Indikatoren für welche Zwecke geeignet wären, da jede Leistungsbewertung und jedes Anreizsystem entscheidend von den verwendeten Kriterien abhängt. Und schließlich wäre zu diskutieren, unter welchen Bedingungen eine erfolgreiche Institutionalisierung aussichtsreich erscheint und unter welchen eher nicht. Daher wurde die Arbeit in drei Schwerpunkte gegliedert:

- 1.) Vorstellung und Einordnung von Konzepten, Modellen und Instrumenten
- 2.) Analyse (potentieller) Leistungskriterien bzw. Indikatoren
- 3.) Diskussion der Bedingungen erfolgreicher Institutionalisierung

Bevor jedoch mit der Bearbeitung des ersten Schwerpunktes begonnen werden konnte, mussten zunächst einige Begrifflichkeiten geklärt werden. Dies erfolgte im zweiten Kapitel. Die Ergebnisse dieser Klärung sollen hier kurz zusammengefaßt werden.

Zunächst waren Begriffsklärungen zu Leistung und Qualität und ihrer Bewertung notwendig, da die bisherigen Diskussionen hierzu auf keiner einheitlichen begrifflichen Grundlage stattfanden. Im Ergebnis dieser Klärung soll Leistung in dieser Arbeit gelten als

„die (von den Hochschulen bzw. ihrem Personal) geleistete Arbeit, gemessen an ihrer Menge (Leistungseinheiten, Output) und Qualität pro Zeiteinheit“. Wenn im Leistungsbegriff explizit Elemente der Bewertung enthalten sind, sollte von *Leistungsbewertung* die Rede sein. Da in dieser Definition die Qualität eine wesentliche Rolle spielt und auch dieser Begriff sehr uneinheitlich verwendet wird, sollte er geklärt werden. Anhand einer Diskussion einschlägiger Publikationen wurde herausgearbeitet, daß für eine Eingrenzung des Qualitätsbegriffs v. a. die Fragen „Qualität von was?“ und „Qualität für wen?“ zu beantworten sind. Die erste Frage bezieht sich auf den Qualitätsgegenstand, die zweite auf die Qualitätszielgruppe.

Der Gegenstand der Qualität von Hochschullehre ist hierbei für die verschiedenen Ebenen des Hochschulsystems differenziert zu betrachten. In dieser Arbeit wurde versucht, Lehrqualität für die Ebene der gesamten Hochschule (Hochschulebene), die Studiengangsebene und die Lehrveranstaltungsebene zu definieren. Während für die Hochschulebene nach wie vor Schwierigkeiten zu konstatieren sind, eine Definition zu finden, die auch operationalisierbar ist, kann für die Studiengangsebene die Definition des Wissenschaftsrates (1996) als Ausgangsbasis genutzt werden. Für die Lehrveranstaltungsebene existiert in der Literatur eine Vielzahl von Definitionen. In dieser Arbeit wird eine Definition favorisiert, in die nicht nur die (didaktische) Gestaltung von Lehrveranstaltungen eingeht, sondern die um die Lehr-, Lern- und Studienbedingungen erweitert wurde.

Zielgruppen für die Qualität der Hochschullehre können ebenfalls sehr vielfältig sein. An erster Stelle sind selbstverständlich die Studierenden zu nennen, die allerdings keine homogenen Interessen haben müssen. Vielmehr können sie je nach Motivationslage und/oder Vorkenntnissen in unterschiedliche Zielgruppen aufgeteilt werden. Eine ebenso wichtige Zielgruppe sind aber auch die Lehrenden selbst, für die durch eine bessere Lehrqualität z.B. eine höhere Motivation und Arbeitszufriedenheit zu erlangen möglich scheint, was aber noch empirisch zu untersuchen wäre. Weitere Zielgruppen sind (auf Studiengangsebene betrachtet) Unternehmen, die potentiell die Absolventen eines Studienganges einstellen oder auch Abiturienten, die ein Studienprogramm nachfragen (sollen). Aber auch Ministerien und Parlament können wichtige Zielgruppen sein. Aus der Vielfalt der Zielgruppen können konfligierende Interessen resultieren, die in einem mehr oder weniger politischen Aushandlungsprozeß geklärt oder gewichtet werden müssen, da in einem solchen Fall nicht alle Ziele gleichermaßen verfolgt werden können.

## **7.1 Vorstellung und Einordnung von Konzepten, Modellen und Instrumenten**

Die Vorstellung und Einordnung von Konzepten, Modellen und Instrumenten für Leistungsbewertung und Leistungsanreize erfolgte im dritten Kapitel. Hierbei wurde Wert darauf gelegt, nicht möglichst viele verschiedene Konzepte, Modelle und Instrumente vorzustellen, sondern sich auf die für das Ziel dieser Arbeit als wesentlich erscheinenden zu konzentrieren. Zur Systematisierung erfolgte eine Einteilung in:

- Instrumente zur Leistungs- und Qualitätsbewertung,
- Modelle des Qualitätsmanagements,
- Leistungsanreizmodelle.

Die einzelnen Instrumente und Modelle zu den ersten beiden Gliederungspunkten wurden aufgrund der hierzu bereits vorliegenden Publikationsfülle nicht allumfassend diskutiert, da ihre Erörterung v. a. der Einordnung in die Qualitätsdebatte dienen sollte. Die sich erst seit einigen Jahren in einer breiteren hochschulpolitischen Diskussion befindenden Leistungsanreizmodelle wurden dagegen ausführlicher thematisiert. Hierfür wurden an Erfahrungsberichten sowohl eher optimistisch ausfallende Veröffentlichungen der hochschulpolitischen Protagonisten solcher Leistungsanreizmodelle wie des Centrums für Hochschulentwicklung (CHE) herangezogen, als auch eher skeptische Publikationen. Aus dem Hochschulbereich liegen jedoch bisher kaum Erfahrungsberichte mit kritischem Grundtenor vor, da überwiegend nur die einführenden Institutionen selbst über ihre Erfahrungen berichten. Diese „Vorreiter“ neigen aus naheliegenden Gründen eher zu optimistischen Einschätzungen. Daher wurden für eine ausführlichere Diskussion auch arbeitswissenschaftliche und gewerkschaftliche Publikationen zu Erfahrungen mit der Einführung von Leistungsanreizen im öffentlichen Dienst allgemein herangezogen. Dabei zeigte sich z.T. eine große Ähnlichkeit in der Diskussion von Problemen und Möglichkeiten.

Zu Möglichkeiten der Weiterentwicklung und Ergänzung von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen kann insgesamt zunächst festgestellt werden, daß relativ häufig der Bedarf einer Weiterentwicklung von Anreizmodellen in Richtung qualitativer Kriterien bzw. subjektiver Indikatoren z.B. im Rahmen von Zielvereinbarungen formuliert wird. Es gibt hierzu jedoch bislang relativ wenige konzeptionelle und empirische Vorarbeiten zur konkreten Operationalisierung solcher Indikatoren.

Eine andere Möglichkeit der Weiterentwicklung und Ergänzung wird in der Schaffung einer stärkeren Verbindlichkeit von Maßnahmen als „Follow-up“ bzw. zur Umsetzung der Ergebnisse von Evaluationen und Qualitätsbewertungen gesehen, die z.B. über Zielvereinbarungen erfolgen könnte.

Nach Einschätzung des Verfassers liegt in der Verbindung dieser beiden bislang meist unverbunden betrachteten Weiterentwicklungsmöglichkeiten ein Potential, die jeweils als Schwächen angesehenen Aspekte von Leistungsanreizmodellen sowie Leistungs- und Qualitätsbewertungen zu mindern. So könnten die Anreizmodelle durch die Nutzung von Erfahrungen aus Qualitätsbewertungen durch Evaluationen u.U. stärker auch Qualität erfassen; die Umsetzung der Ergebnisse von Evaluationen könnte umgekehrt durch Verknüpfung mit Anreizen verbindlicher gestaltet werden. Hierzu gibt es international bereits über längere Zeiträume Erfahrungen, die in die Diskussion einbezogen werden sollten.

Die aus anderen Ländern vorhandenen Vorarbeiten können aber nicht ohne weiteres übernommen werden und so wäre zunächst auszuloten, welche Bedingungen ggf. erfüllt sein müßten, um eine solche Verbindung bislang unverbundener Ansätze an *deutschen* Hochschulen möglich und sinnvoll werden zu lassen. Da ein solches Wissen nicht bei allen Mitgliedern der Hochschule, sondern vor allem bei einigen Experten zu finden ist, bietet sich hier die Durchführung von Experteninterviews an. Hierbei wurde der Schwerpunkt v.a. auf Anreizmodelle und Kriterien bzw. Indikatoren gelegt. Die Experteninterviews bilden ein separates viertes Kapitel; sie können aufgrund der gewählten Schwerpunktsetzung als Bindeglied zwischen der Diskussion von Konzepten, Modellen und Instrumenten und der Analyse von Kriterien angesehen werden.

Bevor einige zentrale Ergebnisse der Experteninterviews vorgestellt werden, wird noch kurz darauf eingegangen, welche Experten einbezogen wurden und nach welchen Grundprinzipien die Interviews ausgewertet wurden. Bei der Auswahl der Gesprächspartner für die Experteninterviews wurde in zwei Schritten vorgegangen: Zunächst waren geeignete Hochschulen und dann geeignete Gesprächspartner (an den Hochschulen) auszuwählen. Bei der Auswahl der Hochschulen für die Experteninterviews wurde davon ausgegangen, daß es bezogen auf das Arbeitsklima und die Leitungskultur zwischen großen und kleinen Universitäten einerseits und Fachhochschulen andererseits deutliche Unterschiede geben kann und deshalb in Anlehnung an das Prinzip des „theoretical sampling“ möglichst jeder dieser Hochschultypen vertreten sein sollte. Bei der Auswahl der einzubeziehenden Personen für die Experteninterviews war Grundprinzip, daß insbesondere Vertreter der

Hochschulleitungen bzw. mit der Einführung von Leistungsbewertungssystemen befaßte Personen der Hochschulen und Personalräte als Vertreter der Betroffenen zu Wort kommen sollten. Als Gesprächspartner wurden außerdem Vertreter der Gruppe der externen, nicht an den betreffenden Hochschulen beschäftigten, aber dennoch als Berater und über wissenschaftliche Publikationen auf diese einwirkende Personen aus der Hochschulforschung einbezogen. Die Auswertung der Experteninterviews erfolgte in Anlehnung an Meuser und Nagel (1994, 2003), wobei in dieser Arbeit die forschungslogisch motivierte Funktion der Experteninterviews die Erlangung von Kontextwissen war und die Experteninterviews nur ein Erhebungsinstrument neben anderen darstellten. Diesem Ziel entsprechend wurden die Interviews nicht vollständig, sondern nur partiell bzw. eher themenzentriert ausgewertet. Als Kernaussagen, die sich aus der Auswertung der Experteninterviews ableiten lassen, können zusammenfassend genannt werden:

- Eine stärkere Leistungsorientierung kann durchaus Akzeptanz finden, wenn der Gedanke der Fairneß und Leistungsgerechtigkeit ausreichend Berücksichtigung findet.
- Die Einheit von Forschung und Lehre könnte weniger starr gesehen werden.
- Wenn Kriterien quantitative *und* qualitative Aspekte enthielten, wären unerwünschte Nebenwirkungen zu vermindern.
- Als qualitative Kriterien könnten auch subjektive Einschätzungen von Studenten und Absolventen geeignet sein.

Zu diesen Ergebnissen der Experteninterviews ist hinzuzufügen, daß sie aufgrund der geringen Fallzahlen, aber auch aufgrund des gewählten Untersuchungsansatzes nicht verallgemeinert werden können. Dies war auch nicht das Ziel. Die Experteninterviews geben eher die Perspektive von „Akteuren“ wieder; die Sicht von „Betroffenen“ erfassen die Lehrendenbefragungen und deren Ergebnisse sollen auch (zumindest für die betreffenden Hochschulen) verallgemeinert werden können. Die Experteninterviews dienten v.a. der Eruierung, ob bestimmte Weiterentwicklungen grundsätzlich als sinnvoll angesehen werden. Ein Beispiel hierfür wäre, daß qualitative Kriterien (auch in Form subjektiver Bewertungen) aus Perspektive der Experten grundsätzlich als sinnvolle Ergänzung der bisher überwiegend verwendeten quantitativen bzw. „objektiven“ Kriterien angesehen werden, um unerwünschte Nebenwirkungen zu vermeiden.

## **7.2 Analyse von (potentiellen) Kriterien und Indikatoren**

Die Bearbeitung des zweiten großen Schwerpunktes dieser Arbeit, die Analyse von (potentiellen) Kriterien und Indikatoren für Leistungsbewertungen und Leistungsanreize in der Hochschullehre, erfolgte im fünften Kapitel. Hier sollten sowohl „objektive“, als auch „subjektive“ Kriterien anhand der vorhandenen Forschungsliteratur genauer auf ihre Eignung untersucht werden. Wo dies notwendig und sinnvoll erschien, sollte dies auch mit einer kurzen Diskussion der Ergebnisse eigener empirischer Untersuchungen ergänzt werden. Als Maß für die Eignung bzw. als Anforderungen an eine Verwendung „objektiver“ und „subjektiver“ Kriterien in Indikatoren wurde hierzu in der vorliegenden Literatur formuliert: Ausreichende Datenqualität, Praktikabilität und dabei insbesondere Vergleichbarkeit, hohe Zeitnähe, sinnvolle Anreizwirkung, Vermeidung unerwünschter Anpassungseffekte und ein vertretbares Verhältnis von Aufwand und Nutzen. Da in den bislang vorliegenden Veröffentlichungen die Eignung von Indikatoren für die genannten Anforderungen häufig nur relativ kurz angesprochen wurde und subjektive Bewertungen bezüglich ihrer Eignung zur Verwendung in Indikatoren so gut wie überhaupt nicht diskutiert wurden, nimmt dieser Aspekt in der Zusammenfassung dieser Arbeit etwas breiteren Raum ein.

### **Objektive Indikatoren**

Zur Datenqualität und Aussagekraft „objektiver“ Kriterien wurde anhand einer Diskussion von häufig verwendeten Kennziffern bzw. Indikatoren wie Studenten- und Studienanfängerzahlen, Betreuungsrelation, Studiendauer und Durchschnittsnote, Erfolgsquote und Absolventenquote herausgearbeitet, daß mit solchen Indikatoren jeweils für sich genommen "nur ein sehr unvollständiges Bild der Studiensituation" zu zeichnen ist (vgl. auch bereits Hornbostel & Daniel 1996). Das liegt zum einen Teil daran, daß bisher den offiziellen Stellen vorliegende hochschulstatistische Daten in ihrer Qualität oft nicht befriedigen können, wie z.B. zur Abbrecherquote. Meist ist lediglich die Berechnung einer Schwundquote möglich. Dabei kann die Zuverlässigkeit zwar durch Mittelung über mehrere Jahre erhöht werden, wie auch eigene Auswertungen anhand von Daten sächsischer Hochschulen zeigten. Aussagen über das Studienabbruchverhalten sind mit Schwundquoten jedoch nicht möglich. Eine Erhöhung der Aussagekraft durch Weiterverarbeitung vorhandener und durch Ergänzung mit weiteren Daten, etwa zum Hochschul- und Fachrichtungswechsel ist z.B. bei

der Absolventenquote zwar grundsätzlich möglich. Dies ist allerdings auch entsprechend aufwendig. Selbst dann erlaubt es jedoch keine exakte, sondern nur eine grobe Aussage über die Wahrscheinlichkeit eines Abschlusses in einem bestimmten Studiengang einer Hochschule und sagt noch nichts über die *Qualität* eines solchen Abschlusses aus und darüber, unter welchen *Bedingungen* er erworben wurde. Ersteres wäre grundsätzlich nur mit relativ aufwendigen Absolventenverbleibsstudien möglich, letzteres wäre auch mit ohnehin durchzuführenden Studentenforschungen zur Qualität der Lehre erfaßbar. Andere, mit relativ geringem Aufwand verfügbare Indikatoren wie die Betreuungsrelation können selbst bei klarer Definition der Aussagekraft nur potentielle Möglichkeiten für eine gute Betreuungssituation aufzeigen, nicht aber die tatsächlich realisierte Betreuung. Diese wiederum könnte möglicherweise durch eine subjektive Einschätzung der Studierenden (oder auch retrospektiv von Absolventen) ergänzt werden.

Die Relation Bewerber pro Studienplatz ist aus inhaltlichen und methodischen Gründen als Indikator für die Attraktivität eines Studienganges ungeeignet, obwohl sie ebenfalls meist leicht zu berechnen wäre. Diese Gründe können auch nicht durch Hinzuziehen anderer Indikatoren ausgeglichen werden. Nach einschlägigen Studien bildet die Relation Bewerber pro Studienplatz eher die Anzahl verfügbarer Studienberechtigter in der Umgebung eines Studienortes bzw. die Attraktivität des Studienortes ab, nicht aber die eines bestimmten Studienganges einer Hochschule. Da hierbei Lebensbedingungen wie niedrige Mieten, kulturelles Angebot usw. eine größere Rolle spielen als die Studienqualität, ist sie kein inhaltlich valides Kriterium für die hier verfolgten Zwecke. Weitere Gründe können darin liegen, daß die Datengrundlage z.T. uneinheitlich ist.

Bei der Fachstudiendauer als einem potentiellen Indikator für die Ergebnisqualität des Studiums sind ebenfalls Probleme der Vergleichbarkeit z.B. aufgrund unterschiedlicher Definitionen des Abschlußzeitpunktes oder der Mittelwertberechnung zu konstatieren. Zudem können Ursachen für Schwankungen auch außerhalb der Verantwortung der Hochschulen liegen, so daß eine Verwendung als alleinige oder zentrale Leistungskennziffer fragwürdig scheint. Zur genaueren Klärung von Ursachen für Schwankungen könnten ggf. auch hier Absolventenforschungen beitragen. Dennoch kennzeichnet die Studiendauer als wesentlicher Nachteil gegenüber der Studierbarkeit (neben der meist geringeren zugrunde liegenden Fallzahl) die mangelnde Zeitnähe. Wenn es Probleme oder auch Verbesserungen bezüglich der Möglichkeit gibt, das Studium in der vorgesehenen Zeit abzuschließen, zeigt sich dies erst, wenn Studierende als Absolventen das Studium beenden. Verlassen

Studierende erst kurz vor der Zwangsexmatrikulation, z.B. im 14. Semester, noch den Studiengang, weil es in einem anderen Bundesland keine solche Regelung gibt, zeigt sich dies in der Studiendauer überhaupt nicht. Die Studierbarkeit kann daher nach o.g. Anforderungen insgesamt als besser geeigneter Indikator angesehen werden als die Fachstudiendauer. Auch die Noten, nach wie vor ein Einstellungskriterium für Arbeitgeber, sind nur unter bestimmten Bedingungen valide und zuverlässig, die in Deutschland meist nicht erfüllt sind. Zudem korrelieren Noten zumindest an einzelnen Hochschulen kaum oder nicht in der gewünschten Richtung mit dem beruflichen Erfolg. Noten können also am ehesten noch ein Maß für die Notenvergabepraxis im jeweiligen Studienfach sein, jedoch keinesfalls ein absolut zu sehendes zuverlässiges Maß für die Leistungsfähigkeit der Absolventen oder des Studienfaches. Dasselbe gilt in ähnlicher Weise auch für die Erfolgsquoten bei den Abschlußprüfungen. Diese sind als Ergänzung zu den Noten wichtig, da in die Durchschnittsnoten der Absolventen nur die bestandenen Prüfungen einfließen. Die Erfolgsquoten wiederum sollten – insbesondere in Fächern mit hohen Erfolgsquoten und dennoch geringen Absolventenzahlen – durch Absolventenquoten ergänzt werden. Umgekehrt ermöglichen die Erfolgsquoten ergänzend zu den Absolventenquoten Aussagen, welcher Anteil des „Schwundes“ auf endgültiges Nichtbestehen von Abschlußprüfungen zurückzuführen ist.

Vergleiche und Interpretationen der sogenannten „objektiven“ Kennziffern und Indikatoren bedürfen also neben einer sorgfältigen und sachkundigen Analyse der gegenseitigen Ergänzung. Außerdem erscheint eine Erweiterung durch „subjektive“ Bewertungen von Absolventen, Studierenden oder auch Lehrenden grundsätzlich sinnvoll. Für eine ausgewogene Erfassung der Studienqualität im Sinne einer Leistungsbewertung und mit dem Ziel, Anreize für Leistung *und* Qualität zu setzen, wäre die Einbeziehung subjektiver Bewertungen darüber hinaus eine wichtige Voraussetzung, da die bisher überwiegend diskutierten „objektiven“ Kennziffern und Indikatoren Aspekte der Qualität nicht erfassen können. Ohne die Einbeziehung von Qualitätsaspekten könnte es leicht zu einer Fehlsteuerung in Richtung höherer „Stückzahlen“ von Absolventen kommen, die nicht in der gewünschten und erforderlichen Qualität ausgebildet werden.

Hierbei dürften Erfahrungen aus anderen Ländern, wie sie bereits im dritten Kapitel im Zusammenhang mit der leistungsorientierten Mittelverteilung angesprochen wurden, teilweise hilfreich sein, da eventuelle Fehlanreizwirkungen von vornherein bedacht werden könnten. Eine direkte Übertragbarkeit von bereits langjährig eingesetzten Leistungsbewertungssystemen und Indikatorenmodellen wie in Großbritannien oder Australien ist

jedoch aufgrund der anderen Struktur und ggf. andersartigen Wirkungsweisen nicht gegeben. So würde beispielsweise die Verwendung von stark gewichteten Studierendenzahlen als Indikator für Leistungsanreize ohne entsprechendes Korrektiv vermutlich zu unerwünschten Nebenwirkungen wie der Absenkung der Qualität führen. In Ländern, die dies bereits anwenden, kann das z.B. durch ein (im Vergleich zu Deutschland) stärkeres Qualitätsmanagement und eine hohe Transparenz über die Qualität von Studium und Lehre an den Hochschulen vermieden oder jedenfalls vermindert werden. In Deutschland werden Qualitätsmanagementmodelle bislang nur an wenigen Hochschulen bzw. Fachbereichen/ Fakultäten eingesetzt.

Als Fazit bleibt festzuhalten, daß von den diskutierten „objektiven“ Kriterien und Indikatoren jeweils für sich genommen aus verschiedenen Gründen keines allein als wirklich objektiv gelten kann. Als am meisten problematisch erscheinen für das Ziel, Leistung einschließlich Qualität zu fördern, die Verwendung von Studenten-, Studienanfänger- und Bewerberzahlen, aber auch die von durchschnittlichen Abschlußnoten. Auch die Betreuungsrelation kann nur das Potential für eine gute Betreuung aufzeigen, nicht aber die tatsächlich realisierte. Ebenfalls nicht unproblematisch, aber etwas besser geeignet erscheinen (im Vergleich zur häufig diskutierten durchschnittlichen Fachstudiendauer) die Studierbarkeit sowie die Absolventenquote, wobei diese unter Einbeziehung zusätzlicher Informationen weiterentwickelt werden sollte. Insgesamt ermöglicht erst ein „Datenkranz“ aus mehreren, sich gegenseitig ergänzenden Kennziffern und Indikatoren unter Einbeziehung subjektiver Indikatoren ein möglichst realistisches Bild der Studiensituation eines Faches an einer bestimmten Hochschule.

### **Subjektive Indikatoren**

Als subjektive Indikatoren ergänzt werden könnten beispielsweise Einschätzungen von Studierenden und/ oder Absolventen zur Betreuungssituation, aber auch Einschätzungen, ob es aus Sicht der Studierenden/ Absolventen in der Studienorganisation Gründe gab, die eine Einhaltung der Regelstudienzeit erschwerten (und ggf. um welchen Zeitraum). Daneben sollten aber zusätzlich zu solchen direkt aus der Ergänzung „objektiver“ Indikatoren resultierenden weitere subjektive Indikatoren herangezogen werden, um Qualität möglichst in allen wesentlichen Dimensionen zu erfassen. Als Voraussetzung für eine Einbeziehung

subjektiver Bewertungen in Indikatoren war jedoch deren Validität und Zuverlässigkeit zu diskutieren.

*Studentenbefragungen* sollten hierbei an erster Stelle thematisiert werden, da hierzu bereits eine Fülle von Literatur vorliegt. Die Ergebnisse empirischer Untersuchungen lassen sich folgendermaßen zusammenfassen, wobei Studentenbefragungen überwiegend als standardisierte schriftliche Befragungen untersucht wurden. Demnach müssen studentische Urteile zur Lehr- und Studiensituation als weitgehend valide und zuverlässig bzw. reliabel eingeschätzt werden. Zunächst wurde bei den Untersuchungen zur Validität festgestellt, daß die Urteile von Lehrenden, Fremdgutachtern und Studierenden bei Anlegen gleicher Kriterien und konkreter Fragen hierzu weit besser übereinstimmen, als häufig angenommen wird. Auch zur Zuverlässigkeit kann Ähnliches berichtet werden: Sowohl bei Meßwiederholung innerhalb eines Semesters, als auch über mehrere Semester hinweg sind die Urteile der Studierenden über das Lehrverhalten ihrer Dozenten weitgehend stabil. Gleichwohl sollten immer mehrere (mindestens fünf) Lehrveranstaltungen zur Beurteilung des Lehrverhaltens von Dozenten herangezogen werden. An einigen Punkten besteht noch weiterer Untersuchungsbedarf – z.B. zur Stabilität der Ergebnisse von Studentenbefragungen an deutschen Hochschulen, die über das Heidelberger Inventar (HILVE) hinausgehen. Will man weitergehende Aussagen zur Eignung von Studentenbefragungen als ein Qualitätskriterium in Leistungsvergleichen treffen, dann sollten aber noch andere Aspekte berücksichtigt werden. So wird auch das Nichtvorhandensein von „verzerrenden“ Bias-Variablen bzw. Einflußfaktoren zu den Eignungskriterien gezählt.

Als die häufigsten genannten Einflußfaktoren, welche die Vergleichbarkeit beeinträchtigen könnten, wurden Sachkenntnis und Reife, Geschlecht, Noten, Leistungsanforderungen, Interesse und Kursgröße genannt. Hierbei finden sich zu unterschiedlicher Sachkenntnis und Reife von Studierenden (gemessen an evtl. Hilfskrafttätigkeit, Studienintensität, Hochschulwechsel oder Alter) und zum Geschlecht kaum nennenswerte Zusammenhänge. Allerdings urteilten Studierende in höheren Fachsemestern tendenziell etwas kritischer. Auch der Einfluß der Noten bzw. Notenerwartung – der zur Vermeidung eines Zirkelschlusses (durch unterschiedliche Lerneffektivität innerhalb eines Kurses verursachte Leistungsunterschiede) nicht innerhalb einer Lehrveranstaltung, sondern über mehrere Lehrveranstaltungen hinweg untersucht werden muß – fiel in verschiedenen Studien eher gering aus oder war nicht nachweisbar.

Für die Leistungsanforderungen war im Einklang mit Erkenntnissen der Lernpsychologie festzustellen, daß keineswegs bei den niedrigsten Leistungsanforderungen die besten Bewertungen vergeben wurden. Vielmehr wurden diese bei als optimal empfundenen und etwas höheren Anforderungen erzielt. Die niedrigsten Leistungsanforderungen gingen dagegen mit den schlechtesten Bewertungen einher. Lediglich der Einfluß des Interesses der Studierenden entsprach den Vermutungen und zeigte sich damit im Verhältnis zu den anderen potentiellen Einflußfaktoren als stärkste Bias-Variable. Studierende mit höherem Fachinteresse beurteilen Lehrveranstaltungen besser als Studierende mit geringerem Interesse.

Zu weiteren genannten Anforderungen kann festgehalten werden, daß Studentebefragungen meist eine vergleichsweise hohe Zeitnähe haben, daß sie in Verbindung mit quantitativen bzw. objektiven Indikatoren eine sinnvolle Anreizwirkung entfalten und daher unerwünschte Anpassungseffekte vermeiden könnten. Da Studentebefragungen zur Lehrqualität ohnehin gesetzlich gefordert sind und meist auch durchgeführt werden, ist ein günstiges Verhältnis von Aufwand und Nutzen festzustellen.

*Absolventenbefragungen* können, trotz einiger aufgrund ihrer Retrospektivität vorzunehmenden Einschränkungen der Aussagekraft, zur Einschätzung der Studiensituation grundsätzlich als geeignet gelten. Während kurz nach Abschluß des Studiums durch den erstmals möglichen Gesamtüberblick bei geringen Retrospektivitätseffekten ihre Aussagekraft als am höchsten eingeschätzt werden muß, kann sie mit größer werdendem zeitlichem Abstand (nach mehreren Jahren oder gar Jahrzehnten) durch Erinnerungsunschärfen oder Verklärungen sinken. Vergleiche der Ergebnisse von Studenten- und Absolventenbefragungen an der TU Dresden sowie der Ergebnisse von Absolventenbefragungen kurz nach Studienabschluß und mit mindestens 5 Jahren Abstand zum Studienabschluß konnten aber lediglich die Tendenz zu einer etwas positiveren Einschätzung mit größer werdendem zeitlichen Abstand und keine grundsätzlich anderen Bewertungsmuster aufzeigen.

Fragenkataloge zum Übergang vom Studium in den Beruf, wie z.B. die bereits über einen längeren Zeitraum genutzten Kasseler Studien, werden als validiert eingeschätzt. Zur Vergleichbarkeit von Absolventenbefragungen kann formuliert werden, daß diese unter Berücksichtigung solcher Aspekte, die eine Hochschule nicht zu beeinflussen in der Lage ist, als gegeben angesehen werden kann. Durch eine Hochschule nicht beeinflusst wird insbesondere die Ausgangssituation der Absolventen schon am Beginn ihres Studiums. Dies bezieht sich v.a. auf soziale Herkunft und Geschlecht, die über das gesamte Studium hinweg

bis zum beruflichen Verbleib wirken und Einfluß auf den beruflichen Erfolg haben können, wie bivariate und multivariate Analysen anhand mehrerer Absolventenstudien zeigten. Aber auch die aktuelle persönliche Situation der Absolventen zum Befragungszeitpunkt, wie z.B. das Vorhandensein von Kindern oder der Partnerstatus, sollten grundsätzlich berücksichtigt werden, da sie ebenfalls Einfluß auf den beruflichen Erfolg haben können. Weiteren Anforderungen wie der Zeitnähe können Absolventenbefragungen nur bezüglich der Angaben zum beruflichen Verbleib und Erfolg genügen, da sich retrospektive Bewertungen des Studiums auf eine bereits länger zurückliegende Situation beziehen. Der Anforderung, sinnvolle Anreize zu schaffen, ohne unerwünschte Anpassungseffekte zu erzielen, kann durch Absolventenbefragungen sehr gut entsprochen werden, da sie häufig als einzige Möglichkeit gelten, nicht nur die Quantität bzw. „Stückzahl“ der Absolventen zu erfassen, sondern auch deren Qualität. Absolventenbefragungen erfordern allerdings einen vergleichsweise hohen Aufwand, der erst mittelfristig durch kostensparende Weiterentwicklungsmöglichkeiten in Richtung Online-Befragungen gesenkt werden könnte. Dem Aufwand stehen als Ertrag jedoch neben der alternativlosen Informationsgewinnung zum beruflichen Verbleib und der (derzeit leider nicht finanziell bezifferbaren) Möglichkeit zur Vermeidung von Fehlanreizwirkungen weitere Nutzungsmöglichkeiten der aus Absolventenbefragungen resultierenden aktuellen Adreßpools und damit auch ein vergleichsweise hoher Ertrag gegenüber.

Zur Eignung von *Lehrendenbefragungen* kann insgesamt festgehalten werden: Lehrende schätzen bezüglich der Situation von Lehre und Studium nach den vorliegenden Ergebnissen offener Befragungen und halboffener Fragestellungen zwar weitgehend dieselben Aspekte als relevant ein, wie sie in Studierendenbefragungen (und z.T. in Absolventenbefragungen) verwendet werden. Die Ergebnisse standardisierter Befragungen mit geschlossenen Fragestellungen zeigen jedoch, daß Lehrende die Lehre insbesondere bei Aspekten, die Selbsteinschätzungen beinhalten, anders bewerten als Studierende und unabhängige Fremdgutachter. Bei anderen Aspekten, z.B. zu Rahmenbedingungen der Lehre, stimmen die Bewertungen der verschiedenen Befragtengruppen besser überein. Da bei Selbsteinschätzungen generell von eingeschränkter Validität gesprochen wird und die vorliegenden Ergebnisse dies bestätigen, ist auch bei Lehrendenbefragungen bezüglich dieser Aspekte nur von eingeschränkter Validität auszugehen. Ähnliche Ergebnisse zeigen Analysen zur Dimensionalität der Befragungsergebnisse. Bei Untersuchungen unter Einbeziehung der Einschätzungen zur Situation von Lehre und Studium konnten mittels Faktorenanalysen

keine stabilen Faktorenstrukturen gefunden werden. Anders sieht dies bei separater Analyse von Aspekten der Arbeitszufriedenheit aus, die für Leistungsbewertungen und Leistungsanreize als relevant eingeschätzt wurden (u.a: Höhe und Leistungsgerechtigkeit des Gehalts, berufliche Anerkennung). Die hierzu existierenden Modelle konnten unter Berücksichtigung der hochschulspezifischen Besonderheiten weitgehend bestätigt werden und zeigten sich auch bei Untersuchungen über mehrere Hochschulen hinweg als relativ stabil.

Studien zur Zuverlässigkeit, die hier in Ermangelung besserer Analysemöglichkeiten nur exemplarisch zu in etwa vergleichbaren Fragestellungen bei zwei unabhängig voneinander durchgeführten Lehrendenbefragungen an derselben Hochschule durchgeführt werden konnten, ergaben in der Grundtendenz ähnliche Ergebnisse und damit zumindest keine Anhaltspunkte für eine mangelnde Retestreliabilität. Auch eine andere Möglichkeit zur Überprüfung der Reliabilität nach dem Grundprinzip der Testhalbierungsmethode zeigte, daß die Urteilerübereinstimmung über mehrere Personen hinweg gesehen (Interraterreliabilität) ab einer Mindestzahl von 10 befragten Lehrenden in etwa ähnlicher Höhe lag wie bei Studierendenbefragungen, wo sie in dieser Höhe als akzeptabel eingeschätzt wurde.

Zur Vergleichbarkeit und dabei evtl. zu beachtenden „verzerrenden“ Einflußfaktoren wurden in vorliegenden Publikationen bisher kaum Einschränkungen formuliert. Wie generell bei Vergleichen ist jedoch die Kontextgebundenheit der Ergebnisse und bei hochschulübergreifenden Vergleichen der Arbeitszufriedenheit von Lehrenden z.B. die Fächerstruktur der Hochschulen zu beachten, da diese Einfluß auf die Arbeitszufriedenheit haben könnten.

Weitere Anforderungen wie die Zeitnähe der Befragungsergebnisse sind derzeit kaum erfüllt, da Lehrendenbefragungen meist nur sporadisch durchgeführt werden. Der Erhebungsaufwand bei Lehrendenbefragungen hält sich jedoch in Grenzen, so daß grundsätzlich eine größere Zeitnähe möglich wäre. Wenn eine Hochschule sich für eine breite Einbeziehung aller Lehrenden in einen Prozeß zur Förderung der Qualitätsorientierung entschließen sollte, könnten Lehrendenbefragungen ähnlich wie Mitarbeiterbefragungen in Unternehmen aber u.a. auch als Kommunikationsinstrument und als Instrument zur Identifikation von Ansatzpunkten für akzeptanzerhöhende Maßnahmen für Reformvorhaben dienen, so daß dem Erhebungsaufwand ein mehrfacher Nutzen gegenüberstände. Da unerwünschte Anpassungseffekte durch Lehrendenbefragungen nicht zu erwarten sind und ihre Ergebnisse eher zur Vermeidung oder Minderung von Fehlanreizwirkungen dienen als selbst Fehlanreizwirkungen auslösen könnten, erscheinen sie grundsätzlich als geeignetes

Instrument zur Förderung der Qualitätsorientierung von Hochschulen. Hierbei ist die Datenqualität der Bewertung von Aspekten der Arbeitszufriedenheit den bisher vorliegenden Untersuchungen zufolge aufgrund der höheren Validität besser als bei Einschätzungen der Situation von Lehre und Studium. Die Untersuchungsergebnisse stellen jedoch z.T. nur erste Analysen dar und können noch nicht verallgemeinert werden.

Als vergleichendes Fazit zur Eignung subjektiver Bewertungen für die Ergänzung „objektiver“ Indikatoren kann formuliert werden, daß sie trotz z.T. vorhandener Retrospektivitäts- und Selbsteinschätzungseffekte gute Validität und Zuverlässigkeit erreichen können. Sie wären damit grundsätzlich für die Verwendung in Indikatoren geeignet. Studentenbefragungen eignen sich v.a. zur zeitnahen Einschätzung der Studiensituation, Absolventenbefragungen zur Analyse von Wirkungen der Hochschulausbildung auf den beruflichen Verbleib und Lehrendenbefragungen zur Einschätzung von Aspekten der Arbeitszufriedenheit. Inwieweit sie tatsächlich Anwendung finden, hängt neben den zugrunde liegenden Konzepten (bzw. welchen Stellenwert Qualität in diesen Konzepten einnimmt) und dem Problembewusstsein für Fehlanreizwirkungen auch von der (erwarteten) Akzeptanz solcher subjektiven Bewertungen durch die hochschulpolitischen Akteure und die betroffenen Lehrenden ab, worauf nachfolgend eingegangen wird.

### **7.3 Bedingungen erfolgreicher Institutionalisierung**

Den dritten Schwerpunkt dieser Arbeit bildete eine Diskussion von Bedingungen erfolgreicher Institutionalisierung, die im sechsten Kapitel vorgenommen wurde. Am Beginn des Kapitels wurde zunächst herausgearbeitet, daß neben der Datenqualität und Eignung von Kriterien für Leistungsbewertung und Leistungsanreize auch die Akzeptanz bei den Betroffenen – insbesondere bei einer relativ starken Rechtsstellung und Existenzsicherung – als wesentlicher Ansatzpunkt für die Diskussion von Bedingungen erfolgreicher Institutionalisierung zu sehen ist. Für eine längerfristige Akzeptanzsicherung oder –förderung wird vermutet, daß es einen Zusammenhang zwischen Akzeptanz und Arbeitszufriedenheit gibt. Deshalb wurde auch die Arbeitszufriedenheit in die Diskussion einbezogen.

Bei einer direkten Akzeptanzuntersuchung mit Hilfe der Frage, inwieweit Lehrende bestimmten hochschulpolitischen Reformvorhaben zustimmen, zeigte sich, daß bei in den Jahren 2001 bis 2003 an drei ostdeutschen Hochschulen durchgeführten Lehrendenbefragungen deutlich weniger Zustimmung zu Reformvorhaben zur leistungsorientierten Mittelzu-

weisung zu finden war als bei einer bundesweiten Lehrendenbefragung. Allerdings hängt die Zustimmung auch wesentlich von der Formulierung ab. Über unterschiedliche Frageformulierungen hinweg stabil ist die relative Abfolge der Zustimmung zu den einzelnen Gruppen von Bewertenden: Die (Einbeziehung der Bewertungen von) Experten, zu denen sich die Professoren selbst zählen können, erfährt die stärkste Akzeptanz, während die Gruppe der Absolventen (zu der sich die Mitarbeiter zählen können) etwas weniger und die der Studenten deutlich weniger Zustimmung erhält. Die Zustimmung zur Verwendung von subjektiven Bewertungen sinkt also mit dem Status der Gruppe der Bewertenden.

Geht man vom derzeitigen Stand der Akzeptanz aus, kann die Einbeziehung subjektiver Bewertungen in Leistungsbewertungen und Leistungsanreize also entweder wie bisher (z.B. bei Evaluationen) durch Experten erfolgen, oder aber zunächst über Absolventenbefragungen etabliert werden. Zwar haben diese den Nachteil, daß Aspekte der Prozeßqualität von Lehre und Studium nur mit großem zeitlichen Abstand erfasst werden können. Eine Durchsetzung studentischer Bewertungen, die eine größere Zeitnähe ermöglichen, würde aber wahrscheinlich eher zu Unterlaufungs- oder Umgehungsstrategien führen und daher unter den derzeitigen Rahmenbedingungen nicht die gewünschten Ziele erreichen.

Auf längere Sicht bleibt aber vermutlich – wenn eine zeitnahe Einschätzung von Aspekten der Prozeßqualität als wichtig eingeschätzt wird – eine Einbeziehung von Studierendenbefragungen weiter in der Diskussion, zumal diese in der Literatur überwiegend als valide beschrieben werden und die Politik in dieser Hinsicht klare Absichtserklärungen formuliert hat. Daher wurde nach Ansatzpunkten gesucht, wie die Akzeptanz möglicherweise längerfristig erhöht werden könnte. Solche Ansatzpunkte ließen sich eventuell aus der Einbeziehung der Arbeitszufriedenheit in diese Diskussion ableiten. Wenn man davon ausgeht, daß langfristig nur solche Reformmaßnahmen (höhere) Akzeptanz finden, die die Arbeitszufriedenheit der Lehrenden nicht gefährden, sondern möglichst erhöhen, dann kann von einer akzeptanzfördernden Wirkung einer hohen Arbeitszufriedenheit ausgegangen werden. Daher wurde untersucht, welche Faktoren die Arbeitszufriedenheit determinieren.

Als Fazit einer Analyse der Determinanten der Arbeitszufriedenheit wurde festgehalten, daß v.a. die berufliche Anerkennung Einfluß auf die Arbeitszufriedenheit hat. Dies ist bei Professoren stärker ausgeprägt als bei wissenschaftlichen Mitarbeitern. Die Höhe und die empfundene Leistungsgerechtigkeit des Gehalts ist dagegen von nachrangiger Bedeutung für die Arbeitszufriedenheit. Da lediglich die Lehrenden dreier Hochschulen befragt werden konnten, sind die Ergebnisse jedoch noch nicht ohne weiteres generalisierbar.

Zwei mögliche Ansatzpunkte ergeben sich aber bereits jetzt aus den vorliegenden Resultaten: Erstens wäre es möglicherweise effektiver, stärker direkt bei der beruflichen Anerkennung als bei den Besoldungszulagen (entspricht Höhe und Leistungsgerechtigkeit des Gehaltes) anzusetzen. Bisher wird diese v.a. über die Forschungsreputation erlangt. Analog Forschungspreisen könnten als ideelle Leistungsanreize in größerem Umfang z.B. Lehrpreise vergeben werden. Zweitens hängt die für Hochschullehrer wie wissenschaftliche Mitarbeiter ebenfalls wichtige berufliche Autonomie mit der Ausstattung bzw. Mittelzuweisung an die Fächer bzw. Fakultäten zusammen. In den letzten Jahren verstärkt eingeführte leistungsorientierte Mittelzuweisungen könnten die berufliche Zufriedenheit daher sowohl in die eine als auch in die andere Richtung beeinflussen. Sie könnten deshalb bei entsprechender Gestaltung für die hochschulpolitischen Ziele genutzt werden.

Um insbesondere dem Qualitätsaspekt der Lehre einen höheren Stellenwert zu verschaffen, könnte die Einbeziehung der Ergebnisse von Absolventenbefragungen hilfreich sein. Auf diese Weise würden die Lehre und ihre Qualität aufgewertet und Leistungen in diesem Bereich auch über leistungsorientierte Mittelzuweisungen noch breitere Anerkennung finden als über Lehrpreise. Diese können nur einzelne Lehrende für hervorragende Lehre „belohnen“ und haben daher wie bereits erwähnt eher symbolischen Wert für die Anerkennung der Lehre. Wenn daneben bei der Verwendung der Anteile leistungsorientierter Mittelverteilung, die zur Honorierung hoher Lehrqualität ausgeschüttet werden, ein hoher Handlungsspielraum der Verantwortlichen für die Lehre gesichert wird, begünstigt dies auch ein höheres Autonomieerleben der Lehrenden. Breitere Anerkennung und ein höheres Autonomieerleben könnten zu einer höheren Arbeitszufriedenheit und Akzeptanz führen.

Möglicherweise kann es so insgesamt gelingen, Leistungsanreizsysteme erfolgreich zu institutionalisieren, die sich in möglichst vielen Aspekten positiv und in möglichst wenigen negativ auf die Lehrqualität und die Arbeitszufriedenheit auswirken.

## 8. Literaturverzeichnis<sup>265</sup>

Ahrens, R. (2000): Eine Gefahr für die Universitäten? Forschungsevaluation in Großbritannien. In: Forschung und Lehre, DHV (Hg.), Ausgabe April 2000, S. 182-184

Aleamoni, L.M. & Hexner, Pamela (1980): A review of the research on Student Evaluation and a report on the effect of different Sets of Instructions on Student Course and Instructor Evaluation, In: Instructional Science 9/1980: 67-84

Arbuckle, J. L./ Wothke, W. (1999): AMOS 4.0 User's Guide, IL: SmallWaters Corporation, Chicago

Arnhold, N./ Kandel, K. (2004): Modellentwicklung für die W-Besoldung am Beispiel der Universität Bremen. Arbeitspapier Nr. 59, Centrum für Hochschulentwicklung (CHE), Gütersloh

Backhaus, Klaus u.a. (2000): Multivariate Analysemethoden, Springer Verlag, Berlin Heidelberg

Bahn Müller, R. (2001): Stabilität und Wandel in der Leistungsentlohnung. In: WSI-Mitteilungen Nr. 7/ 2001, Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung (Hg.), Düsseldorf, S. 426-433

Baltes-Götz, B. (2002): Analyse von Strukturgleichungsmodellen mit Amos 4.0, Universitäts-Rechenzentrum Trier (Hg.), (Online: <http://www.uni-trier.de/urt/user/baltes/docs/amos/v40/amos40.htm>)

Bank, V. (2000): Evaluationsforschung in historischer und zeitgenössischer Betrachtung. In: Bank, V./ Lames, M. (2000): Über Evaluation. BajOsch-Hein Verlag für Berufs- und Wirtschaftspädagogik, Kiel, S. 49-83

Bargel, T. (1994): Das Studium der Medizin - Eine Fachmonographie aus studentischer Sicht, Schriftenreihe Studien zur Bildung und Wissenschaft; BMBW, Bonn

Bargel, T. (1995): Studierende in den alten und neuen Bundesländern - Erfahrungen und Orientierungen Datenalmanach, herausgegeben von der Universität Konstanz, Sozialwissenschaftliche Fakultät, Arbeitsgruppe Hochschulforschung, Konstanz

Bargel u.a. (1996): Bargel, T./ Multrus, F./ Ramm, M.: Studium und Studierende in den 90er Jahren. bmbf, Bonn

Bargel u.a. (2002): Bargel, T./ Multrus, F./ Ramm, M.: Studiensituation und studentische Orientierungen. 7. Studierendensurvey an Universitäten und Fachhochschulen. *Langfassung*. bmb+f, Bonn

Bargel, T./ Ramm, M. (1998): Ingenieurstudium und Berufsperspektiven. Sichtweise, Reaktionen und Wünsche der Studierenden, bmb+f

Bartel-Lingg, G. (1996): Die Mitarbeiterorientierung im Total Quality Management, Rainer Hampp Verlag, München

Barth, G. (1999): Die Bedeutung der Qualitätssicherung für Unternehmen. In: Hochschulrektorenkonferenz (Hg.) (1999): Ein Schritt in die Zukunft- Qualitätssicherung im Hochschulbereich. Berliner Bildungsdialoge. Hochschulrektorenkonferenz und Veranstaltungsforum der Verlagsgruppe Georg von Holtzbrinck. Berlin, 26. Oktober 1998. In: Beiträge zur Hochschulpolitik 3/1999. Bonn. S.27-31

Barz, A./ Carstensen, D./ Reissert, R. (1997): Lehr- und Evaluationsberichte als Instrumente zur Qualitätsförderung. Bestandsaufnahme zur aktuellen Praxis, CHE/HIS GmbH, Bielefeld 1997

Barz, A. (1998): Strategische Positionierung und Organisationsentwicklung: Projekte des CHE. In: Hochschulrektorenkonferenz (Hg.) (1998): Evaluation und Qualitätssicherung an den Hochschulen in Deutschland- Stand und Perspektiven. Nationales Expertenseminar der Hochschulrektorenkonferenz, Bonn, 29. Mai 1998. In: Beiträge zur Hochschulpolitik 6/1998. Bonn: HRK S. 49 - 55

---

<sup>265</sup> alphabetisch geordnet

- Bayer, C. R. (2000): Anspruch und Wirklichkeit von Hochschul-Ranking: Vorschlag einer allgemeinen Methodik, In: Die Betriebswirtschaft (DBW), 60. Jg. (2000), H. 5, S. 547-69.
- Bayer, C. R. (2001): Eine Analyse des entscheidungsunterstützenden Informationsgehaltes deutscher und US-amerikanischer Hochschul-Rankings. Discussion Paper DP-WIOR Nr. 593. Institut für Wirtschaftstheorie und Operations Research Universität Karlsruhe
- Bayer, C. R. (2004): Hochschul-Ranking: Vorschlag eines ganzheitlichen Ranking-Verfahrens. Band 14 der Reihe Abhandlungen zu Bildungsforschung und Bildungsrecht, Duncker & Humblodt, Berlin
- Becher, T./ Kogan, M. (1992): Process and Structure in Higher Education, Routledge, London
- Behrens, T. (2001): Leistungsbezogene Mittelverteilung. In: Pasternack, P. (Hrsg.): Flexibilisierung der Hochschulhaushalte. Handbuch für Personalräte und Gremienmitglieder. Schüren-Verlag, Marburg, S. 68-77
- Benkhoff, B./ Huhle, A./ Kühn/ K. (2001): Erklärungsansätze für die Mitarbeiterzufriedenheit mit Arbeitszeitregelungen, In: Dresdner Beiträge zur Betriebswirtschaftslehre Nr. 54/01, TU Dresden, Fakultät Wirtschaftswissenschaften, Die Professoren der Fachgruppe Betriebswirtschaftslehre (Hg.)
- Benninghaus, H. (1994): Einführung in die sozialwissenschaftliche Datenanalyse, Oldenbourg Verlag München Wien
- Berdahl, R.O./ Moodie, G. C./ Spitzberg, I. J., Jr (1991): Quality and Access in Higher Education: Comparing Britain and the United States, The Society for Research into Higher Education & Open University Press
- Berendt, B. (1992): Stärkung der Qualität der Lehre durch hochschuldidaktische Weiterbildung und Beratung. In: Altrichter & Schratz (Hg.): Qualität von Universitäten. Innsbruck, Österreichischer Studien Verlag, S. 256-278
- Berghoff, S./ Federkeil, G./ Giebisch, P./ Hachmeister, C.-D./ Müller-Böling, D. (2004): Das Hochschulranking. Vorgehensweise und Indikatoren. CHE (Hg.), Arbeitspapier Nr. 54, Gütersloh
- Beywl, W. (Hrsg.) (2000): Handbuch der Evaluationsstandards. Die Standards des Joint Committee on Standard for Education Evaluation. 2., durchgesehene Auflage, Leske+Budrich, Opladen
- bmb+f (1991): Zum Stellenwert der Hochschullehre. In: Bildung und Wissenschaft aktuell Nr. 4/ 1991, Bonn
- bmb+f (2000a): Bericht der Expertenkommission Reform des Hochschuldienstrechts, Bonn
- bmb+f (2000b): Hochschuldienstrecht für das 21. Jahrhundert. Das Konzept des BMBF, Bonn
- Böhm, M. (2000): Demotivierende Wirkung. Monetäre Leistungsanreize im internationalen Vergleich. In: Forschung und Lehre, DHV (Hg.), Ausgabe März 2000, S. 116 ff.
- Bourdieu, P. (1998): homo academicus. Suhrkamp, Frankfurt/ Main
- Borg, I. (2000): Affektiver Halo in Mitarbeiterbefragungen, ZUMA-Arbeitsbericht 2000/03, Zentrum für Umfragen, Methoden und Analysen (Hg.), Mannheim
- Bortz, J./ Döring, N. (2002): "Forschungsmethoden und Evaluation" Springer Verlag, Berlin und Heidelberg
- Bortz, J. (1999): "Statistik für Sozialwissenschaftler" Springer Verlag, Berlin Heidelberg
- Bos, W. (1995): Was fällt Ihnen zur akademischen Lehre ein? Ein inhaltsanalytischer Beitrag zur Validierung der Evaluation der Hochschullehre durch Befragungen. In: Mohler, H. (1995): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung, Waxmann Verlag, Münster, S. 61-82

- Bothfeld, S./ Eisend, M. (2002): Zwischen Effizienz und Selbstbestimmung – Promovieren im WZB. Wissenschaftszentrum Berlin für Sozialforschung gGmbH, veröffentlicht unter Publikations- Nr. P 02 - 001
- Brosius, G. & Brosius, F. (1995): SPSS - Base System and Professional Statistics International Thomson Publishing, Bonn
- Brosius, F. (2002): SPSS 11, mitp Verlag, Bonn
- Bruce, T. (1999): Financing higher education - British and German perspectives, British Council Gemany (Hg.)
- Brüderl, J. (1996): Münchner Soziologinnen und Soziologen auf dem Arbeitsmarkt: Langfristige Veränderungen und beruflicher Erfolg. In: "Soziologie" 3/96
- Brüderl, J./ Reimer, D. (2002): Soziologinnen und Soziologen im Beruf. Ergebnisse ausgewählter Absolventenstudien der 90er Jahre. In: Stockmann, R./ Meyer, W./ Knoll, T. (2002): Soziologie im Wandel. Universitäre Ausbildung und Arbeitsmarktchancen in Deutschland. Leske+Budrich, Opladen, S. 201-214
- Büchel, F./ Weißhuhn, G. (1997): Ausbildungsinadäquate beschäftigung von Absolventen des Bildungssystems: Berichterstattung zu Struktur und Entwicklung unterwertiger Beschäftigung in West- und Ostdeutschland, Duncker & Humblodt, Berlin
- Bühl, A. & Zöfel, P. (1994): SPSS für Windows Version 6 - Praxisorientierte Einführung in die moderne Datenanalyse Addison-Wesley, Bonn/Paris
- Bülow-Schramm, M. (1992): Unter anderem Lehrevaluation - vielfältige Maßnahmen zur Förderung der Lehre an der Universität Hamburg. In: Grünh & Gattwinkel: Evaluation von Lehrveranstaltungen - Überfrachtung eines sinnvollen Instrumentes? Berlin, FU-Dokumentationsreihe, 1992: 185-196
- Bund-Länder-Arbeitsgruppe (1998): Bericht der Bund-Länder-Arbeitsgruppe Leistungsorientierte Besoldung von Professoren sowie Mitgliedern von Hochschul- und Fachbereichsleitungen. Eckpunkte eines leistungsorientierten Besoldungssystems. Drucksache BMBF - 312 - 41212-19/15, Bonn
- Burkhardt, A./ Schomburg, H./ Teichler, U. (Hg.) (2000): Hochschulstudium und Beruf – Ergebnisse von Absolventenstudien. Bonn: bmb+f.
- Buschor, E. (2002): Evaluation und New Public Management. In: Zeitschrift für Evaluation, Nr. 1/2002, S. 71-86 ([www.zfev.de](http://www.zfev.de), Stand 18.02.2004)
- Conrady, H. (2003): Performance Measurement. Nur die Besten. In: Die Mitbestimmung Nr. 12/2003, Hans-Böckler-Stiftung (Hg.), Düsseldorf
- Cicourel, A. (1970): Methode und Messung in der Soziologie, Suhrkamp, Frankfurt/ Main, 1970: 110- 152
- Cremers, E. (1998): Befragung der Studenten der TU Dresden zur Einschätzung der Studienbedingungen im Rahmen eines Forschungsseminars am Institut für Soziologie der TU Dresden (unveröffentlichte Daten).
- Dahlgaard, J.J./ Madsen, H. (1998): Erfahrungen mit der Implementierung von TQM an Hochschulen, Paper zum Vortrag auf der HRK-Tagung 28./29. 9. 1998 an der Universität Kaiserslautern (Tagungsband erschien als Band 1/1999 "Qualität an Hochschulen" in der Schriftenreihe der HRK)
- Daniel, H.-D. (1996a): Evaluierung der universitären Lehre durch Absolventen und Studenten. In: ZSE, Heft 2/ 1996: 149-164
- Daniel, H.-D. (1996b): Korrelate der Fachstudiendauer von Betriebswirten - Ergebnisse einer Absolventenbefragung an der Universität Mannheim. In: Zeitschrift für Betriebswirtschaft (ZfB) Ergänzungsheft 1/1996: 95-115

- Daniel, H.-D. (1998): Studentische Beurteilung von Lehrveranstaltungen – Anlage, Durchführung und Ergebnisse eines Modellprojektes an der Universität Mannheim. In: HRK (Hg.) Evaluation und Qualitätssicherung an den Hochschulen in Deutschland – Stand und Perspektiven Beiträge zur Hochschulpolitik 6/1998: 79-104
- Daniel, H.-D. (1999): 16 Fragen für 12374 Studenten. Ranking-Experte Hans-Dieter Daniel über die Entstehung der Spiegel-Ranglisten. In: Der Spiegel Nr. 15/ 1999, S. 97 (ausführlichere Dokumentation hierzu vgl. Hornbostel/ Daniel 1996)
- Daxner, M. (1999): Evaluation, Indikatoren und Akkreditierung. Auf dem Weg in die Rechtfertigungsgesellschaft. In: Hochschulrektorenkonferenz (Hg.) (1999): "Viel Lärm um nichts?" Evaluation von Studium und Lehre und ihre Folgen. Tagung an der Universität Rostock vom 6. Bis 8. September 1998. Beiträge zur Hochschulpolitik 4/1999. Bonn. S. 41-49
- DeGEval (2002): Standards für Evaluation, Deutsche Gesellschaft für Evaluation e.V. (Hg.), Köln
- Diekmann, Andreas (2000): "Empirische Sozialforschung - Grundlagen, Methoden, Anwendungen", Rowohlt Taschenbuch Verlag, Reinbeck bei Hamburg
- Deeke, A. (1995): Experteninterviews – ein methodologisches und forschungspraktisches Problem. In: Brinkmann, C./ Deeke, A./ Völkel, B.: Experteninterviews in der Arbeitsmarktforschung. Diskussionsbeiträge zu methodischen Fragen und praktischen Erfahrungen. Beiträge zur Arbeitsmarkt- und Berufsforschung 191, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. Nürnberg. S. 7-22
- Diehl, J. M. und Kohr, H. U. (1977): "Entwicklung eines Fragebogens zur Beurteilung von Hochschulveranstaltungen", Psychologie in Erziehung u. Unterricht 24/ 1977: 61-75
- Diehl, J. M. (1994): Fragebogen zur studentischen Evaluation von Hochschulveranstaltungen, (Manual, Überarbeitete Fassung), Institut für Psychologie, Universität Gießen, 1994
- Diekmann, A. (2002): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen, 8. Auflage, Rowohlt, Reinbek bei Hamburg
- DIN (1995): "Qualitätsmanagement, Statistik, Zertifizierung: Begriffe aus DIN-Normen" Deutsches Institut für Industrienormung
- Doerry, M./ Mohr, J. (1999): Das aktuelle Spiegel-Ranking. Die besten Hochschulen in Deutschland. Walhalla und Praetoria Verlag, Regensburg 1999
- Donsbach, W. (1995): Studentenumfrage 1995, Institut für Kommunikationswissenschaft TU Dresden (Hg.)
- Dubs, R. (1998): Qualitätsmanagement für Schulen. St. Gallen: Institut für Wirtschaftspädagogik
- Duden (2000): Wissenschaftlicher Rat der Dudenredaktion (Hg.): Duden. Die deutsche Rechtschreibung, 22., völlig neu bearbeitete und erweiterte Auflage, Dudenverlag, Mannheim, Leipzig, Zürich, Wien
- Eberle, K.-R./ Patzak, K./ Unger, C. (1997): Methodik: So wurden die Daten erhoben und verrechnet. In: Focus Nr. 16/ 1997, S. 155 (außerdem ausführlicher dazu die leider unbetitelte Dokumentation, die auf Anfrage von der Redaktion zugesendet und von denselben Autoren verantwortet wurde)
- Ebers, N./ Streitferdt, L. (1999): Controlling. Ein leistungsfähiges Instrument zur Steuerung von Hochschulen. Dokumentation einer Ringvorlesung an der Universität Hamburg, hg. vom Präsidenten der Universität Hamburg - Projekt Universitätsentwicklung -
- Ederer, F. (1999): Balanced Scorecard. In: Wirtschaft und Unterricht, Ausgabe September 1999, Institut der Deutschen Wirtschaft (Hg.), Köln

- Enders/ Bormann (2001): Enders, Jürgen/ Bormann, Lutz: Karriere mit Dokortitel? Ausbildung, Berufsverlauf und Berufserfolg von Promovierten, Campus Verlag, Frankfurt/ New York
- Enders, J./ Teichler, U. (1993): Der Hochschullehrerberuf im internationalen Vergleich. Ergebnisse einer Befragung über die wissenschaftliche Profession in 13 Ländern, BMBF (Hrsg.), Bonn
- Enders, J./ Teichler, U. (1995) (Hrsg.): Der Hochschullehrerberuf. Aktuelle Studien und ihre hochschulpolitische Diskussion, Neuwied, Luchterhand
- Endruweit, Günter (1993): Programmevaluation als Laienspiel. Bemerkungen über Meinungsforschung, Sozialforschung und Pfusch bei Studentenforschungen. In: MittHV 2/93: 129-132
- Endruweit, Günter (2002): Lehrevaluation und Lehrziele. Neue Wege der Semesterauswertung. In: Forschung & Lehre (MittHV) 8/2002
- Engel, Uwe/ Krekeler, Gaby (1999): Der Wechsel vom Grundstudium ins Hauptstudium. Studentische Einschätzungen zu Beginn des Hauptstudiums an der Universität Potsdam im Wintersemester 1998/ 99. Potsdam: Universität Potsdam Arbeitsgruppe "Qualität der Lehre"
- Engel, U./ Pohlenz, P. (2001): Lehre und Studium im Spiegel studentischer Bewertungen: Über das Potsdamer Modell der Lehrevaluation. In: Spiel, Christiane (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck, Waxmann, Münster, S. 131-150
- Engler, S. (2003): Zur Kombination von qualitativen und quantitativen Methoden. In: Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München
- Esser, Hartmut (1997): Zweifel an der Evaluation der Lehre. In: WiSt (2), Heft 1, Januar 1997, S. 45-49
- Europäisches Komitee für Normung (CEN) (1994): Europäische Norm EN ISO 9000-1. Deutsche Fassung. Normen zum Qualitätsmanagement und zur Qualitätssicherung/ QM- Darlegung. Teil 1. Leitfaden zur Auswahl und Anwendung (ISO 9000-1:1994). Brüssel
- Expertenkommission „Reform des Hochschuldienstrechts“ (2000): Bericht der Expertenkommission
- Faist, U. (2000): Klassifikation der Lehrevaluation an wirtschaftswissenschaftlichen Fakultäten in Deutschland. Abschlußarbeit im Rahmen des Seminars „Methodische Probleme bei der Erstellung von Ranking-Modellen“, Leitung: Bayer, C.R./ Eichhorn, W., Institut für Wirtschaftstheorie und Operations Research, Universität Karlsruhe
- Feldmann, K.A. (1989): Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. In: Reserch in Higher Education. 30 (2). S. 137-194
- Felger, S./ Paul-Kohlhoff, A. (2004): Human Ressource Management. Konzepte, Praxis und Folgen für die Mitbestimmung. In: edition Hans-Böckler-Stiftung Nr. 102, Hans-Böckler-Stiftung, Düsseldorf
- Ferland, J.A. & Fleurent, C. (1994): SAPHIR: A Decision Support System for Course Scheduling. In: Interfaces 24: 105-115
- Fey, M. (1998): Verhältnis interne – externe Revision in öffentlichen Verwaltungen. USA, Deutschland, Österreich, Schweiz. Schriftenreihe der Treuhand-Kammer, Band 157, Verlag Hans Schellenberg, Zürich
- Finkemeier, T. (2003): Aus Erfahrung skeptisch. Variabilisierte Entlohnung ist kein Thema, bei dem Betriebsräte euphorisch werden... In: Die Mitbestimmung Nr. 6/2003, Hans-Böckler-Stiftung, Mitbestimmungs-, Forschungs- und Studienförderungswerk des DGB (Hg.), Düsseldorf

- Fischer, W./ Voss, H.-P. (1998): Evaluation und Enthusiasmus Studienkommission für Hochschuldidaktik an Fachhochschulen in Baden-Württemberg (Hg.)
- Flick, U. (1992): Entzauberung der Intuition. Systematische Perspektiven-Triangulation als Strategie der Geltungsbegründung qualitativer Daten und Interpretationen. In: Hoffmeyer-Zlotnik, J. (Hg.): Analyse verbaler Daten, Opladen, S. 11-55
- Flick, U. (1995): Handbuch qualitative Sozialforschung, 2. Aufl. München: Beltz, Psychologie-Verl.-Union, 1995
- Forbeck, S. (2000): Nach der Thatcher-Revolution. Hochschulen in Großbritannien. in: Forschung und Lehre, DHV (Hg.), Ausgabe März 2000, S. 178-181
- Forberg, T. (2002): Modellierung eines Wertpapierhandelssystems als Grundlage zur Gestaltung eines computergestützten komplexen Lehr-Lern-Arrangements für die Wirtschaftslehre. Diplomarbeit, Lehrstuhl Wirtschaftspädagogik, Fakultät Wirtschaftswissenschaften TU Dresden
- Frackmann, E. (1988): Die Bedeutung des Rankings für den Wettbewerb im Hochschulbereich. In: "Betriebswirtschaftliche Forschung und Praxis" 2/ 1988, S. 138-177
- Franzen, A./ Hecken, A. (2002): Studienmotivation, Erwerbsspartizipation und der Einstieg in den Arbeitsmarkt. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 54, Dezember 2002, S. 733-752
- Frey, W. (1995): ...und die Arbeit macht der Computer. In: Handbuch Hochschullehre Highlights, Band 1: Evaluation der Lehre. Ziele - Akzeptanz – Methoden
- Frederiks, M. (2001): Qualitätssicherung in Großbritannien. In: Beiträge zur Hochschulpolitik 10/2001, Hochschulrektorenkonferenz (Hg.), Bonn
- Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München
- Gnahn, D./ Kregel, E./ Wolter, A. (1995): Qualitätsmanagement im Bildungswesen" In: Schriften des Internationalen Begegnungszentrums Sankt Marienthal (IBZ), Band 2, Hannover/ Berlin/ Dresden
- Goedegebuure u.a. (1992): Hochschulpolitik international: Trends - Probleme – Lösungsansätze. Eine länderübergreifende Synopse von: Goedegebuure, Leo/ Kaiser, Frans/ Maasen, Peter/ Meck, Lynn/ van Vught, Frans/ de Weert, Egbert herausgeg. von der Bertelsmann Stiftung, Gütersloh
- Goedegebuure, L./ Kaiser, F./ Maasen, P./ Meck, L./ van Vught, F./ de Weert, E. (1993): Hochschulpolitik im internationalen Vergleich - Eine länderübergreifende Untersuchung im Auftrag der Bertelsmann Stiftung. Verlag Bertelsmann Stiftung, Gütersloh
- Gralki, H./ Grünh, D./ Hecht, H. (Hg.) (1993): "Evaluation schafft Autonomie" Berlin
- Greenwald, A. (1997): Validity Concerns and Usefulness of Student Ratings of Instruction. In: American Psychologist, Vol. 52, No. 11., Nov. 1997, S. 1182- 1186
- Grieshop, H. (2004): Vorbild Großbritannien? Akkreditierung und Qualitätskontrolle an britischen Hochschulen. In: Forschung und Lehre, DHV (Hg.), Ausgabe Oktober 2004, S. 550-551
- Grünh, D./ Gattwinkel, H. (1992): Evaluation von Lehrveranstaltungen - Überfrachtung eines sinnvollen Instrumentes?: Informationen aus Lehre und Forschung 2/1992, FU Berlin
- Grüning, M. (2002): Performance-Measurement-Systeme. Wiesbaden: Gabler, Deutscher Universitäts Verlag, Gabler Edition Wissenschaft.

Hage, N. (1996a): Lehrevaluation und studentische Veranstaltungskritik - Projekte, Instrumente und Grundlagen Bonn, bmb+f

Hage, N. (1996b): Studentische Urteile über die Lehr- und Studienqualität - Analysen zu ihrer Validität anhand der Indikatoren und Skalen im Studierendensurvey. In: Hefte zur Bildungs- und Hochschulforschung (19) Universität Konstanz,

Hägele, H.. (1995): Experteninterviews in der öffentlichen Verwaltung: Ausgewählte praktische Probleme. In: Brinkmann, C./ Deeke, A./ Völkel, B.: Experteninterviews in der Arbeitsmarktforschung. Diskussionsbeiträge zu methodischen Fragen und praktischen Erfahrungen. Beiträge zur Arbeitsmarkt- und Berufsforschung 191, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. Nürnberg.

Hartmann, H. (1989): Mängel im soziologischen Lehrangebot. In: Soziale Welt, 1989: 220-232

Hertmann, M./ Kopp, J. (2001): Elitenselektion durch Bildung oder durch Herkunft? Promotion, soziale Herkunft und der Zugang zu Führungspositionen in der deutschen Wirtschaft. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 53, Heft 3, 2001, S. 437-466

Hartung, D./ Teichler, U. (1981): Einige aktuelle Aspekte zum Verhältnis von Bildungs- und Beschäftigungssystem. In: Hartung, D./ Nuthmann, R./ Teichler, U. (Hrsg.): Bildung und Beschäftigung. Probleme, Konzepte, Forschungsperspektiven. München, New York, London u.a.: K G Saur. S. 43 – 64

Hartwig, L./ Küpper, H.-U. (2002): Neue Modelle der Hochschulfinanzierung und -Steuerung. In: Beiträge zur Hochschulforschung, Bayrisches Staatsinstitut für Hochschulforschung und Hochschulplanung (Hg.), München, Heft 3/2002, S. 4-5

HIS (1992): Das soziale Bild der Studentenschaft in der Bundesrepublik Deutschland - 13. Sozialerhebung Schriftenreihe Studien zu Bildung und Wissenschaft, Bonn, bmbw

HIS (1992): Dokumentation Evaluation der Lehre, Teil.2, HIS GmbH, Hannover

HIS (1994): Evaluation der Lehre - Interne Selbstevaluation und externe Begutachtung durch Peers, HIS-Kurzinformationen A8/1994, HIS GmbH, Hannover

HIS (1995): Das soziale Bild der Studentenschaft in der Bundesrepublik Deutschland - 14. Sozialerhebung. Schriftenreihe Studien zu Bildung und Wissenschaft, Bonn, bmbw

HIS (1995a): Interne und externe Evaluation - Modell und Praxis - Eine Zwischenbilanz aus Sicht von HIS, HIS-Kurzinformationen A16/1995, HIS GmbH, Hannover

HIS (1995b): Studium im Freistaat Sachsen. Urteile von Absolventen sächsischer Hochschulen über ihr Studium, HIS GmbH, Hannover

HIS (1995c): Absolventenreport Magisterstudiengänge. Ergebnisse einer Längsschnittuntersuchung zum Berufsübergang von Absolventinnen und Absolventen der Magisterstudiengänge, bmb+f

HIS (1995d): Absolventenreport Rechtswissenschaft. Ergebnisse einer Längsschnittuntersuchung zum Berufsübergang von Absolventinnen und Absolventen der Rechtswissenschaft, bmb+f

HIS (1995e): Studienabbruch: Gründe und anschließende Tätigkeiten. Ergebnisse einer bundesweiten Befragung im Studienjahr 1993/94, Hannover, HIS-Kurzinformation A1/95

HIS (1996): Absolventenreport Ingenieure. Ergebnisse einer Untersuchung zum Berufsübergang von Absolventen ingenieurwissenschaftlicher Diplomstudiengänge, bmb+f

HIS (1997): Defizite der Hochschulausbildung aus Sicht der Hochschulabsolventen HIS-Pressemitteilung vom 22. 5. 1997, HIS GmbH, Hannover

- HIS (1998a): Praxis der internen und externen Evaluation. Handbuch zum Verfahren. HIS-Kurzinfation spezial (März 1998), HIS GmbH, Hannover
- HIS (1998b): Ausbildung und Qualifikation von Ingenieuren: Herausforderungen und Lösungen aus transatlantischer Perspektive. HIS-Kurzinfomation A 6/98, bmb+f
- HIS (1998c): Ingenieurstudium - Daten, Fakten, Meinungen
- HIS (1998d): Das soziale Bild der Studentenschaft in der Bundesrepublik Deutschland - 15. Sozialerhebung. Schriftenreihe Studien zu Bildung und Wissenschaft, Bonn, bmb+f
- HIS (1998e): Studienabbruch – Typologie und Möglichkeiten der Abbruchquotenbestimmung, HIS-Kurzinformation A5/98, HIS GmbH, Hannover
- HIS (1999): Studienanfänger 1998/99. HIS-Kurzinformation A7/99, HIS GmbH, Hannover
- HIS (2000): Absolventenreport 2000, HIS GmbH, Hannover
- HIS (2000a): Lebensorientierungen und Studienmotivation von Studienanfängern, HIS-Kurzinformation A5/2000, HIS GmbH, Hannover
- HIS (2001): Das soziale Bild der Studentenschaft in der Bundesrepublik Deutschland - 16. Sozialerhebung. Schriftenreihe Studien zu Bildung und Wissenschaft, Bonn, bmb+f
- HIS (2002): Studienabbruchstudie 2002. HIS-Kurzinformation A5/2002, HIS GmbH, Hannover
- HIS (2003): Indikatoren zur Ausbildung im Hochschulbereich. Bericht für 2003 im Rahmen des Berichtssystems zur Technologischen Leistungsfähigkeit Deutschlands. HIS-Kurzinformation A3/2003, HIS GmbH, Hannover
- HIS (2004): Evaluation der leistungsbezogenen Mittelvergabe auf der Ebene Land-Hochschulen in Berlin. Gutachten im Auftrag der Berliner Senatsverwaltung für Wissenschaft, Forschung und Kultur, HIS-Kurzinformation A4/2004, HIS GmbH, Hannover
- Hochschild, A. R. (1998): Der Arbeitsplatz wird zum Zuhause, das Zuhause zum Arbeitsplatz. Total Quality Management bei der Arbeit, Taylorismus im Familienalltag – kann das auf Dauer gut gehen? In: Harvard Business Manager, Nr. 3/1998, S. 29-41
- Hochschulrektorenkonferenz (HRK) (Hg.) (1993): Zur Evaluation im Hochschulbereich unter besonderer Berücksichtigung der Lehre. Entschließung des 176. Plenums der HRK vom 3.7.1995. Bonn: HRK
- Hochschulrektorenkonferenz (Hg.) (1998a): Qualitätsmanagement in der Lehre, TQL 98. Tagung des hessischen Arbeitskreises "Qualitätsmanagement in der Lehre" und der Hochschulrektorenkonferenz, Wiesbaden, 4. Februar 1998. Beiträge zur Hochschulpolitik 5/1998. Bonn
- Hochschulrektorenkonferenz (Hg.) (1998b): Evaluation und Qualitätssicherung an den Hochschulen in Deutschland - Stand und Perspektiven. Nationales Expertenseminar der Hochschulrektorenkonferenz, Bonn, 29. Mai 1998. Beiträge zur Hochschulpolitik 6/1998. Bonn: HRK
- Hochschulrektorenkonferenz (Hg.) (1998c): Empfehlungen zum Dienst- und Tarif-, Besoldungs- und Vergütungsrecht sowie zur Personalstruktur in den Hochschulen. Bonn: HRK
- Hochschulrektorenkonferenz (Hg.) (1999a): Qualität an Hochschulen. Fachtagung der Universität Kaiserslautern und der Hochschulrektorenkonferenz. Kaiserslautern, 28./ 29. September 1998. Beiträge zur Hochschulpolitik 1/1999. Bonn

Hochschulrektorenkonferenz (Hg.) (1999b): Ein Schritt in die Zukunft - Qualitätssicherung im Hochschulbereich. Berliner Bildungsdialoge. Hochschulrektorenkonferenz und Veranstaltungsforum der Verlagsgruppe Georg von Holtzbrinck. Berlin, 26. Oktober 1998. Beiträge zur Hochschulpolitik 3/1999. Bonn

Hochschulrektorenkonferenz (Hg.) (1999c): "Viel Lärm um nichts?" Evaluation von Studium und Lehre und ihre Folgen. Tagung an der Universität Rostock vom 6. bis 8. September 1998. Beiträge zur Hochschulpolitik 4/1999. Bonn.

Hochschulrektorenkonferenz (Hg.) (1999d): Gemeinsame Ziele. Evaluation, Qualitätssicherung und Akkreditierung in Deutschland und der Mongolei, Beiträge zur Hochschulpolitik 7/1999. Bonn.

Hochschulrektorenkonferenz (Hg.) (1999e): Qualitätsentwicklung in der Ingenieurausbildung, Beiträge zur Hochschulpolitik 12/1999. Bonn.

Hochschulrektorenkonferenz (Hg.) (2000a): Wegweiser 2000 durch die Qualitätssicherung in Lehre und Studium. Dokumente und Informationen 2/2000. Bonn.

Hochschulrektorenkonferenz (Hg.) (2000b): Erfahrungsberichte zum Qualitätsmanagement im Hochschulbereich. Beiträge zur Hochschulpolitik 4/2000. Bonn.

Hochschulrektorenkonferenz (Hg.) (2000c): ... und im Streben immer der Erste. Beiträge zur Hochschulpolitik 5/2000. Bonn.

Hochschulrektorenkonferenz (Hg.) (2000d): Voneinander lernen. Hochschulübergreifende Qualitätssicherung in Netzwerken und Verbänden, Beiträge zur Hochschulpolitik 7/2000. Bonn.

Hochschulrektorenkonferenz (Hg.) (2001): Frauen - Technik - Evaluation. Frauenförderung in technisch-naturwissenschaftlichen Studiengängen. In: Beiträge zur Hochschulpolitik 3/2001. Bonn

Hochschulrektorenkonferenz (Hg.) (2001a): Zukunftsaufgabe Qualitätsentwicklung – 3 Jahre Projekt Q. In: Beiträge zur Hochschulpolitik 4/2001. Bonn.

Hochschulrektorenkonferenz (Hg.) (2001b): Selbstvergewisserung. Zur Einführung eines Qualitätsmanagement-Systems am Fachbereich III – Wirtschaftswissenschaften – der Fachhochschule Mainz. In: Beiträge zur Hochschulpolitik 5/2001. Bonn.

Hochschulrektorenkonferenz (Hg.) (2002): Evaluation, was nun? Erfahrungen mit der Umsetzung von Evaluationsergebnissen. In: Beiträge zur Hochschulpolitik 1/2002. Bonn.

Hochschulrektorenkonferenz (Hg.) (2003a): Evaluation und ihre Konsequenzen, Beiträge zur Hochschulpolitik 2/2003. Bonn

Hochschulrektorenkonferenz (Hg.) (2003b): Wegweiser 2003. Qualitätssicherung an Hochschulen, Beiträge zur Hochschulpolitik 7/2003. Bonn

Hochschulrektorenkonferenz (Hg.) (2004a): Evaluation und Akkreditierung: Bluffen – vereinheitlichen – profilieren?, Beiträge zur Hochschulpolitik 1/2004. Bonn.

Hochschulrektorenkonferenz (Hg.) (2004b): Metaevaluation. Evaluation von Studium und Lehre auf dem Prüfstand, Beiträge zur Hochschulpolitik 4/2004b. Bonn.

Hofmann, J. M. (1988): Studienmotivation und Veranstaltungsbeurteilung. In: Psychologie in Erziehung und Unterricht 35, S. 119 - 126

Hohmann-Dennhardt, C. (1998): Die Verantwortung der Hochschulen für die Qualität der Lehre- und die Mitverantwortung des Staates. In: Hochschulrektorenkonferenz (HRK) (Hg.) (1998): Qualitätsmanagement in der Lehre, TQL 98. Tagung des hessischen Arbeitskreises "Qualitätsmanagement in der Lehre" und der Hochschulrektorenkonferenz, Wiesbaden, 4. Februar 1998. Beiträge zur Hochschulpolitik 5/1998. Bonn

- Holtkamp, R./ Minks, K.-H./ Schaeper, H. (1998): Fachhochschulabsolventen im Strukturwandel des Beschäftigungssystems. Eine Untersuchung des Berufsübergangs der Absolventenkohorten 1989 und 1993. Mit einem tabellarischen Anhang: Absolventen der Universitäten, bmb+f
- Hornbostel, S./ Daniel, H.-D. (1996): Die Studienbedingungen in der Soziologie und die Bewertung der soziologischen Lehre durch die Studenten in Artus, Helmut M. und Herfurth, Matthias (Hrsg.): Soziologielehre in Deutschland - Lehre, Studium, beruflicher Verbleib, Lehrangebot, Studien- und Prüfungsordnungen. Opladen: Leske + Budrich 1996: 11-57
- Hornbostel, S. (1998): Der Uni-Test Europa des SPIEGEL: Infotainment oder Entscheidungshilfe? In: Doerry, M. und Mohr, J. (Hg.): Uni-Test Europa. Wo sich das Studieren im Ausland lohnt. Hamburg: Hoffmann & Campe, S.149 - 162.
- Hornbostel, S. (1999): Das SPIEGEL-Ranking deutscher Hochschulen und die Folgen: Interaktionsprozesse zwischen Öffentlichkeit und Wissenschaft. In: Gerhards, Jürgen und Hitzler, Ronald (Hg.): Die Eigenwilligkeit sozialer Prozesse. Friedhelm Neidhardt zum 65. Geburtstag. Opladen: Westdeutscher Verlag
- Hornbostel, S. (1999): Evaluation und Ranking- Führen sie zu mehr Transparenz und Vergleichbarkeit? In: Hochschulrektorenkonferenz (Hg.) (1999): "Viel Lärm um nichts?" Evaluation von Studium und Lehre und ihre Folgen. Tagung an der Universität Rostock vom 6. bis 8. September 1998. Beiträge zur Hochschulpolitik 4/1999. Bonn. S. 81-95
- Hornbostel, S. (1999): Welche Indikatoren zu welchem Zweck: Input, Throughput, Output. In: Röbbcke, Martina und Dagmar Simon: Qualitätsförderung durch Evaluation? Ziele, Aufgaben und Verfahren im Wandel, Wissenschaftszentrum Berlin, P99-003.
- Hornbostel, S. (2001): Der Studienführer des CHE – ein multidimensionales Ranking. In: Engel, Uwe (Hg.): Hochschulranking. Zur Qualitätsbewertung von Studium und Lehre. Frankfurt/M.: Campus.
- Hornbostel, S. (2001): Hochschulranking: Beliebigkeit oder konsistente Beurteilungen? Rankings, Expertengruppen und Indikatoren im Vergleich, In: Müller-Böling, Detlef, Stefan Hornbostel und Sonja Berghoff (Hg.): Hochschulranking - Aussagefähigkeit, Methoden, Probleme. Gütersloh: Verlag Bertelsmann Stiftung.
- Horvath, P. (1999): Wissensmanagement steuern: Die Balanced Scorecard als innovatives Controllinginstrument. In: innovation-aktuell.de (Fachbeiträge), Symposium Verlag, Düsseldorf
- Horvath & Partner (Hrsg.) (2000): Balanced Scorecard umsetzen. Schäffer-Pöschel Verlag, Stuttgart
- Hugl, Ulrike (1995): Qualitative Inhaltsanalyse und mind mapping: Ein neuer Ansatz für Datenauswertung und Organisationsdiagnose. Gabler, Wiesbaden
- Hübner, P./ Rau, E. (2001): Leistungsorientierte Mittelvergabe an der Freien Universität Berlin. Ergebnisse und Perspektive einer Befragung von Mitgliedern der Universität. In: Beiträge zur Hochschulforschung, Heft 1/ 2001, München, S. 95-118
- Jackson, M.P. (1997): The Impact of discretionary pay in UK Universities. In: Higher Education Management, July 1997, Vol. 9, No. 2, Organisation for Economic Co-operation and Development (OECD), Paris
- Jahoda, M./ Lazarsfeld, P. F./ Zeisl, J. (1960): Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit. Mit einem Anhang zur Geschichte der Soziographie. Suhrkamp, New York
- Jochmann-Döll, A./ Tondorf, K. (2004): Monetäre Leistungsanreize im öffentlichen Sektor. Bestandsaufnahme – Analysen – Gestaltungsempfehlungen. edition der Hans-Böckler-Stiftung (Hg.) Nr. 119, Düsseldorf

- Johnes/ Taylor (1990): Performance Indicators in Higher Education. The Society for Research into Higher Education & Open University Press
- Kahl, R. (2002): Finnische Lektionen. In: Die Zeit Nr. 23/ 2002, Zeit-Verlag, Hamburg
- Kaplan, R.S./ Norton, D.P. (1997): Balanced Scorecard: Strategien erfolgreich umsetzen. Schäffer-Pöschel Verlag, Stuttgart
- Karpen, U. (1991): The Financing of Higher Education: Between Public Administration and Self-Administration. In: Karpen, U.: The Financing of Higher Education in the Federal Republic of Germany, Nomos-Verlags-Gesellschaft, Baden-Baden
- Keller, A. (2000): Hochschulreform und Hochschulrevolte. Selbstverwaltung und Mitbestimmung in der Ordinarienuiversität, der Gruppenshochschule und der Hochschule des 21. Jahrhunderts, BdWi Verlag, Marburg
- Kellermann u. a. (1991): Studienmotivation und Arbeitsperspektiven. In: Buchenberger, Helmut: "Hochschulzugang und Studienwahl" in "Klagenfurter Beiträge"
- Kellermann, Paul (1992): Methodologische Überlegungen zu Evaluationen im Hochschulbereich. In: 11. GEW-Sommerschule "Qualität von Studium und Lehre - Gespräche über eine vernachlässigte Aufgabe der Hochschulen" Materialien und Dokumente Hochschule und Forschung, Bd. 66, GEW (Hg.) 1992: 135-144
- Kempen, B. (2004): Fairneß statt Willkür. Der Präsident des Deutschen Hochschulverbandes schlägt ein völlig neues Modell der Finanzierung der Hochschulen vor. Das Vorbild: Der öffentlich-rechtliche Rundfunk. In: Die Zeit Nr. 37/ 2004, Hamburg
- Kerst, C./ Minks, K.-H. (2004): Fünf Jahre nach dem Studienabschluß – Berufsverlauf und aktuelle Situation von Hochschulabsolventen des Prüfungsjahrganges 1997. Projektbericht, HIS GmbH, Hannover
- Klein-Schneider, H. (1999): Leistungs- und erfolgsorientiertes Entgelt. Analyse und Handlungsempfehlungen. edition der Hans-Böckler-Stiftung (Hg.), Düsseldorf
- Klockner, C. (1997): Studieren im nächsten Jahrtausend. Perspektiven zur Fortentwicklung des Hochschulstudiums. In: Goebel, J.: Hochschulen in Deutschland. Forschung und Lehre für das nächste Jahrtausend. Reihe Schmalkalden (4), Fachhochschule Schmalkalden (Hg.), LIT Verlag, Hamburg
- Knoll, J. (1999): Qualitätsmanagement an Universitäten. Leitfaden zur Einführung eines Qualitätsmanagementsystems nach DIN EN ISO 9000 ff. im Arbeitsbereich von Professuren bzw. Lehrstühlen (Schwerpunkt: Lehre), Professur für Erwachsenenpädagogik, Erziehungswissenschaftliche Fakultät der Universität Leipzig (Hg.)
- Koebke, J./ Neugebauer, E./ Lefering, R. (1996): Die Qualität der Lehre in der Medizin. München, Wien, Baltimore; Urban und Schwarzenberg
- König, J. (2000): Einführung in die Selbstevaluation. Lambertus, Freiburg im Breisgau
- Kogan, M./ Hanney S. (1999): Reforming Higher Education. Jessica Kingsley Publishers, London
- Kopp, R. & Weiß, M. (1993): Der Arbeitsplatz Universtät und die Zukunft der Hochschulen aus Sicht von Hochschullehrern, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt/ Main
- Krafft, A./ Ulrich, G. (1995): Akteure in der Sozialforschung. In: Brinkmann, C./ Deeke, A./ Völkel, B.: Experteninterviews in der Arbeitsmarktforschung. Diskussionsbeiträge zu methodischen Fragen und praktischen Erfahrungen. Beiträge zur Arbeitsmarkt- und Berufsforschung 191, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. Nürnberg.

- Krause, D./ Stettler, J.-C. (1998): Beschäftigungsstatistik 1997, Teil A: Absolventinnen mit Diplom 1997, Absolventinnen mit Diplom 1994-97, Vergleich Absolventinnen mit Diplom und Doktorat 1994-97, Informationsmanagement ETH Zürich
- Krauth, H. (1998): Arbeitsmarkt für Physikerinnen und Physiker. Statistik, Analysen, Trends 1998. In: "Physikalische Blätter" November/1998
- Krempkow, R. (1997): Ist gute Lehre meßbar? - Untersuchungen zur Validität, Zuverlässigkeit und Vergleichbarkeit studentischer Lehrbewertungen. Diplomarbeit, vorgelegt 1997 am Institut für Soziologie der TU Dresden (1999 unter o.g. Titel veröffentlicht im Tectum Verlag, Marburg)
- Krempkow, R.: Ist "gute Lehre" meßbar? Die Verwendbarkeit studentischer Lehrbewertungen zur Darstellung der Lehrqualität und weiteren Maßnahmen, In: Das Hochschulwesen, 46. Jahrgang, 4. Quartal 1998, Neuwied, Luchterhand Verlag 1998, S. 195-199
- Krempkow, R./ Grimm, H. (1999): Die Bewertung der Studienbedingungen am Institut für Soziologie der TU Dresden aus Sicht der Studierenden, Institut für Soziologie TU Dresden
- Krempkow, R. (1999): Zwischenbericht. Befragung der Absolventen der TU Dresden zur Öffentlichkeitsarbeit, zur retrospektiven Bewertung des Studiums und zum beruflichen Verbleib. Hg.: Universitätsmarketing/ Dezentrat Forschungsförderung und Öffentlichkeitsarbeit TU Dresden
- Krempkow, R. (2000): Absolventenstudien Maschinenwesen und Mathematik/ Naturwissenschaften TU Dresden 1999, Universitätsmarketing TU Dresden (unveröffentlichte Präsentationsmappen)
- Krempkow, R./ Heldt, M. (2000a): Was kam raus? - Eine erste Auswertung der Studentenbefragungen zur Qualität der Lehre an der TU Dresden, Hg.: Prorektor Bildung/ Dezentrat Akademische Angelegenheiten
- Krempkow, R./ Heldt, M. (2000b): Das Lehrevaluationsverfahren an der TU Dresden - Entwicklung, Erfahrungen und Ergebnisse, Vortrag anlässlich des 1. Ingenieurpädagogischen Kolloquiums am 10.02.2000 an der TU Dresden
- Krempkow, R./ Heldt, M. (2000c): Einflussfaktoren auf studentische Lehrbewertungen an sächsischen Hochschulen. In: hochschule ost Nr. 1/2 2000, Leipziger Beiträge zu Hochschule & Wissenschaft, Leipzig
- Krempkow, R./ Heldt, M. (2000d): Bericht über die Ergebnisse der Studentenbefragungen zur Qualität der Lehrveranstaltungen im Wintersemester 99/00, Studentisches Evaluationsbüro Sachsen - SES, Dresden
- Krempkow, R./ König, K. und Winter, J. (2000): Studienführer Sachsen: Ingenieurwissenschaften - Eine Gegenüberstellung von Daten aus den Lehrberichten der Hochschulen und ausgewählten Ergebnissen studentischer Lehrbewertungen, Studentisches Evaluationsbüro Sachsen - SES, Dresden
- Krempkow, R./ Winter, Jana (2000): Bericht über die Ergebnisse der Studentenbefragungen zur Qualität der Lehrveranstaltungen im Sommersemester 2000 - Gegenüberstellung der Ergebnisse Sommersemester 1999 bis Sommersemester 2000, Studentisches Evaluationsbüro Sachsen – SES, Dresden
- Krempkow, R. (2001): Ist "gute Lehre" meßbar? Die Verwendbarkeit studentischer Lehrbewertungen zur Darstellung und Verbesserung der Lehrqualität in der Architektur In: Johannes, Ralf (Hg.): Architekturausbildung im Entwerfen in Europa – von Vitruv bis zum bauhaus und danach
- Krempkow, R./ Thieme, L. (2001): Zwischen den Stühlen? Berufseinstieg und Berufserfahrungen der DHfK-Studienanfänger des Jahrganges 1988. In: Sportwissenschaftliche Beiträge der Universität Leipzig
- Krempkow, R. (2001): Dresdner Absolventenstudien 2000: Wirtschaftswissenschaften. Abschlußbericht der Befragung von Absolventen der Fakultät Wirtschaftswissenschaften der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>

Krempkow, R. (2001): Dresdner Absolventenstudien 2000: Elektrotechnik. Abschlußbericht der Befragung von Absolventen der Fakultät Wirtschaftswissenschaften der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>

Krempkow, R./ Popp, J. (2001): Kontinuität und Veränderung in über 40 Jahren Elektrotechnik – Ein Vergleich von Absolventenbefragungen der 50er und 90er Jahre. Zusatzauswertung zur Dresdner Absolventenstudie 2000: Elektrotechnik, TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>, Dresden

Krempkow, R. (2001): Abschlußbericht. Befragung von Lehrenden an der TU Dresden zu Arbeitssituation, Fortbildungsinteressen und zur Akzeptanz hochschulpolitischer Reformvorhaben, Institut für Soziologie TU Dresden (Hg.),

Krempkow, R./ König, K./ Winter, J. (2001): Studienführer Sachsen: Wirtschafts-/Rechts-/Geistes- und Sozial- sowie Ingenieurwissenschaften und Informatik - Gegenüberstellung von Daten aus den Lehrberichten der Hochschulen und ausgewählten Ergebnissen studentischer Lehrbewertungen, TU Dresden, Institut für Soziologie

Krempkow, R. (2002): Befragung von Lehrenden an der Europa-Universität Viadrina Frankfurt/ Oder zu Arbeitssituation, Fortbildungsinteressen und zur Akzeptanz hochschulpolitischer Reformvorhaben, Kurzzusammenfassung der Ergebnisse, Institut für Soziologie TU Dresden

Krempkow, R./ Popp, J. (2002):, Dresdner Absolventenstudien 2001: Geowissenschaften. Abschlußbericht zur Befragung der Absolventen der Fachrichtung Geowissenschaften der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>, Dresden

Krempkow, R./ Fückler, M. (2002):, Dresdner Absolventenstudien 2001: Verkehrswissenschaften. Abschlußbericht zur Befragung der Absolventen der Fakultät Verkehrswissenschaften der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>, Dresden

Krempkow, R./ Brüggemann, H. (2002):, Dresdner Absolventenstudien 2001: Architektur. Abschlußbericht zur Befragung der Absolventen der Fakultät Architektur der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>, Dresden

Krempkow, R./ Popp, J. (2002):, Dresdner Absolventenstudien 2001: Wasserwesen. Abschlußbericht zur Befragung der Absolventen der Fachrichtung Wasserwesen der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, auch im Internet: <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>, Dresden

Krempkow, R. (2002): "Die Einschätzung der Studienbedingungen durch Studierende am Institut für Soziologie der TU Dresden 2002 und 1999", Institut für Soziologie TU Dresden

Krempkow, R./ König, K (2002): Studienführer Sachsen: Mathematik/ Naturwissenschaften und Medizin, Ingenieurwissenschaften und Informatik sowie Wirtschafts-/Rechts-/Geistes- und Sozialwissenschaften - Gegenüberstellung von Daten aus den Lehrberichten der Hochschulen und ausgewählten Ergebnissen studentischer Lehrbewertungen, TU Dresden, Institut für Soziologie

Krempkow, R. (2002): Sonderauswertung zum Studienführer Sachsen: Mathematik/ Naturwissenschaften und Medizin, Ingenieurwissenschaften und Informatik, Wirtschafts- und Rechtswissenschaften, Geistes- und Sozialwissenschaften – Zusammenhangsanalysen von Daten aus den Lehrberichten der Hochschulen und Ergebnissen von Studentenforschungen, TU Dresden, Institut für Soziologie

Krempkow, R. (2003a): Abschlußbericht zur Lehrendenbefragung an der Hochschule Zittau-Görlitz, Befragung der Lehrenden zu Arbeitssituation, Fortbildungsinteressen und zur Akzeptanz hochschulpolitischer Reformvorhaben, Institut für Soziologie TU Dresden

Krempkow, R. (2003b): Bessere Didaktik trotz schlechterer Rahmenbedingungen - Ergebnisse zweier Befragungen zur Situation von Lehre und Studium an der TU Dresden. In: Zeitschrift für Evaluation Nr. 2/ 2003: 257-278, Leske+Budrich, Saarbrücken, <http://www.zfev.de/>

Krempkow, R. (2003c): Evaluation und Professorenbesoldung. Vortrag im Rahmen der Ringvorlesung zur Hochschulpolitik an der TU Dresden, Integrale – Institut für studium generale, Dresden

Krempkow, R./ Puschmann, A./ Fücker, M. (2003): Dresdner Absolventenstudien 2002: Bauingenieurwesen. Abschlußbericht zur Befragung der Absolventen der Fakultät Bauingenieurwesen der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R./ Dettmer, A./ Eberhardt, J./ Schmidt, S. (2003): Dresdner Absolventenstudien 2002: Medizin. Abschlußbericht zur Befragung der Absolventen der Medizinischen Fakultät der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R./ König, K. (2003): Studienführer Sachsen 2003: Mathematik/ Naturwissenschaften und Medizin, Ingenieurwissenschaften und Informatik sowie Wirtschafts-/Rechts-/Geistes- und Sozialwissenschaften - Gegenüberstellung von Daten aus den Lehrberichten der Hochschulen und ausgewählten Ergebnissen studentischer Lehrbewertungen, TU Dresden, Institut für Soziologie

Krempkow, R./ Kühne, A./ Reiche, C. (2003): Dresdner Absolventenstudien 2002: Mathematik/ Naturwissenschaften. Abschlußbericht zur Befragung der Absolventen der Fakultät Mathematik/ Naturwissenschaften der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R. /Pastohr, M. (2003): Hochschulbindung an der TU Dresden. Bindungspotential, Weiterbildungsinteressen und Versuch einer Typologisierung – Eine Sonderauswertung der Dresdner Absolventenstudien 2000 - 2002. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R. /Popp, J. (2003a): Was macht Hochschulabsolventen erfolgreich? Analyse der Bestimmungsgründe beruflichen Erfolges anhand der Dresdner Absolventenstudien 2000 - 2002. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R. /Popp, J. (2003b): Tabellenband. Sonderauswertung der Dresdner Absolventenstudien 2000 - 2002. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R./ Popp, J./ Puschmann, A. (2004): Dresdner Absolventenstudien 2003: Informatik. Abschlußbericht zur Befragung der Absolventen der Fakultät Informatik der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html> ,

Krempkow, R./ Heidemann, L./ Ellwardt, L./ Eberhardt, J. (2004): Dresdner Absolventenstudien 2003: Maschinenwesen. Abschlußbericht zur Befragung der Absolventen der Fakultät Maschinenwesen der TU Dresden zum beruflichen Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R./ Dettmer, A./ Reiche, C. / Schmidt, S. (2004): Dresdner Absolventenstudien 2003: Jura. Abschlußbericht zur Befragung der Absolventen der Juristischen Fakultät der TU Dresden zum beruflichen

Verbleib und zur retrospektiven Bewertung der Studienqualität. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie, <http://www.tu-dresden.de/phfis/lenz/fo/Absolventenstudie.html>,

Krempkow, R. /Pastohr, J. (2004): Was macht Hochschulabsolventen erfolgreich? Analyse der Bestimmungsgründe beruflichen Erfolges anhand der Dresdner Absolventenstudien 2000 - 2003. TU Dresden, Institut für Soziologie, Lehrstuhl für Mikrosoziologie (unveröffentlichtes Manuskript)

Krempkow, R. (2004): Leistungsbewertung und Hochschulsteuerung,. In: Fröhler, N./ Hürtgen, S./ Schlüter, C./ Tiedke, M. (2004) (Hg.): Dokumentation der Promovierendenkonferenz 2002 der Hans-Böckler-Stiftung, Düsseldorf, S. 99-108

Krempkow, R. (2004): Fachliche und fachunabhängige Kompetenzen in den Dresdner Absolventenstudien 2000-2003. Langfassung des Vortrages am 17.06.2004 in der Tagungsdokumentation des Projektes der Bund-Länder-Kommission „Leistungspunkte“, Hochschule Zittau-Görlitz

Kriz, J (1995): Die Wirklichkeit von (Vor-)Urteilen. Über die inhaltlichen und methodischen Hintergründe der Stern-Image-Analyse. In: Mohler, H. (1995): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung. Waxmann Verlag, Münster , S. 11-28

Kromrey, H. (1993): Studentische ‚Lehrevaluation‘ oder (nur) Teilnehmerbefragungen in Lehrveranstaltungen? - Methodische Probleme bei der Bewertung von Lehrqualität. In: Gralki u.a. (1993): Evaluation schafft Autonomie. Berlin

Kromrey, H. (1995): Evaluation. Empirische Konzepte zur Bewertung von Handlungsprogrammen und die Schwierigkeiten ihrer Realisierung. In: Zeitschrift für Sozialisationsforschung und Erziehungssoziologie (ZSE) 15/1995: 313 - 336

Kromrey, H. (1996a): Qualitätsverbesserung in Lehre und Studium statt sogenannter Lehrevaluation. In: Zeitschrift für Pädagogische Psychologie 10 (3/4) 1996, S. 153 - 166, (Vgl. auch ebenda: Kromrey, H.: „Gute oder schlechte Sozialforschung“ S. 171 - 173)

Kromrey, H. (1996b): Von der Lehrevaluation zur Entwicklung von Lehrqualität - Eine Bilanz bisheriger Evaluationsverfahren und ein Ausblick. Langfassung des Vortrags auf dem 28. Kongreß der Deutschen Gesellschaft für Soziologie in Dresden am 9. 10. 1996, zu beziehen über: FU Berlin, Institut für Soziologie

Kromrey, H. (1998): Empirische Sozialforschung: Modelle und Methoden der Datenerhebung und Datenauswertung, Leske+Budrich, Opladen

Kromrey, H. (2001): Evaluation – ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. In: Sozialwissenschaften und Berufspraxis (SuB) Nr. 2/ 2001, S. 105-132

Kromrey, H. (2002): Empirische Sozialforschung: Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung, Leske+Budrich, Opladen

Kultusministerkonferenz (KMK) (1996): Beschluß zur leistungs- und belastungsbezogenen Verteilung der Haushaltsmittel vom 26.01.1996, In: Forschung und Lehre, DHV (Hg.), Heft 8/ 1996, S. 426-428

Kundelbusch, A. (2001): Variable Vergütung bedeutet Wettbewerb und Risiko. In: Die Mitbestimmung, Heft 6/2001, Hans-Böckler-Stiftung (Hg.), Düsseldorf, S.21-25

Lames, M. (2000): Evaluationswissenschaft in methodischer Rekonstruktion. In: Bank, V./ Lames, M. (2000): Über Evaluation. BajOsch-Hein Verlag für Berufs- und Wirtschaftspädagogik, Kiel, S. 8-48

Lamnek, S. (Hrsg.), (1993): Soziologie als Beruf in Europa. Ausbildung und Professionalisierung von Soziologinnen und Soziologen im europäischen Vergleich. edition sigma, Berlin

Lamnek, S. (1995): Qualitative Sozialforschung. Band 2: Methoden und Techniken. Weinheim, München

- Langer, M/ Ziegele, F./ Hennig-Thurau, T. (2001): Hochschulbindung – Entwicklung eines theoretischen Modells, empirische Überprüfung und Ableitung von Handlungsempfehlungen für die Hochschulpraxis, Abschlußbericht zum Kooperationsprojekt Hochschulbindung, Universität Hannover, Lehrstuhl Marketing I und CHE - Centrum für Hochschulentwicklung (Hg.)
- Leitow, B. (1996): Studentische Haltungen zur Studieneffizienz - ein Beitrag zur Typologie studentischer Orientierungen. In: Hefte zur Bildungs- und Hochschulforschung, Bd. 18, hg. von der Universität Konstanz
- Lemke, W. (2000): Leistungen der Hochschullehrer sind unbestritten. In: Forschung und Lehre, DHV (Hg.), Ausgabe April 2000, S. 172-174
- Lenz, K./ Wolter, A./ Winter, J. (2000): Das Image der Technischen Universität Dresden. Teil 1: Die Technische Universität Dresden aus Sicht ihrer Angehörigen, Abschlußbericht. TU Dresden, Institut für Soziologie, Institut für Allgemeine Erziehungswissenschaft
- Lerchenmüller, M. (2002): Die Balanced Scorecard als Instrument zur Steuerung von Hochschulen. In: VHW-Mitteilungen 1/2002, S. 4-10
- Leszczensky, M./ Orr, D. (2003): Paradigmenwechsel in der Hochschulfinanzierung (Leszczensky); Verfahren der Forschungsbewertung im Kontext neuer Steuerungsverfahren im Hochschulwesen: Analyse von vier Verfahren aus Niedersachsen, Großbritannien, den Niederlanden und Irland (Orr). In: HIS-Kurzinformation A1/2003, HIS GmbH, Hannover
- Leszczensky, M./ Orr, D. (2004): Staatliche Hochschulfinanzierung durch indikatorgestützte Mittelverteilung, Dokumentation und Analyse der Verfahren in 11 Bundesländern. In: HIS-Kurzinformation A2/2004, HIS GmbH, Hannover
- Leszczensky, M./ Orr, D./ Schwarzenberger, A./ Weitz, B. (2004): Staatliche Hochschulsteuerung durch Budgetierung und Qualitätssicherung: Ausgewählte OECD-Länder im Vergleich. Projektbericht, HIS GmbH, Hannover
- Lichka, C./ Kühn, H./ Karagiannis, D. (2002): IT-gestützte Balanced Scorecard. In: WISU-Kompakt Nr. 7/2002, Lange, Düsseldorf
- Luhmann, N. (1992): Universität als Milieu. Haux, Bielefeld
- Malorny, C. (1997): Der Weg zum umfassenden Qualitätsmanagement. In: „Spektrum der Wissenschaft“ Januar 1997, S. 96-105
- Marsh, H. W. (1982a): Effects of Expressivness, Content Coverage and Incentive on Multidimensional Student Rating Scales: New Interpretation of the Dr. Fox Effect. In: Journal of educational psychology Nr.74 (1), 1982: S. 126- 134
- Marsh, H. W. (1982b): Validity of students' evaluations of college teaching - a multitrait multimethod analysis. In: Journal of educational psychology Nr.74(2), (1982) S. 264 - 279
- Marsh, H. W. (1983): Multidimensional Ratings of Teaching Effectiveness by Students From Different Academic Settings and Their Relation to Student/ Course/ Instructor Characteristics. In: Journal of Educational Psychology, 1983, Vol. 76, No. 1 pp. 150- 1666
- Marsh, Herbert W. (1984): Students' Evaluations of University teaching: Dimensionality, Reliability, Validity, Potential Biases, and Utility. In: Journal of Educational Psychology, 1984, 76, Vol. 5, pp. 707-754
- Marsh, H.W. & Hocevar, D. (1991): Students evaluation of teaching effectiveness: the stability of of mean ratings of the same teachers over a 13-year period. In: Teaching & Teacher Education 7 (4), 1991: 303-314
- Marsh, H. W. & Roche, L. A.(1997): Making Students Evaluations of Teaching Effectiveness Effective. In: American Psychologist, 11/ 1997: 1187-1197

- Maul, K.-H. (2000): Wissensbilanzen als Teil des handelsrechtlichen Abschlusses. In: DStR 47/2000: 2009-2016
- Mayring, P. (1995): Qualitative Inhaltsanalyse: Grundlagen und Techniken. 5. Aufl. Weinheim: Dt. Studien-Verlag
- McKeachie, W. (1999): Teaching tips - A guidebook for the beginning college teacher, Lexington, D.C. Heath & Company
- Meer, D. (1999): Ein institutionelles Ritual. Beobachtungen zu mündlichen Abschlussprüfungen an Hochschulen. In: Forschung und Lehre, DHV (Hg.), Ausgabe September 1999: 478-480
- Meinefeld, W. (2000): Hochschulranking. Eine unsichere Basis für Entscheidungen. In: Forschung und Lehre, DHV (Hg.), Ausgabe Januar 2000: 26
- Meinefeld, W. (2002): [www.spaeter-mal-taxifahrer.de](http://www.spaeter-mal-taxifahrer.de)? Eine Untersuchung der beruflichen Situation Erlanger Soziologie-Absolventen. In: Sozialwissenschaften und Berufspraxis (SuB), Heft 1-2/2002, Leske+Budrich, S. 59-83
- Mende, M./ Stier, S. (2002): Den Vertrieb steuern mit der Balanced Scorecard. Wie das Versandhaus Quelle mit diesem Managementinstrument einen Wandel im Denken und Handeln seiner Vertriebspartner erreichte. In: Harvard Business Manager 2 /2002, S. 96-107
- Merkens, H. (2003): Stichproben bei qualitativen Studien. In: Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München
- Meulemann, H. (1995): Die Geschichte einer Jugend. Lebenserfolg und Erfolgsdeutung ehemaliger Gymnasiasten zwischen dem 15. und 30. Lebensjahr. Westdeutscher Verlag, Opladen
- Meuser, M./ Nagel, U. (1994): Expertenwissen und Experteninterview. In: Hitzler, R./ Honer, A./ Maeder, C. (Hg.): Expertenwissen. Die institutionalisierte Kompetenz zur Konstruktion von Wirklichkeit. Westdeutscher Verlag, Opladen 1994: 180-192
- Meuser, M./ Nagel, U. (2003): Das Experteninterview – Wissenssoziologische Voraussetzungen und methodische Durchführung. In: Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München, S. 481-491
- Meyer, H.-J. (1999): Mehr Leistungsanreize bieten In: Forschung & Lehre Nr. 11/1999, S. 636
- Miller, R. I. (1999): Major American Higher Education Issues and Challenges in the 21<sup>st</sup>. Jessica Kingsley Publishers, London
- Minssen, H./ Wilkesmann, U. (2003): Folgen der indikatorisierten Mittelverteilung in nordrhein-westfälischen Hochschulen. In: Beiträge zur Hochschulforschung Nr. 3/ 2003: S. 106-129
- Mohler, H. (1995): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung. Waxmann Verlag, Münster
- Morse, J. M. (1994): Designing Funded Qualitative Research. In: Denzin, N.K./ Lincoln Y. S.: Handbook of Qualitative Research, Thousand Oaks, S. 220-235
- Multrus, F. (1995): Zur Lehr- und Studienqualität. Dimensionen, Skalen und Befunde des Studierendensurveys. In: Hefte zur Bildungs- und Hochschulforschung (12), Arbeitsgruppe Hochschulforschung, Sozialwissenschaftliche Fakultät, Universität Konstanz

- Multrus, F. (2001): Skalenentwicklung zur Messung der Lehr- und Studienqualität. Vorgehensweise zur Identifizierung von Dimensionen und Skalen auf der Basis des Studierendensurveys im WS 2000/01. In: Hefte zur Bildungs- und Hochschulforschung (36), Arbeitsgruppe Hochschulforschung, Sozialwissenschaftliche Fakultät, Universität Konstanz
- Murray, H.G. & Lawrence, Ch.: Speech and drama training for Lectures as a means of Improving University Teaching. In: Research in Higher Education , 13 (1), 1980: 73-90
- Müller, U./ Ziegele, F. (2003): Standardisierung und Umsetzung der Berichtspflichten im Rahmen der Zielvereinbarungen in Nordrhein-Westfalen, CHE-Arbeitspapier Nr. 49, Gütersloh
- Müller-Böling, D. (1994): Leistungsbemessung - Leistungstransparenz -Leistungsfolgen - 7 Thesen Mai 1994, zwischenzeitlich veröffentlicht in: Hochschulen im Wettbewerb, Jahresversammlung 1994 der Hochschulrektorenkonferenz, Ansprachen und Diskussionen, Halle, 5.-7. Mai 1994, Dokumente zur Hochschulreform 96/1994, S. 49-63
- Müller-Böling, D. (Hrsg.) (1995): Qualitätssicherung in Hochschulen. Forschung - Lehre - Management. Eröffnungsveranstaltung des CHE Centrum für Hochschulentwicklung am 25. / 26. Januar 1995. Verlag Bertelsmann Stiftung, Gütersloh 1995: 119-198
- Müller-Böling, D. (1996): Evaluationen zur Rechenschaftslegung oder Qualitätsverbesserung. Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen, CHE (Hg.), Bielefeld
- Müller-Böling, D. (2000): Die entfesselte Hochschule, Verlag Bertelsmann Stiftung, Gütersloh
- Müller-Böling, D./ Hornbostel, S. (2000): Fehlinterpretationen und Vorurteile. Vom Umgang mit Hochschulrankings und deren Nutzen. In: Forschung und Lehre, DHV (Hg.), Ausgabe Februar 2000
- Müller-Böling, D./ Hornbostel, S./ Giebisch, P./ Buhr, P. (2000): Der Studienführer 2000: Vorgehensweise und Indikatoren CHE (Hg.), Arbeitspapier Nr. 22, Gütersloh
- Müller-Böling, D./ Hornbostel, St./ Berghoff, S. (2001): Hochschulranking. Aussagefähigkeit, Methoden, Probleme. Dokumentation der Wissenschaftlichen Tagung des Centrum für Hochschulentwicklung (CHE) am 25./26. September 2000 in Berlin, Verlag Bertelsmann Stiftung, Gütersloh
- Müller-Böling, D. (2001): Hochschulrankings wichtige Orientierungshilfe für Abiturienten, CHE-Pressemitteilung vom 5. 9. 2001, Centrum für Hochschulentwicklung CHE, Gütersloh
- Müller-Böling, D. (2004): Ranking statt Akkreditierung? Pro. In: Forschung und Lehre, Ausgabe 10/2004, DHV (Hg.), S. 548
- MWK/ Arnold: (undatierte Arbeitsanweisung): Hessisches Ministerium für Wissenschaft und Kunst , Referat I/7, Bearbeiter: Arnold, Erläuterung der Schwundberechnung nach dem Hamburger Verfahren, Wiesbaden
- Nießen, M. (1977): Gruppendiskussion. Interpretative Methodologie, Methodenbegründung, Anwendung. Fink, München
- Noelle-Neumann, E./ Peterson, T. (1996): Alle, nicht jeder - Einführung in die Methoden der Demoskopie. Deutscher Taschenbuch Verlag, München
- OECD (1998): University Research in Transition. Paris
- Olbertz, J.-H./ Otto, H.-U. (Hrsg.) (2001): Qualität von Bildung. Vier Perspektiven. In: Arbeitsberichte des Institutes für Hochschulforschung (HoF) Wittenberg
- Orr, D. J. (2001): Die Finanzierungsmethodik im englischen Universitätssektor: eine verfahrensanalytische Untersuchung ihrer Implikationen und Folgen, Inaugural Dissertation zur Erlangung des Grades eines Doktors der Philosophie an der Fakultät Erziehungswissenschaften der TU Dresden

- Palandt (2000): Bürgerliches Gesetzbuch, Verlag C.H. Beck, München
- Pasternack, P. (Hrsg.) (2001): Flexibilisierung der Hochschulhaushalte. Handbuch für Personalräte und Gremienmitglieder. Schüren-Verlag, Marburg
- Pasternack, P. (2004): Qualität an Hochschulen. In: Arbeitsberichte des Institutes für Hochschulforschung (HoF) Wittenberg
- Pastohr, M./ Wolter, A. (2003): Die Entwicklung der Studiennachfrage in den Ingenieurwissenschaften. Eine vergleichende Analyse von Entwicklungstrends beim Ingenieurwachstum im Freistaat Sachsen und in Deutschland. Arbeitsgruppe Standortprognose, Technische Universität Dresden
- Pawlowski, P. (1999): Absolventenbefragung 1998. Ein Projekt der Stipendiatengruppe Chemnitz, Stiftung der Deutschen Wirtschaft, TU Chemnitz, Fakultät für Wirtschaftswissenschaften, Professur BWL VI: Personal und Führung, Chemnitz
- Pellert, A. (1999): Die Universität als Organisation: Die Kunst, Experten zu managen. Studien zur Politik und Verwaltung; Bd. 67. Böhlau Verlag, Wien - Köln - Graz
- Pfennig, U./ Renn, O./ Mack, U. (2002): Zur Zukunft technischer und naturwissenschaftlicher Berufe. Strategien gegen den Nachwuchsmangel. Akademie für Technikfolgenabschätzung in Baden-Württemberg, Stuttgart
- Plath, H.-E.. (1995): Zum „Experteninterview“ - Fragen und Anmerkungen zur Diskussion. In: Brinkmann, C./ Deeke, A./ Völkel, B.: Experteninterviews in der Arbeitsmarktforschung. Diskussionsbeiträge zu methodischen Fragen und praktischen Erfahrungen. Beiträge zur Arbeitsmarkt- und Berufsforschung 191, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. Nürnberg.
- Plicht, H./ Schreyer, F. (2002): Methodische Probleme der Erfassung von Adäquanz der Akademikerbeschäftigung. In: Kleinhenz, G. (Hrsg.): IAB-Kompendium Arbeitsmarkt- und Berufsforschung. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesanstalt für Arbeit (Beiträge zur Arbeitsmarkt- und Berufsforschung BeitrAB250). S. 531 – 545
- Pohlenz, P. (2002): Exmatrikulierte der Universität Potsdam im Wintersemester 2001/02. Befragung zu studentischen Qualitätsurteilen über Lehre und Studium. Servicestelle für Lehrevaluation, Universität Potsdam
- Pohlenz, P. (2003): Exmatrikulierte der Universität Potsdam im Wintersemester 2001/02. Befragung zu studentischen Qualitätsurteilen über Lehre und Studium. Teil II. Servicestelle für Lehrevaluation, Universität Potsdam
- Pötschke, M. (2004): Akzeptanz hochschuldidaktischer Weiterbildung. Ergebnisse einer empirischen Studie an der Universität Bremen. In: Das Hochschulwesen Nr. 3/2004: 94-99, Luchterhand Verlag, Neuwied
- Rammelt, K. (2003): Controlling an der TU Dresden – Erfahrungen aus 10 Jahren. Tagungsbeitrag zum 2. Controllertag in Görlitz, Hochschule Zittau/ Görlitz
- Rawls, J. (2003): Gerechtigkeit als Fairneß. Ein Neuentwurf, (Hrsg. von Erin Kelly. Aus dem Amerikanischen von Joachim Schulte), Suhrkamp, Frankfurt a. M.
- Reuband , K.-H. (2000): Telefonische und postalische Befragungen in Ostdeutschland. Auswirkungen auf das Antwortverhalten. In: Hüfken, V.: Methoden in Telefonbefragungen. Westdeutscher Verlag, Wiesbaden
- Reuband, K.-H. (2001): Möglichkeiten und Probleme des Einsatzes postalischer Befragungen. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie (KZfSS), Jg. 53, Heft 2, S. 307-333
- Richter, R. (Hrsg.). (1994): "Qualitätssorge in der Lehre. Leitfaden für die studentische Lehrevaluation" Neuwied, Luchterhand

- Rindermann, H. & Amelang, M. (1994): "Das Heidelberger Inventar zur Lehrveranstaltungsevaluation" (HILVE), Handanweisung, Heidelberg, Asanger
- Rindermann, H. (1996): Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen anhand des Heidelberger Inventars zur Lehrveranstaltungsevaluation, Landau, Verlag empirische Pädagogik
- Rindermann, H. (1997): Die studentische Beurteilung von Lehrveranstaltungen: Forschungsstand und Implikationen für den Einsatz von Lehrevaluationen. In: R. S. Jäger, R. H. Lehmann und G. Trost (Hg.): Tests und Trends (Jahrbuch der Pädagogischen Diagnostik) 11, 1997: 12-53
- Rindermann, H. (1997): Lehrveranstaltungsbewertungen: Erfahrungen und ihre Nutzung für die Fakultätsentwicklung. In: H. Altrichter, M. Schratz & H. Pechar (Hg.) "Hochschulen auf dem Prüfstand" Studienverlag, Innsbruck 1997: 179-196
- Rindermann, H. (1998): Das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität: Entwicklung, Begründung und Überprüfung. In: Beiträge zur Hochschulforschung (hg. vom Bayrischen Staatsinstitut für Hochschulforschung und Hochschulplanung) 3/1998: 189-224
- Rindermann, H. (1999): Was zeichnet gute Lehre aus? Ergebnisse einer offenen Befragung von Studierenden und Lehrenden. In: Zeitschrift für Hochschuldidaktik (ZSfHD) 1/99: 136-155
- Rindermann, H. (2000): Das Selbstobjektivierungsproblem im akademischen Milieu, In: Das Hochschulwesen, 48. Jahrgang, Neuwied, Luchterhand Verlag, 3/ 2000: 74-82 sowie 4/ 2000: S. 117-123
- Rindermann, H. (2001): Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierter Unterrichts. Landau: Empirische Pädagogik.
- Rindermann, H./ Kohler, J. (2003): Lässt sich die Lehrqualität durch Evaluation und Beratung verbessern? Überprüfung eines Evaluations-Beratungs-Modells. Psychologie in Erziehung und Unterricht, 50 (1), 71-85
- Rindermann, H. (2003): Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. In: Zeitschrift für Evaluation Nr. 2/ 2003: 233-256, Leske+Budrich, Saarbrücken, <http://www.zfev.de/>
- Rindermann, H. (2004): Konsequenzen aus der studentischen Veranstaltungskritik. In: Hochschulrektorenkonferenz (Hg.) (2004): Evaluation und Akkreditierung, Beiträge zur Hochschulpolitik 1/2004, Bonn, S. 83-96
- Ronge, V. (2000): Zielvereinbarungen: Contra. In: Forschung und Lehre, DHV (Hg.), Ausgabe März 2000. S. 189
- Rostampour, P./ Lemberg, A. (2003): Berufserfolg. Objektive und subjektive Dimensionen. In: Krüger, H.-H. u.a.: Diplom-Pädagogen in Deutschland. Survey 2001. Juventa Verlag, Weinheim und München
- Sauter, E. (1995): Ansätze der Qualitätssicherung im Überblick – Vergleichende Betrachtung. In: Gnahn, D./ Krekel, E./ Wolter, A.: Qualitätsmanagement im Bildungswesen, Schriften des Internationalen Begegnungszentrums Sankt Marienthal (IBZ), Band 2, Hannover/ Berlin/ Dresden, S. 135-148
- Schaeper, H. (1995): Zur Arbeitssituation von Lehrenden an westdeutschen Universitäten. Ergebnisse einer empirischen Untersuchung in fünf ausgewählten Disziplinen. In: Enders/Teichler (Hrsg.) (1995): "Der Hochschullehrerberuf" Aktuelle Studien und ihre hochschulpolitische Diskussion, Neuwied, Luchterhand
- Schaeper, H./ Briedis, K. (2004): Kompetenzen von Hochschulabsolventinnen und Hochschulabsolventen, berufliche Anforderungen und Folgerungen für die Hochschulreform. Projektbericht, HIS GmbH, Hannover

- Schenker-Wicki, A. (1996): Evaluation von Hochschulleistungen - Leistungsindikatoren und Performance Measurements. Wiesbaden: Deutscher Universitäts- Verlag
- Schenker-Wicki, A. (2002): Finanzierungs- und Steuerungssysteme der universitären Hochschulen in der Schweiz. In: Beiträge zur Hochschulforschung Nr. 4/ 2002, S. 18-39
- Schmidberger, J. (1994): Controlling für Öffentliche Verwaltungen. Funktionen – Aufgabenfelder – Instrumente, Gabler, Wiesbaden
- Schmidt, C. (2003): „Am Material“: Auswertungstechniken für Leitfadeninterviews. In: Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München
- Schmidt; D. (1992): Studienzeitverkürzung in Baden-Württemberg. In: 11. GEW-Sommerschule "Qualität von Studium und Lehre - Gespräche über eine vernachlässigte Aufgabe der Hochschulen" Materialien und Dokumente Hochschule und Forschung, Bd. 66, GEW (Hg.) 1992: 109-112
- Schmitz, W. (Hg) (1995): Evaluation der Lehre - ein Kolloquium an der TU Dresden. Dresdner Universitätverlag
- Schneekloth, U./ Leven, I. (2003): Woran bemisst sich eine „gute“ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu de Hintergründen von Nonresponse in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS. In: ZUMA-Nachrichten, November 2003, S. 16-57
- Schnell, R./ Hill, P. B./ Esser, E. (1999): Methoden empirischer Sozialforschung. Oldenburg Verlag, München und Wien
- Schreiber, J. (1994): Der Arbeitsmarkt für Soziologen, In: Sozialwissenschaften und Berufspraxis, Heft Nr. 4/1994
- Schreyer, F. (2001): Grundzüge des Akademikerarbeitsmarktes. Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesanstalt für Arbeit (ibv Nr. 26), Nürnberg
- Schröder, T. (2004): Der Einsatz leistungsorientierter Ressourcensteuerungsverfahren im deutschen Hochschulsystem. Eine empirische Untersuchung ihrer Ausgestaltung und Wirkungsweise. In: Beiträge zur Hochschulforschung, Bayrisches Staatsinstitut für Hochschulforschung und Hochschulplanung (Hg.), München, Heft 2/2004, S. 28-59
- Schründer-Lenzen, A. (2003): Triangulation und idealtypisches Verstehen in der Rekonstruktion subjektiver Theorien. In: Friebertshäuser, B./ Prengel, A. (2003): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Juventa Verlag, Weinheim und München
- Schulz, R. (Hrsg.) (1996): Verbesserung von Studium und Lehre (Arbeitstagung am 26. 11. 1993 an der Universität Oldenburg) Bibliotheks- und Informationssystem der Universität Oldenburg
- Schwarzkopf, J. (2001): Studiengänge auf dem Prüfstand. Wie Studierende ihr Studium bewerten. (Diplomarbeit) Institut für Soziologie Universität Potsdam
- Scott, P. (1995): The Meanings of Mass Higher Education. The Society for Research into Higher Education & Open University Press
- Seeger, T./ Walter, M./ Liebe, R./ Ebert, G. (1998): Kosten-, Leistungsrechnung und Controlling. Ein Erfahrungsbericht für die Praxis über die Einführung der Standard-KLAR am Beispiel der Bundesverwaltung. R. v. Deckers, Heidelberg
- Seidenschwarz, B. (1992): Entwicklung eines Controllingkonzeptes für öffentliche Institutionen – dargestellt am Beispiel einer Universität, Vahlen, München

- Sendldorfer, K.R. (1997): Controlling in Unternehmen und öffentlichen Verwaltungen und Möglichkeiten der Übertragung auf die Universität. In: Franz, K-P./ Feser, H.-D./ Fahse, H./ Sendldorfer, K.R.: Arbeitspapiere zu Hochschulfragen und Hochschulcontrolling Nr. 1, Universität Kaiserslautern, Hochschulcontrolling
- Sims, S.J.& Sims, R.R. (1995): Total quality management in higher education: Is it working? Why or why not? edited by Serbrenia J. Sims and Ronald R. Sims, Praeger Publishers, Westport, Connecticut, London 1995
- Sixl, H. (1998): DPG-Berufsumfrage 1997. In: Physikalische Blätter, Juni 1998: 504 ff.
- Souvignier, E./ Gold, A. (2002): Fragebögen zur Lehrevaluation: Was können sie leisten? In: Zeitschrift für Evaluation Nr. 2/ 2002: 265-280, Leske+Budrich, Saarbrücken, <http://www.zfev.de/>
- Speckbacher, G./ Bischof, J. (2000): Die Balanced Scorecard als innovatives Managementsystem. Konzeptionelle Grundlage und Stand der Anwendung in deutschen Unternehmen. In: Die Betriebswirtschaft (DBW) Nr. 4/ 2000, S. 795-810
- Spöhring, W. (1989): Qualitative Sozialforschung. Teubner (Teubner-Studienskripten; 133), Stuttgart
- Spiel, C./ Gössler, M. P. (1999): Allgemeingültige Qualitätskriterien von Lehre – eine Fiktion? In: Zeitschrift für Hochschuldidaktik (ZSfHD) 2/99: 72- 95
- Spiel, C./ Gössler, M. P. (2000): Zum Einfluß von Biasvariablen auf die Bewertung universitärer Lehre durch Studierende. In: Zeitschrift für Pädagogische Psychologie, 14 (1), S. 38-47
- Spiel, C. (2001) (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck, Waxmann, Münster
- Spiel, C./ Wolf, P./ Popper, V. (2002): Lehre und Lehrevaluation – (un)geliebt?: Die Perspektive der Universitätslehrenden. In: Zeitschrift für Psychologie, 210, S. 27-39
- Spradley, J.P. (1979): The Ethnographic Interview. Holt, Rinehart Winston. New York
- Sprenger, R. K. (2000): Mythos Motivation. Wege aus einer Sackgasse. Campus Verlag, Frankfurt/ Main
- Stachowiak, H. (1973): Allgemeine Modelltheorie. Springer Verlag, Wien – New York
- Stachowiak, H. (1980) (Hrsg.): Modelle und Modelldenken im Unterricht. Anwendungen der Allgemeinen Modelltheorie auf die Unterrichtspraxis. Band 4 der Reihe Forschen und Lernen, Verlag Julius Klinkhardt, Bad Heilbrunn/ Obb.
- Stadler, M. (2003): Leistungsorientierte Besoldung von Hochschullehrern auf der Grundlage objektiv messbarer Kriterien? In: WiSt, Heft 6/2003, S. 334-339
- Stawicki, M. (1998): Ansätze zu einem Qualitätsmanagement in der Lehre. Grundsätzliche Überlegungen zu Möglichkeiten, Nutzen und Problemen. In: HRK (Hg.) (1998): Qualitätsmanagement in der Lehre, TQL 98. Tagung des hessischen Arbeitskreises "Qualitätsmanagement in der Lehre" und der Hochschulrektorenkonferenz. Wiesbaden, 4. Februar 1998. Beiträge zur Hochschulpolitik 5/1998. Bonn. S. 67-78
- Stengler, R. (1998): Aufbau eines Qualitätsmanagements für einen technischen Fachbereich. In: HRK (Hg.) (1998): Qualitätsmanagement in der Lehre, TQL 98. Tagung des hessischen Arbeitskreises "Qualitätsmanagement in der Lehre" und der Hochschulrektorenkonferenz. Wiesbaden, 4. Februar 1998. Beiträge zur Hochschulpolitik 5/1998. Bonn. S. 97-120
- Stockmann, R. (Hg.) (2000): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Leske + Budrich, Opladen
- Stockmann (2002a): Zur Notwendigkeit und Konzeption einer deutschsprachigen „Zeitschrift für Evaluation“. In: Zeitschrift für Evaluation Nr. 1, Leske + Budrich, Leverkusen, S. 1-9

- Stockmann (2002b): Qualitätsmanagement und Evaluation – Konkurrierende oder sich ergänzende Konzepte? In: Zeitschrift für Evaluation Nr. 2, Leske + Budrich, Leverkusen, S. 209- 244
- Strobel, A./ Westhoff, K. (2003): Rückmeldung zum freien wissenschaftlichen Vortrag. Ein erster Schritt zur systematischen Verbesserung der Lehre. In: Report Psychologie, 6, S. 372-378
- Strych, M./ Weiß, J. (1998): Meyers großes Taschenlexikon, B.I.-Taschenbuchverlag, Mannheim, Leipzig, Wien, Zürich
- Süllwold, Fritz (1997): Fehler und Fallen. In: Forschung & Lehre 4/97: 185-188
- Teichler, U. (1993): Absolventenbefragungen als Instrument für die Analyse der Wirkungen von Hochschulen” In: Altrichter/ Schratz (Hg.) Qualität von Universitäten, Studien Verlag, Innsbruck-Wien
- Teichler, U./ Schomburg, H. (1997): Evaluation von Hochschulen auf der Basis von Absolventenstudien” In: Altrichter/ Schratz/ Pechar (Hg.): Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen?, Studien Verlag, Innsbruck-Wien
- Teichler, U./ Daniel, H.-D./ Enders, J. (Hrsg.) (1998): Brennpunkt Hochschule. Neuere Analysen zu Hochschule, Beruf und Gesellschaft. Frankfurt a. M.; New York: Campus
- Teichler, U./ Schomburg, H. (2000): Studieren lohnt sich. Bericht über die Deutsche Teilstudie zum internationalen Forschungsprojekt ”Higher Education and Graduate Employment in Europe” des Wissenschaftlichen Zentrums für Berufs- und Hochschulforschung der Universität Gesamthochschule Kassel, In: Der Spiegel Nr. 46 / 2000, Seiten 54 bis 88
- Teichler, U./ Schomburg, H. u. a. (Hg.) (2001): Erfolgreich von der Uni in den Job. Regensburg/ Düsseldorf, Walhalla Verlag
- Teichler, U. (2002): Potentiale und Erträge von Absolventenstudien. In: Sozialwissenschaften und Berufspraxis (SuB), Heft 1-2/2002, Leske+Budrich, S. 9-32
- Teichler, U. (2003): Die Entstehung eines superkomplexen Systems der Qualitätsbewertung. Ein Beitrag aus Sicht der Hochschulforschung. In: hochschule innovativ, Ausgabe 9, Juni 2003, S. 5-6
- Thoma, M./ Zimmermann, M. (1996): Zum Einfluß der Befragungstechnik auf den Rücklauf bei schriftlichen Umfragen – Experimentelle Befunde zur ”Total-Design-Methode” In: ZUMA-Nachrichten, November 1996: 141-158
- Tinto, V. (1993): Leaving College: Rethinking the Causes and Cures of Student attrition, 2<sup>nd</sup> ed., Chicago, The University of Chicago Press
- Tondorf, K. (2003): Einführung leistungsbezogener Vergütung auf der Basis von Zielvereinbarungen. Arbeitspapier Nr. 64 der Hans-Böckler-Stiftung (Hg.), Düsseldorf
- Tondorf, K./ Jochmann-Döll, A. (2004): Monetäre Leistungsanreize im Öffentlichen Sektor. In: WSI-Mitteilungen Nr. 8/ 2004, Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung (Hg.), Düsseldorf, S. 428-434
- Tropp, G. (2002): Kennzahlensysteme des Hochschulcontrolling – Fundierung, Systematisierung, Anwendung. Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (Hg.), München
- TU Dresden (2001): Zielvereinbarung zur Einführung des Globalhaushaltes mit dem Sächsischen Staatsministerium für Wissenschaft und Kunst (SMWK), unveröffentlichtes Dokument, Dresden

- Universität Dortmund (2000): Ausbau der volumen- und leistungsorientierten Mittelverteilung unter Einbeziehung qualitativer Indikatoren (VLM und QPS), Beschluß des Senates der Universität Dortmund vom 6. 7. 2000 zum Projekt des Stifterverbandes für die deutsche Wissenschaft
- Universität Oldenburg (2000): Konzept zur leistungsorientierten Mittelvergabe, Oldenburg
- Vogel, B. (1995): „Wenn der Eisberg zu schmelzen beginnt...“ – Einige Reflexionen über den Stellenwert und die Probleme des Experteninterviews in der Praxis der empirischen Sozialforschung. In: Brinkmann, C./ Deeke, A./ Völkel, B.: Experteninterviews in der Arbeitsmarktforschung. Diskussionsbeiträge zu methodischen Fragen und praktischen Erfahrungen. Beiträge zur Arbeitsmarkt- und Berufsforschung 191, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. Nürnberg.
- Wagner, B. (1999): Studieren in Ostsachsen - Sonderauswertung der 15. Sozialerhebung des Deutschen Studentenwerkes für die Hochschulstandorte Dresden, Zittau und Görlitz, Studentenwerk Dresden (Hg.)
- Wagner, R. (2000): Zielvereinbarungen: Pro. In: Forschung und Lehre, DHV (Hg.), Ausgabe März 2000. S. 188
- Wahl, D. et al (1995): Erwachsenenbildung konkret - Mehrphasiges Dozententraining: Eine neue Form erwachsenendidaktischer Ausbildung von Referenten und Dozenten. Weinheim, Deutscher Studien Verlag,
- Wahrig, G. (2004) (Hrsg.): Brockhaus - Wahrig Deutsches Wörterbuch, Brockhaus, Deutsche Verlags-Anstalt, Wiesbaden
- Weber, J. (1996): Hochschulcontrolling. Das Modell WHU, Schäffer-Poeschel, Stuttgart
- Webler, W.-D.(1991): Kriterien für gute akademische Lehre. In: Das Hochschulwesen 6/ 1991: 243-249
- Webler, W.-D.(1992): Evaluation der Lehre - Zwischenbilanz einer unübersichtlichen Entwicklung. In: Das Hochschulwesen 1992: 153-173
- Webler, W.-D./ Domeyer, V./ Schieber, B. (1993): Lehrberichte. Empirische Grundlagen, Indikatorenauswahl und Empfehlungen zur Darstellung der Situation der Lehre in Lehrberichten. bmbf, Bad Honnef: Bock Verlag
- Webler, W.-D. (1995): Das Modell eines Lehrberichts über die Evaluation von Lehre und Studium und erste Ergebnisse. In: Das Hochschulwesen 4/1995: 258-266
- Webler, W.-D. (1996): Qualitätssicherung in Lehre und Studium an deutschen Hochschulen. In: Zeitschrift für Sozialisationsforschung und Erziehungssoziologie (ZSE) 1996: 119-148
- Weinert, A. B. (1998): Organisationspsychologie. Beltz, Psychologie Verlags Union, Weinheim
- Westermann, R. u.a. (1996): Identifikation und Erfassung von Komponenten der Studienzufriedenheit. In: Psychologie in Erziehung und Unterricht, 43. Jg. 1996: 1-22
- Weule, H. (1999): Blick aus der Praxis. Qualitätssicherung an Hochschulen. In: Forschung und Lehre, DHV (Hg.), Ausgabe Oktober 1999
- Winter, M. (1999): Engagierte Studis – zufriedene Profs? Resultate aus der Hochschullehrer-Befragung, In: Universitätszeitung der Martin-Luther-Universität Halle/ Wittenberg Nr. 13/1999: 3
- Winter, M. (2000): Evaluation und Qualitätsentwicklung von Lehre und Studium, In: Das Hochschulwesen, 48. Jahrgang, Nr. 6/ 2000: 185-191, Neuwied, Luchterhand Verlag,
- Wissenschaftsrat (1996): Empfehlungen zur Stärkung der Lehre in den Hochschulen durch Evaluation. Berlin: Wissenschaftsrat

- Wissenschaftsrat (2002): Die Entwicklung der Fachstudiedauer an Universitäten von 1990 bis 1998, im Internet unter: <http://www.wissenschaftsrat.de/texte/4770-01.pdf>
- Wissenschaftsrat (2003): Prüfungsnoten an Hochschulen 1996, 1998 und 2000 nach ausgewählten Studienbereichen und Studienfächern, Arbeitsbericht (Drs. 5526/03), im Internet unter: [http://www.wissenschaftsrat.de/texte/5526\\_03.pdf](http://www.wissenschaftsrat.de/texte/5526_03.pdf)
- Witte, J./ Schreiterer, U./ Müller-Böling, D. (2001): Gestaltungsfragen bei der Umsetzung des Professorenbesoldungsgesetzes, Arbeitspapier zur Dienstrechtsreform 1
- Witzel, A. (1995): Auswertung problemzentrierter Interviews: Grundlagen und Erfahrungen. In: Strobl, R./ Böttger, A. (Hrsg.) Wahre Geschichten, Baden-Baden, S. 49-75
- Wolter, A. (1995): Evaluierung der Hochschullehre - Mehr Staatskontrolle oder mehr Hochschulautonomie? In: Gnahn, D./ Krekel, E./ Wolter, A.: Qualitätsmanagement im Bildungswesen, Schriften des Internationalen Begegnungszentrums Sankt Marienthal (IBZ), Band 2, Hannover/ Berlin/ Dresden
- Wottawa, H./ Thierau, H. (1990): Lehrbuch Evaluation. Huber Verlag, Bern
- Wottawa, H./ Thierau, H. (1998): Lehrbuch Evaluation. 2., vollst. überarb. Auflage, Huber Verlag, Bern
- Wuppertaler Kreis e.V./ CERTQUA (Hrsg.) (2002): Qualitätsmanagement und Zertifizierung in der Weiterbildung nach dem internationalen Standard ISO 9000:2000, Luchterhand Verlag
- Ziegele, F. (1997): Untersuchungen des CHE zum Studienabbruch, Gütersloh, CHE (Hg.)
- Ziegele, F. (2002): Indikatoren für formelgebundene Finanzaufweisungen. CHE (Hg.) [www.evanes.de/infoboerse/dok/htm/CHE-Indikatoren.htm](http://www.evanes.de/infoboerse/dok/htm/CHE-Indikatoren.htm) (Stand 10.01.2002)
- Ziegele, F. (2003a): Das Berichtswesen im Verhältnis Hochschulen – Staat in Brandenburg: Soll-Konzept und Vorschläge der Weiterentwicklung, CHE-Arbeitspapier Nr. 47, Gütersloh
- Ziegele, F. (2003b): Konzeption eines neuen Modells der staatlichen Mittelvergabe an den Hochschulen in Brandenburg. Ergebnisbericht, CHE-Arbeitspapier Nr. 51, Gütersloh
- Ziegele, F./ Handel, K. (2004): Anreizsysteme im Hochschuleinsatz. Grundlagen – Chancen und Grenzen – Empfehlungen zu Auswahl und Konzeption. In: Benz, W./ Kohler, J./ Landfried, K. (Hrsg.): Handbuch Qualität in Studium und Lehre, Berlin, S. 1-22
- Zimmermann, M. (1999): Berufliche Eingliederung – Zur Entwicklung einer erziehungswissenschaftlichen Theorie des beruflichen Verbleibs. Rainer Hampp Verlag, München/ Mering
- Zimmermann, W./ Fries, H.-P./ Hoch, G. (2003): Betriebliches Rechnungswesen. Bilanz- und Erfolgsrechnung, Kosten- und Leistungsrechnung, Wirtschaftlichkeits- und Investitionsrechnung. 8. Auflage, Oldenbourg Verlag, München
- Zink, K.-J. (1998): Paper zum Vortrag auf der HRK-Tagung 28./29. 9. 1998 an der Universität Kaiserslautern (Tagungsband erschien als Band 1/1999 "Qualität an Hochschulen" in der Schriftenreihe der HRK)
- Zurwehme, A. (2000): Möglichkeiten der Steuerung beschäftigungsadäquater Qualifizierung im Lernort Schule. Überlegungen zum Einsatz einer Balanced Scorecard in beruflichen Schulen. Wissenschaftliche Begleitung des Projekts zur Stärkung der regionalen Leistungsfähigkeit bei der beschäftigungsadäquaten Qualifizierung junger Menschen in Sachsen (Regionalprojekt). In: Dresdner Beiträge zur Wirtschaftspädagogik Nr. 2/2000, Fakultät Wirtschaftswissenschaften TU Dresden (Hg.)

## **Publikationen von Massenmedien zu Hochschul-Rankings in Deutschland**

„Der Spiegel“ (1989/90): Untersuchung der Qualität des Lehrbetriebes durch Befragung von ca. 6000 Studierenden (12 pro Fachbereich) an 51 westdeutschen Universitäten. In: Spiegel Nr. 50/1989: 70-87 (zur Methode vgl. Spiegel special Nr. 3/ 1990 sowie Hornbostel & Daniel 1996)

„Der Stern“ (1993): Befragung von 700 Professoren zur Qualität der 15 größten Fachbereiche an 51 deutschen Hochschulen, Nr. 2/1993

„Der Spiegel“ (1993): „Wie gut sind unsere Hochschulen? Befragung von 1185 Professoren und 11828 Studierenden (18 pro Fachbereich) Spiegel special Nr. 3/1993

„Forbes“ (1993): Befragung von bundesweit 1650 Professoren zur Einschätzung der Lehrqualität „ihrer“ Fachbereiche, Nr. 6/1993: 39-44

„Focus“ (1993): Bewertung der Forschungsqualität deutscher Universitäten anhand der Anzahl von Publikationen ihrer Professoren in renommierten Fachzeitschriften, Nr. 39/1993: 128-140

„Focus“ (1997): Untersuchung von 1760 Professoren, ca. 26000 Studierenden, und zusätzlich von Personalchefs, außerdem Erhebung „objektiver“ hochschulstatistischer Daten (Einschätzung der 20 wichtigsten Fachbereiche (Jura, Wirtschaft, Medizin, Psychologie, Geschichte, Pädagogik, Geographie, Informatik, Naturwissenschaften, Ingenieurwissenschaften, Germanistik, Anglistik) erschien in einer Serie ab Nr. 16 (vom 14. 4. 97) und in den 9 folgenden Ausgaben)

„Der Spiegel“ (Nr. 19/1998: 94-117): „Uni-Test Europa“ erster europäübergreifender Uni-Test für die Studienfächer Jura, Wirtschaftswissenschaften, Ingenieurwissenschaften und Sprachwissenschaften auf der Basis einer Befragung von ca. 1000 Professoren und rund 7500 Studenten, wobei jeweils 50 Studierende pro Fachbereich befragt wurden;

Stiftung Warentest/ Centrum für Hochschulentwicklung (Hrsg.): „TEST“-Sonderheft: „Studienführer Chemie/ Wirtschaftswissenschaften“ 1998: 3, 13, 159-161; Erhoben wurden hochschulstatistische Daten sowie die Urteile von einer großen Anzahl Studierender und Professoren;

„Der Spiegel“ (Nr. 15/1999): Uni-Ranking für die beliebtesten Studienfächer Deutschlands ähnlich dem 1993 durchgeführten Ranking für die Fächer Jura, Wirtschaft, Anglistik, Germanistik, Geschichte, Erziehungswissenschaften, Informatik, Maschinenbau, Mathematik, Biologie, Medizin, Psychologie

Centrum für Hochschulentwicklung (CHE)/„Stern“/„Start“ 1999 (Hrsg.): „Start“-Sonderheft: Interaktiver Studienführer für Naturwissenschaften, Informatik, Jura und Wirtschaft: Aufbauend auf dem Studienführer von CHE/ Stiftung Warentest (1998) wurden weitere Kriterien hinzugezogen und zusätzlich per CD-ROM eine interaktive Suche nach den persönlichen Bedürfnissen am besten entsprechenden Hochschule ermöglicht.

Centrum für Hochschulentwicklung (CHE)/„Stern“/„Start“ 2000 (Hrsg.): „Start“-Sonderheft: Interaktiver Studienführer Ingenieurwissenschaften: Aufbauend auf dem Studienführer von CHE/ Stiftung Warentest (1999) wurden die Fächer Maschinenbau, Elektrotechnik, Bauingenieurwesen und Architektur einbezogen

Centrum für Hochschulentwicklung (CHE)/„Stern“/„Start“ 2001 (Hrsg.): „Start“-Sonderheft: „Studienführer 2001“: Aufbauend auf den Studienführern von 1998, 1999 und 2000 wurden die Fächer Anglistik/ Amerikanistik, Erziehungswissenschaft, Germanistik, Geschichte und Psychologie einbezogen.

„Focus“ (2002): Der aktuellste Hochschulführer. 1300 Studienangebote im Vergleich (Einschätzung der 20 meiststudierten Fächer, erschien in einer Serie ab Nr. 16 (vom 15. 4. 2002)

Centrum für Hochschulentwicklung (CHE)/„Stern“/„Start“ 2002 (Hrsg.): „Start“-Sonderheft: „Der Studienführer 2002“: Aufbauend auf den Studienführern von 1998 bis 2001 wurden die Fächer Soziologie/ Sozialwissenschaft, Politologie und Sozialwesen einbezogen sowie Jura und Wirtschaft aktualisiert.

Centrum für Hochschulentwicklung (CHE)/"Stern"/"Start" 2003 (Hrsg.): "Start"-Sonderheft: "Der Studienführer 2003": Aufbauend auf den Studienführern von 1998 bis 2002 wurden die Fächer Medizin, Pharmazie und Pflege einbezogen sowie Mathematik, Naturwissenschaften und Informatik aktualisiert.

Centrum für Hochschulentwicklung (CHE)/"Stern"/"Start" 2004 (Hrsg.): "Start"-Sonderheft: "Der Studienführer 2004": Aufbauend auf den Studienführern von 1998 bis 2003 wurden die Fächer Anglistik, Architektur, Bauingenieurwesen, Elektro- und Informationstechnik, Erziehungswissenschaft, Germanistik, Geschichte, Maschinenbau/ Verfahrenstechnik und Psychologie aktualisiert.

"Focus" (2004): Untersuchung "objektiver" hochschulstatistischer Daten, neu ist die Einbeziehung eines Zitationsindex (Einschätzung der 20 wichtigsten Fachbereiche, erschien in einer Serie ab Nr. 39 (vom 20.09.2004))

„Der Spiegel“ (2004): Spiegel-Rangliste: Wo studieren die Besten? Elite-Studenten an deutschen Unis. (Nr. 48 vom 22.11.2004, S. 178-191) Auswertung einer Online-Befragung von ca. 50.000 Studenten der 15 meist-studierten Fächer zwischen April und Juli 2004 zu ihren Werdegängen, Erfahrungen und Qualifikationen (u.a. Abitur- und Prüfungsnoten, Studiendauer und Alter, Praktika und Auslandsaufenthalte, Stipendien/ Preise und Veröffentlichungen, EDV- und Sprachkenntnisse, Berufserfahrung, Engagement außerhalb der Universität)

Centrum für Hochschulentwicklung (CHE)/ Die Zeit (Hrsg.): Studienführer 2005/06: Aufbauend auf den Studienführern von 1998 bis 2004 wurden die am häufigsten studierten Fächer der Rechts- Wirtschafts- und Sozialwissenschaften aktualisiert.

Hinzu kam in den letzten Jahren eine unüberschaubare Anzahl von Fachbereichs-Rankinglisten, die sich nur mit einer oder wenigen Fachrichtungen beschäftigten. Beispielhaft sollen hier deshalb nur einzelne genannt werden, wie das "Manager-Magazin": Imageanalyse von Wirtschaftsfakultäten aufgrund Beurteilung durch Manager und Führungskräfte der Freien Wirtschaft in Nr. 1/1995: 70-87 (Neuaufgabe 1997), "Capital": Vergleich von Wirtschaftsfakultäten hinsichtlich ihres Images bei Personalchefs, der Abschlußnoten der Absolventen, des Prüfungssystems, der Absolventenzahlen und der Studiendauer in Nr. 5/97: 204 ff, oder die "Wirtschaftswoche", die 1997 und 2004 ebenfalls Rankings bezogen auf das Image bei Personalchefs veröffentlichte.

## Anhang A: Tabellarische Übersicht prägnanter Textpassagen

Fragen	1. Fall: B
1 Grund- position	B nimmt keine explizite Selbsteinschätzung vor. Im Verlauf des Gespräches wird aber deutlich, daß B <b>grundsätzlich für eine stärkere Leistungsorientierung</b> ist. B würde jedoch <b>weitere Kriterien einbeziehen</b> und prozessbegleitend evaluieren: „Also wenn ich mich in die Situation der Studierenden begeben – und das wäre mein Ansatzpunkt – was ist denn für die eine Leistung, die was wert ist oder die Nutzen hat für sie?“ (9) „...ich denke, das ist für Studierende wesentlich wichtiger als diese anderen Sachen, die wir erheben.“ (10) „Ich denke Qualitätssicherung in den Studiengängen kann tatsächlich nur dadurch passieren, daß <b>prozessbegleitend evaluiert</b> wird – und nicht nur am Ende, wenn eine Lehrveranstaltungsevaluation gemacht wird. Es muß im Grunde eine Rückmeldekultur entwickelt werden. Und Professoren dürfen das auch nicht als Angriff auf ihre fachliche Kompetenz verstehen, wenn Studierende sagen: Nein, so nicht. Das verstehe ich nicht.“ (21)
2 geeignete Kriterien	B formuliert <b>Kriterien</b> als Fragen: Ist eine Lehrkraft „ <b>ansprechbar</b> “ d.h., „ob sie tatsächlich Ihre Sprechstunden abhalten“? (9) „Werden in den Veranstaltungen <b>die aktuellen Forschungsergebnisse</b> ausreichend diskutiert? Ist ein Professor, eine Professorin in der Lage, <b>verschiedene Arbeitsformen</b> einzusetzen? Wie werden Medien in der Lehre eingesetzt, (...) <b>didaktische Kompetenz</b> nachgewiesen?“ (10) ...“was in der Diskussion überhaupt noch nicht zum Tragen gekommen ist, wofür ich aber sehr sensibilisiert bin, ist <b>Gender-Mainstreaming</b> . Das heißt also: Berücksichtigen die Studieninhalte, die Ziele und die Methoden die Interessen männlicher und weiblicher Studierender?“ (13) „...die Hochschule braucht eine <b>Internationalisierungsstrategie</b> .“ (14) „Dann haben wir in Bezug auf Praxiskontakte auch differenziert: Also dieses Kriterium, der Hochschullehrer hat <b>Kontakte zur Praxis</b> , daraus ergeben sich Diplomarbeiten, Belegarbeiten, Praxisstellen usw. Das haben wir eigentlich ausgeschlossen, weil das machen alle - oder die meisten. Also es muß noch ein bißchen weitergehen: Also haben die Leute in der Praxis gefunden, die z.B. die Lehrer materiell unterstützen, Geräte herstellen o.ä. oder in irgendeiner anderen Weise wirksam werden? Hat ein Hochschullehrer irgend jemanden aus der Praxis überzeugen können, sich an der Lehre zu beteiligen?“ (16) B spricht

<sup>267</sup> B erläutert im Zusammenhang mit der Schaffung eines Stellenpools ausführlicher, wie Anreizmechanismen für Innovationen/ Profilbildung geschaffen wurden: „Ich denke, wenn es absolut keine Veränderung von Rahmenbedingungen gibt, dann gibt es ja auch keinen Grund sich etwas Neues auszudenken. Also wir (...) zeigen eben Notwendigkeiten auf. Zum Beispiel, daß sich nur etwas ändert, wenn wir einen **Stellenpool** haben, über den wir frei verfügen können und der nicht wieder in irgendeinen Fachbereich geht und blockiert ist bis der Kollege, die Kollegin in Pension geht - und solche Dinge.“ (28) „Wir haben beschlossen, daß wir bestimmte **Kompetenzfelder oder Profildfelder** - wie immer man das am Ende auch nennen wird - (hier) für die Hochschule beschreiben werden. Das werden maximal fünf sein. Da sind einmal Traditionen der Hochschule drin, aber auch die neueren Entwicklungen.“ (29) „Also der **Gruppenbonus** kommt ja in zweierlei Weise. Also einmal in Bezug auf Stellen insgesamt. – Wir haben einfach gesagt: es muß ein paar Kernkompetenzen geben und darum herum gibt es dann Weiteres. Also dieses Drumherum wird auf die Gruppe bezogen sein. Dann zum Beispiel Lehrauftragsmittel, investive Mittel. Das sind alles Dinge, die der Gruppe zugute kommen können. Aber dann natürlich investive Mittel oder wir haben dann so einen Topf, um die Drittmittelfähigkeit herzustellen. - Das wäre ja auch wieder etwas, was die Einzelnen in Anspruch nehmen können. Also ich denke man könnte ziemlich gut abgrenzen was also ein **Bonus für einen Einzelnen** wäre und dann wieder für die Gruppe. Im Grunde haben wir das jetzt zum Teil ja auch schon so.“ (31)

	mehrfach <b>Arbeitsmarktchancen</b> für die Studierenden an: „Also z.B. wo kommen die unter, wie gut kommen die unter?“ (13, 50 ff.)
3 Akzeptanz von Stud.- und Absolventenbefragungen	B formuliert in Zusammenhang mit der Aktualität der Lehrinhalte: „... das <b>können Studierende sehr wohl beurteilen.</b> “ (9) Und was sicher auch wichtig wäre: wie sozusagen dieser Einsatz des Hochschullehrers/ der Hochschullehrerin sich auswirkt auf die Chancen für die Studierenden. Also z.B. <b>wo kommen die unter, wie gut kommen die unter?</b> Kommen die <b>in ihrem Fachgebiet unter oder woanders usf.?</b> Gibt es da auch über solche Kontakte von Professoren und Professorinnen gute Praxisstellen, gute Diplomarbeiten usw.? Also ist das eine unmittelbare Wirkung eines Engagements eines Professors oder einer Professorin? (13)
4 Gewichtung Leistungsanteil	B äußert sich hierzu nicht explizit, spricht jedoch mehrfach große Unterschiede in der Aufgabenerfüllung an und daß dies <b>entsprechend Aufgabenerfüllung anerkannt</b> bzw. honoriert werden müsse. (31 ff., 37)
5 Gewichtung Lehre, Forschung usw.	B sieht die verschiedenen Aufgaben der Fachbereiche bzw. „Profilfelder“ prinzipiell gleichermaßen als wichtig an: „so ein Profilfeld muß sich auszeichnen durch: Profil in Lehre, Forschung und Weiterbildung. Also die drei Dinge müssen als Profil zusammen passen.“ (29)
6 Art der Kopplung Bewertung und Mittelvergabe	Von B wird <b>direkte Kopplung abgelehnt:</b> „Also, daß die Starken noch mehr kriegen und die Schwachen noch weniger. Also das ist ein System, das die Ungleichheiten verfestigt, aber nicht hilft sie aufzulösen.“ (36) „Also in Bezug auf die Fachbereiche sehe ich da eine gute Möglichkeit. Das <b>diskutieren</b> wir jetzt auch. Das kann z.B. Inhalt von <b>Zielvereinbarungen</b> sein. Die Fachbereiche machen Vorschläge: was sie alles für tolle Sachen machen werden und wir werden mit denen in der Zielvereinbarung festschreiben über welchen Zeitraum die die Chance haben das zu verwirklichen. Und an dem Punkt gucken wir noch einmal und entweder sagen sie dann: Wir sind ganz toll in den Startlöchern und jetzt läuft es aber – (so) und dann werden sie vielleicht eine Verlängerung des Vertrages kriegen unter ähnlichen Bedingungen. Und dann wird eben entschieden: Ja oder Nein. Also das haben wir schon ganz klar so besprochen. So eine Regulation kann man ja ganz gut über solche Zielvereinbarungen bekommen. Wie man das jetzt auf der individuellen Ebene hinkriegt (, das) weiß ich nicht so ganz genau. Aber z.B. wenn ich mir überlege (...) irgend jemand hat über Lehrveranstaltungsevaluation bescheinigt gekriegt: die didaktische Qualität ist mittelmäßig, es könnte sich noch verbessern, dann könnte der über den Nachweis, daß sich das verändert (z.B.) in eine positive Richtung – weil er eben zwischendrin eine Weiterbildung gemacht hat, also einen direkten Effekt von Weiterbildung auf Lehre, das auch anerkannt bekommen.“ (37) <sup>267</sup>
7 Art der Spreizung	B äußert sich hierzu nicht explizit.

Fragen	2. Fall: N
1 Grund- position	N nimmt hier eher eine <b>Mittler-Position</b> ein und formuliert hierzu, was sie als Extrempositionen sieht: „Also, was ich sehr schwierig finde ist, wenn man versucht das eine Modell, das sich als schlecht erweist durch das andere Kontrastmodell zu ersetzen, was aus meiner Sicht auf jeden Fall zeigen wird, daß es schlecht ist, weil es sozusagen, einen Schwenk <b>von dem einen Extrem ins andere</b> Extrem macht. Also von der völligen Unkontrolliertheit quasi - momentan ist es ja relativ egal, wie viele Absolventen man produziert - in die sozusagen ökonomisierte Kontrolle und setzt damit auch Leistungs- bzw. überhaupt Handlungsanreize, die problematisch sind, find ich.“ (55) Das greift sie zum Schluß auf: „...ich kann mich in keines der beiden Lager schlagen“ (89)
2 geeignete Kriterien	N formuliert implizit ein Kriterium: „Also setzt man die z.B. in einem Dekanat am besten dafür ein, daß man mit den Studierenden gemeinsam an der <b>Weiterentwicklung von Studiengängen</b> arbeitet...“(56), würde aber insgesamt „...gar <b>nicht so auf die Leistung abzielen, sondern auf die Belastung</b> , ob einer hundert Diplomanden hat oder einen...“ (57). Im übrigen favorisiert N nicht Absolutwerte, sondern eher Relationen: „ <b>Relation</b> ist immer das vorher, wo diese Frage, ob jemand besonders gut oder besonders schlecht ist, einfach gar keine Rolle spielt. Den <b>Vergleich</b> dazu finde ich sinnvoller.“ (63)
3 Akzeptanz von Stud.- und Absolven- tenbefrag- ungen	„...es ist unterschiedlich zwischen den Fachbereichen. Grundsätzlich gibt es aber eine <b>sehr große Zustimmung zu studentischen Befragungen</b> . Die Frage ist nur, welche Relevanz haben sie im Endeffekt, also werden sie eingesetzt, um zum Beispiel leistungsbezogene Mittelzuwendungen zu machen, was wir innerhalb der Uni nicht machen...“ (67) „es gibt eigentlich keine Resistenz gegenüber Studentenforschungen, außer bei Einzelnen.“ (67) „Widerstände, die so geäußert werden, sind in der Regel die, daß man auch da droht quasi naiv mit umzugehen, daß man praktisch jetzt die Studierendenbefragungen zum Maß aller Dinge macht...“ (68) Als Indikatoren für Mittelzuweisungen hält sie sie für wenig geeignet: „Ich glaube nicht, daß es einen wirklichen, großen Effekt hat, dies wiederum über Indikatoren zu sanktionieren zu versuchen, ich glaube, daß dies nicht funktioniert.“ (70) „Also dann müssen halt die Dekane mit den einzelnen Lehrenden sprechen.“ (70) Insgesamt würde sie es eher weniger formalisiert zu lösen versuchen, z.B. über „mindestens einmal im Semester eine Lehrendenkonferenz“ (71). „Ich merke halt, daß <b>diese starken Formalisierungen</b> von solchen Techniken zu diversen Unterlaufungsstrategien führen, genauso wie die Unis das zwischen Land und Uni unterlaufen, tun das die Lehrenden oft auf dieser Ebene“ (71) Sie ist jedoch nicht grundsätzlich gegen Indikatoren: „...ein paar Indikatoren finde ich wichtig, zum Beispiel die <b>Absolventen finde ich eine relevante Frage</b> , weil wenn die Uni einfach gar keine Absolventen produziert, dann finde ich sie hochproblematisch“ (76) „Und da finde ich immer die Modelle ganz gut, die halt einen Teil vergleichbarer Indikatoren haben, so daß man einfach nicht als Schlechtester starten kann und einen Teil Individualindikatoren. Das finde ich eigentlich immer (ein) ganz gutes Modell, daß man sich da nicht einfach rausstellen kann.“ (88)
4 Gewichtung Leistungs-	Hierzu möchte sich N nicht festlegen, da es von vielen Randbedingungen abhängt und daher schwierig einzuschätzen ist: „ <b>Das kommt sehr darauf an</b> , wie die Grundausstattung ist.“ (80) „...also ich

anteil	wüßte es einfach nicht wieviel sinnvoll wäre.“ (81)
5 Gewichtung Lehre, Forschung usw.	„...ich finde eine Uni und eine Professur an der Uni hat sich darauf eingelassen und eigentlich auch einzulassen, daß die <b>Lehre die Hälfte der Tätigkeit</b> ausmacht und ich würde es fast zu defensiv finden zu sagen man nimmt die Lehre raus. Also eine Universität ist eben eine Hochschule und kein Max-Planck-Institut. (...) Ich meine klar, man könnte natürlich pragmatisch sagen gut, wenn einer nicht lehren kann, dann muß man sich etwas überlegen. Ich würde aber nicht als Institution sozusagen nachgeben wollen, (denn) das finde ich eine ganz schwierige Kiste, weil das auch zwischen den Institutionen sehr starke Rivalitäten produziert.“ (85) „Ich würde schon sagen <b>Korridor, aber beides muß sehr prominent vertreten sein...</b> (bei der Gewichtung von Lehre und Forschung)“ (86)
6 Art der Kopplung Bewertung und Mittel- vergabe	N betont, sie: „...finde, daß das jetzige System nicht fair ist, weil es sich zu wenig an unterschiedlichen Belastungen und unterschiedlichen Engagements von Lehrenden orientiert. Und es setzt auch die falschen Leistungsanreize, weil man ja, wenn man das System beobachtet, sagen könnte, eigentlich ist es egal, ob ich was mache oder nicht, es kommt eh dasselbe dabei raus. Da könnte man umgekehrt formulieren, die <b>Anreizstruktur ist eigentlich, nix zu machen</b> , weil man kriegt das gleiche Geld, die gleichen Forschungsmittel und hat da nicht so viel Streß am Hals.“ (57) „Und man muß auch mal die Frage stellen, will man die Guten belohnen oder will man die Schlechten sozusagen .. irgendwie nach vorne holen (...). Und wir machen das momentan eigentlich lieber so, daß wir, wenn es so wirkliche Ausreißer gibt also wo wir das Gefühl haben jetzt müssen wir mit den Leuten reden, das auch tun und sonst eher <b>versuchen über positive Anreize zu gehen</b> .“ (70) Ich glaube man kann schon als Institution sich Ziele setzen .. also wir haben zum Beispiel eine Zielvereinbarung für diese Uni für uns selbst geschrieben, nicht für das Ministerium o.ä. Und man kann dann natürlich sagen, diejenigen, die daran mitarbeiten diese Ziele wirklich umzusetzen, (...) die können dann damit rechnen an den Töpfen, die wir zur Verfügung haben (...), stärker partizipieren zu können.“ (73) „Ich finde auch, daß man da sozusagen (versch. Arten der Kopplung) <b>kombinieren</b> muß und das machen alle schlaun Modelle. Auch die schlaun Modelle des CHE, die es sogar versuchen, sozusagen das „ist“ auch irgendwie zu berücksichtigen. Also zum Beispiel jetzt eine Uni, die schon dreißig Prozent Ausländer hat oder so, die kann natürlich ihren Ausländeranteil nicht noch beliebig steigern. Und wenn man da dann nur das Positive/ die Weiterentwicklung bemißt, dann kommt man irgendwann zu einem konträren Ziel. Und ich glaube das muß man immer mit in Betracht ziehen (,d.h.:) auf welcher Schwelle sich das „ist“ befindet und das einrechnen in die Verbesserungen.“ (87)
7 Art der Spreizung	Diese Frage entfiel, da Zulagen- und Mittelverteilung nach einem solchen, als viel zu stark formalisiert eingeschätzten Modell <b>grundsätzlich abgelehnt</b> würde. (89)

Fragen	3. Fall: M
1 Grund- position	<p>Für M sind <b>Leistungsbewertungsverfahren notwendig</b>: „Nun, daß diese Verfahren Unfug seien, ist sicherlich nicht haltbar. Denn man <i>muß</i> zu Steuerungsinstrumenten dieser Art im Rahmen der internen und externen Hochschulreform notwendigerweise kommen. (Also) Das ist überhaupt keine Frage. Quantitative Kriterien haben natürlich immer den Charme der Meßbarkeit und auch den Charme der Möglichkeit, daß man mit ihnen Dinge in unterschiedlichen Wissenschaftskulturen vergleichen kann. Obwohl das schon manchmal nicht ohne Probleme ist.“ (91) „Das Qualitative hat das Problem, daß es sich doch nach den einzelnen Wissenschaftskulturen stark unterscheidet und, daß man da wahrscheinlich, jedenfalls über eine sehr komplexe und große Universität, es sehr schwer haben wird, zu Maßstäben zu kommen, die unterschiedslos auf alle Fakultäten und alle Teile anwendbar sind. Gleichwohl muß man dann trotzdem in der Richtung weiterarbeiten und (dann) eben in kleineren Räumen mehrere Vergleichsmaßstäbe werten.“ (92)</p>
2 geeignete Kriterien	<p>M schätzt <b>quantitative und qualitative Kriterien als unverzichtbar</b> ein und erläutert dies am Beispiel angestrebter Internationalisierung: „Es ist sicherlich richtig, also Köpfe allein zählen. Das ist ein relativ grobes Instrument, aber es ist als solches auch nicht verzichtbar. Denn wenn wir an der Universität einen Anteil von zwei Prozent bei den (ausländischen) Dozenten haben, dann ist das eben, auch im kritischen Vergleich, ganz außerordentlich wenig – viel zu wenig. (...) Im übrigen wird es ja sicherlich auch so sein, daß der Erfolg einer solchen Internationalisierungsstrategie nicht nur von Marketing abhängt, sondern eben auch von inhaltlich qualitativen Gesichtspunkten. Und dazu gehört der curriculare Aspekt unbedingt dazu. Dazu gehört auch der sprachliche Aspekt. Dazu gehört sicherlich auch das Umfeld, (das) was die jetzt als <i>Noneducational Services</i> bezeichnen.“ (93)</p> <p>„Ja, Sie wollten das gerne noch ausdifferenziert haben für die Ebene der Besoldung. Das ist nicht nur eine Frage, die an die Hochschulen zu stellen ist, sondern an den öffentlichen Dienst generell. Das hat natürlich alle Probleme der Meßbarkeit und der Vergleichbarkeit von Leistungskriterien und mitunter spielen da auch noch <b>gesellschaftliche Rahmenbedingungen</b> eine Rolle. Es gibt Fächer, in denen – nehmen wir die ingenieurwissenschaftlichen Fakultäten – die außeruniversitären Besoldungsskalen ganz anders sind, als wenn man sich im Bereich der Sozial- und Geisteswissenschaften befindet. Als man kann dies auch nicht als isoliertes Problem der Universitätsbesoldung oder auch der Besoldung des öffentlichen Dienstes sehen, sondern man muß die gesellschaftlichen Rahmenbedingungen (vielleicht) noch stärker bedenken, als das bis jetzt der Fall gewesen ist. (...) Also man kann das noch kleinräumig ausdifferenzieren, aber grundsätzlich wollte ich darauf abheben, daß – weil hier wieder Beamte etwas für Beamte erdacht haben – dieser Aspekt vielleicht nicht ganz – und auch der internationale Aspekt, wenn man Mobilität fördern will – in der Weise mitbedacht worden ist, wie man dies für wünschenswert halten kann.“ (93)</p> <p>„Also mit diesem sowohl Steuerungs- als auch Bewertungsinstrumentarium stehen wir relativ am Anfang, wenn man das vergleicht mit (etwa) Dingen, die man schon immer gemacht hat, - Ich meine, weil man sie im Berufsungsver-</p>

	fahren braucht, nämlich die Bewertung einer individuellen Lehr- und Forschungsleistung. - dann können wir da sehr viel mehr, auch wenn es eine Reihe von Kollegen gibt, die immer abstreiten, daß man da (ja) vergleicht. Aber das tut man natürlich, denn sonst könnte man auch eine Münze werfen, wen man beruft. Das heißt, <b>wir haben die Kriterien, man muß sie ein bißchen klarer stellen.</b> “
3 Akzeptanz von Stud.- und Absol- ventenbe- fragungen	Hierzu erläutert M, dies sei in den USA bereits seit langem der Fall, und „es <b>bringt das Hochschulsystem nicht zum Zusammenbruch</b> “. Einschränkend wird jedoch angemerkt, daß dabei <b>evtl. Probleme der Vergleichbarkeit zu berücksichtigen</b> sind. Z.B. seien eher unbeliebte Pflichtveranstaltungen und Veranstaltungen, die überwiegend aus Interesse besucht werden, nicht ohne weiteres vergleichbar. (95)
4 Gewichtung Leistungs- anteil	Hierzu führt M aus, daß das Grundgehalt auf jeden Fall wettbewerbsfähig sein müsse (vgl. auch weiter vorn im Interview die Ausführungen hierzu). Der <b>Leistungsanteil sollte <math>\geq 15\%</math> sein</b> , dies hänge aber von finanzpolitischen Rahmenbedingungen ab, die derzeit (und wohl auch auf absehbare Zeit) komplett unter der Prämisse stehen, daß Anreizsysteme „kostenneutral zu sein haben“. Für die Ebene von Studiengängen, Fakultäten bis hin zur Hochschulebene merkt M an, daß eine zu starke Gewichtung eines leistungsbezogenen Finanzierungsanteils verfassungsrechtlich bedenklich sein könnte, wenn er in die Freiheit von Forschung und Lehre eingreift. (95)
5 Gewichtung Lehre, Forschung usw.	Hier preferiert M eine <b>Gewichtung von Forschung und Lehre zusammengenommen in jedem Fall <math>\leq 75\%</math></b> (um Raum für andere zusätzliche Ziele und diesbezügliche Anreizmöglichkeiten zu schaffen, auf der Hochschulebene z.B. Wissenstransfer, Berufseinstiegsunterstützung u.a.). Zur Erwähnung der Diskussion um die Bedeutung des Humboldt'schen Bildungsideal für diese Gewichtungsabwägung merkt sie an, daß die Aufgabenteilung inzwischen sehr viel mehr ausdifferenziert ist. (95)
6 Art der Kopplung Bewertung und Mittel- vergabe	M <b>lehnt eine direkte Kopplung von Leistungsbewertung und Mittelvergabe ab</b> . Wegen nicht intendierter Nebenwirkungen sei dies das weniger geeignete Instrument. Eine Kombination der anderen beiden genannten Kopplungsarten (und evtl. weiterer) wäre vorstellbar, jedoch nur bis zur Ebene von Fakultäten und (maximal) Instituten herunter möglich, da es hier um Lehre geht. Hierbei bringt M das Instrument Zielvereinbarung in die Diskussion ein. Was dieses leisten kann, hänge jedoch von den Entscheidern ab. (96)
7 Art der Spreizung	Zu dieser Frage möchte sich M nicht äußern. Diese Diskussion wird für „noch nicht spruchreif“ gehalten, zuvorderst wären viele andere Dinge zu klären. (96)

Fragen	4. Fall: G
1 Grund- position	G äußert sich als Befürworter von Leistungsbewertungsverfahren: „ <b>Leistungsbewertungen sollte man unbedingt machen</b> , denn ansonsten gibt es ja überhaupt keine Motivation für Leute, denen es ein bisschen schwerer fällt (auch) <sup>268</sup> einen Anreiz zu schaffen. Unter anderem für die, die ein bisschen lasch mit sich umgehen würden. Also, das muss eine Leistung sein. In der Leistungsgesellschaft da auch die Hochschulen einzubeziehen.“ (101)
2 geeignete Kriterien	G äußert sich zunächst für Individual- statt Gruppenzulagen: „Ja und da gibt es ja nur diese Individualregelung. Weil, wenn man sonst eine Gruppenprämie macht – das hatten wir vorher schon einmal – das bringt überhaupt nichts. Und das ist auch schwer vergleichbar. Und ein Mensch ist von Grund auf materiell bedacht. Also, <b>wenn er individuell nichts weggenommen bekommt, dann macht er nichts.</b> “ (102) Welche Kriterien im einzelnen als geeignet scheinen, dazu gibt es lediglich indirekte Äußerung anhand von Aspekten, die wichtig scheinen:  „Also, was mir nicht gefällt ist, daß da zu wenig <b>Abstimmungen in den Lehrveranstaltungen</b> , in den angrenzenden Lehrveranstaltungen miteinander durchgeführt werden. Das heißt also, es gibt Überschneidungen zu bestimmten Problemen. Oder andere Dinge fallen dann unter den Tisch, die mitkommen müssten. ... <sup>269</sup> . Und man baut dann aufeinander auf. Und so gibt es auch andere Leute. Bauchemie z.B., da müssten die Leute sich mehr zusammensetzen. Und das gleiche betrifft die Informatik. Die Informatiker ... <sup>270</sup> in der Grundauflösung, in der Dienstleistung in bestimmte angrenzende Wissenschaften. Was bringen die da? Was sollte ich zielgerichtet <sup>271</sup> , um vielleicht die Informatik im Maschinenwesen besser zu unterstützen? Diese Abstimmungen sind noch viel zu wenig. Und damit entstehen Lücken, die nicht sein müssten. Und damit entstehen Überschneidungen, die auch vielleicht nicht sein müssten. (...) Und das hängt auch mit den Professoren zusammen. Daß die Lehrenden, wie Sie auch vorhin schon sagten - da ist jeder sehr schlau auf seinem Gebiet. Und jeder läßt sich nicht gerne in die Karten kucken. Und deshalb fasst der das als Angriff auf seine Persönlichkeit auf (und nicht als irgendwie förderlich). Und er sagt: <i>Ich möchte das beste machen.</i> Und er ist nicht kritikfähig. Je höher der ... <sup>272</sup> , um so weniger kritikfähig sind sie. Das ist mein Eindruck.“ (106)
3 Akzeptanz von Stud.- und Absolven- tenbefrag- ungen	G äußert Bedenken zur <b>rechtlichen Absicherung</b> der Einbeziehung von Studentenbefragungen als Leistungskriterium (103). Zudem kam sinngemäß der Einwand, daß Ergebnisse studentischer Lehrbewertungen <b>durch evtl. Einflüsse „verzerrt“</b> sein können, die nichts mit der Lehrqualität zu tun haben (104). Zur konkreten Nachfrage zur Einschätzung der Akzeptanz der Befragung von Studenten und auch Absolventen zur Einbeziehung praktischer Erfahrungen aus dem Berufsleben wurde geäußert: „Aber, <b>wenn dann nur gemeinsam mit den</b>

<sup>268</sup> Einschub: i.S.v.: für sich selbst

<sup>269</sup> unverständlich (2 Sätze) G zählt verschiedene aufeinander aufbauende Teilbereiche der Chemie auf. (Physikalische Chemie. Organische Chemie. Analytische Chemie.)

<sup>270</sup> unverständlich (1 Wort) Vorschlag: schauen

<sup>271</sup> Einschub: tun

<sup>272</sup> unverständlich (1 Teilsatz) vermutlich: „der Status ist“

	<b>Fachleuten.</b> Nicht bloß, daß jemand angerufen wird für die Studieninhalte, sondern vielfach Leute, um die es eigentlich geht, die sollten bei diesem Gespräch mit dabei sein. Und die sollten das dann auch selber mit tragen können, damit sie unmittelbar eingreifen können und nachhaken können.“ (107)
4 Gewichtung Leistungs- anteil	Hier äußert sich G spontan und klar, bezogen auf die Individualebene: „ <b>75% als Grundgehalt.</b> “ (113) Bezogen auf die institutionelle Ebene (von Studiengängen bzw. Fakultäten) schätzt G ein: „Und da sollte man also achtzig Prozent nicht unterschreiten. <b>80% Prozent Grundzuführung</b> sollte es geben.“ (114)
5 Gewichtung Lehre, Forschung usw.	Hier ist G, bezogen auf die Individualebene: „nach wie vor der Meinung, die <b>Lehre</b> ist das, wofür es sich eigentlich ... <sup>273</sup> . Und die soll <b>mindestens zu sechzig Prozent</b> dazu beitragen. Zwanzig Prozent Forschung und zwanzig Prozent Beteiligung an Gremientätigkeit.“ (114) Bezogen auf die institutionelle Ebene meint G: <i>Siebzig – dreißig. Siebzig Prozent Lehre</i> und dreißig Prozent Forschung.“(115)
6 Art der Kopplung Bewertung und Mittel- vergabe	G wurde auf die Bitte zur Erläuterung der Arten der Kopplung am Beispiel der niederländischen Variante zu Ihrer Einschätzung der Eignung befragt. (Eine <b>zeitlich verzögerte Kopplung</b> wäre dann z.B.: Zwei Jahre erst einmal nur im Fachbereich, wo festgestellt wird, daß die Qualität unter aller Kanone ist <sup>274</sup> . Nach zwei Jahren gibt es einen Zwischenbericht und wenn es bis dahin verbessert wird, ist es in Ordnung. Wenn nicht, gibt es noch einmal eine letzte Frist von zwei Jahren. Und wenn es dann immer noch nicht besser ist, wird es eingestellt.) Hierauf antwortet G: „Das <b>finde ich gut.</b> “ (116) G hält aber durchaus auch eine Kopplung von Varianten für vorstellbar. (116)
7 Art der Spreizung	Befragt zur vorstellbaren Spreizung antwortet G: „ <b>Individuell auf jeden Fall für die, die Verantwortung tragen. Und die anderen, denen sollte man die Grundzuführung geben.</b> “ (117) Aber: „dann müsste jeder einen neuen Arbeitsvertrag bekommen“ (118)

<sup>273</sup> unverständlich (1 Teilsatz) vermutlich: lohnen soll

<sup>274</sup> meint: sehr schlecht

Fragen	5. Fall: S
1 Grund- position	<p>S spricht sich klar für Leistungsbewertungsverfahren aus: „Ich bin vom Grundsatz her jemand, der eine <b>Leistungsbewertung schon immer unterstützt</b> hat. Weil ich der Meinung bin, daß man damit Anreize schaffen kann, den Einzelnen - sagen wir einmal - in Hochform zu bringen. Und wirklich etwas zu Stande zu bringen, was er vielleicht ohne einen Anreiz in Form einer Leistungszulage oder Prämie (...) nicht oder nicht in dem Maße intendiert hätte.“ (121)</p> <p>„Wenn man sagt, die Frage ist nicht <i>ob</i>, sondern <i>wie</i> – (dann) hat man natürlich ein Hindernis auf diesem Wege. Und dieses <b>Hindernis ist</b> erst einmal <b>prinzipiell das gegenwärtige Gehaltsgefüge</b> an den Hochschulen (...) Und man (müßte) demzufolge, wenn man so etwas machen will, nicht nur eine Dienstrechtsreform machen. Das beträfe ja nur die Beamten, sondern man müßte auch entsprechende ähnliche Regelungen treffen für die Angestellten im Hochschulwesen.“ (121)</p>
2 geeignete Kriterien	<p>„Das Jüngste sind die ganzen <b>Beurteilungskriterien, die für Beamte in Form von Dienstbeurteilungen festgeschrieben sind</b> und die man eben jetzt für Angestellte auch sukzessive einführen will. Den Kriterien hat der Hauptpersonalrat bereits zugestimmt – gezwungenermaßen. Er hat es lange hinausgezögert. Weil ja meine <b>Meinung, die meine private ist</b>, daß ich also seit langem was für Leistungsanreize übrig habe, durchaus <b>nicht die allgemeine Meinung</b> ist. Es kann ja dann auch bedeuten, denn manche können sich sehr gut einschätzen und wissen demzufolge, daß sie vielleicht dadurch doch etwas schlechter wegkommen würden. Und oft sitzen diese Leute an den Schalthebeln.“ (124) „Was für meine Begriffe notwendig wäre, ist, mal zu sagen: <i>Das ist der ideale Professor. Das ist der ideale Mitarbeiter.</i> (...) Und für diesen idealen, oder durchschnittlichen vielleicht besser (...) Also, was kann ich von einem Professor verlangen?“ (125) Dann folgt eine Retrospektive: „Ich gehe einmal in „alte Zeiten“ zurück. Jedes Jahr im September ungefähr (...) gab es in den Arbeitsgruppen, Lehrkollektiven große Diskussionsrunden, weil es am 7. Oktober immer Prämie gab. (...) Da bekam vornweg jeder drei, vier Blatt. Dort standen harte Zahlen. <b>Wieviel Lehrveranstaltungen welcher Art. Wieviel Belegarbeiten usw.</b> <i>Was wurde jeweils in der Forschung gemacht?</i> Wobei <b>Forschung qualitative Ziele</b> sind, Erfüllung qualitativer Ziele – im Endeffekt. Das ist nichts quantitatives mehr. Da stand zwar im Endeffekt auch etwa: <i>So und soviel Zeit habe ich für Forschung aufgebracht.</i> Aber ansonsten ist das qualitativ, weil der Chef ja gesagt hat, ob es gut oder (ob es) schlecht ist, ob die Ziele erreicht sind oder nicht.“ (126) Insgesamt sieht S qualitative und quantitative Kriterien beide als notwendig: „Leistung, die dann quali- und quantitativ bewertet werden müßte - das ist sicherlich eine Einheit“ (128).</p>
3 Akzeptanz von Stud.- und Absolv.- befrag- ungen	<p>„Es ist – Studentenbefragung und Absolventenbefragung (, beides) – sicherlich <b>das am meisten subjektiv befrachtete Gebiet</b>. Und das muß man berücksichtigen. Ich kann unmöglich, das ist meine Meinung, auf der Grundlage einer Beurteilung eines Studenten oder eines solchen Fragebogens sowohl eine Lehrveranstaltung (verdammten), als auch einen Lehrenden verdammten. (...) Also ich übertreibe jetzt. Das ist ja hier gestattet. Ich weiß, daß das gar nicht passieren kann. Aber ich könnte mir auch nicht vorstellen, daß man das machen</p>

	<p>kann, aufgrund der subjektiven Befragung. <b>Ich halte eigentlich viel von Absolventenbefragungen, wenn man sie 1 Jahr später macht.</b> Und zwar deshalb, weil die Leute dann erst einmal mitbekommen, was eigentlich notwendig gewesen ist in ihrem Studium. Was sie vielleicht als Studenten gar nicht so gesehen haben und was sie jetzt durch das Messen an der Praxis erst einmal merken, daß sie dort eigentlich viel zu wenig Wert draufgelegt haben.“ (130) Auf den Hinweis, daß es bei Vergleichen der Lehrbewertungen von Studenten und Absolventen kaum Unterschiede gab, räumt S ein: „Das kann ich mir auch vorstellen. Entweder der Hochschullehrer ist gut oder der ist nicht gut. Und das ändert sich dann nicht zwischen (...) dem Studenten, der am Fach interessiert ist und dem Absolventen, der (...) an dem Fach interessiert war(...). Das kann im Höchsthalle einmal eine Verschiebung geben...“ (131/ 132)</p>
<p>4 Gewichtung Leistungs- anteil</p>	<p>„Also das sehe ich als ganz kompliziert an. Ich könnte mir im Höchsthalle - also was ich jetzt sage, das ist aus dem Bauch – vorstellen, daß man das so stufenartig einführt. Weil man natürlich die Älteren, die jetzt da sind, in/ bei so einer Sache anders anfassen muß als die Jungen.“ (133) „Dann meine ich, wenn das eingehalten ist, daß das Grundgehalt automatisch so hoch ist, daß er davon existieren kann. So daß der Leistungsanteil wirklich ein Anreiz ist für ihn, irgendwie in den Genuß von etwas Besonderem zu kommen.“ (134) „Ich könnte mir durchaus vorstellen, daß der <b>Leistungsanteil etwa 1/5</b> sein könnte.“ (135) S führt dann im Rahmen Ihrer Retrospektive auf die DDR-Zeit aus: „Und unter dem Strich stand, dieser erhält tausend Mark und jener (kriegt) hundert Mark. (Wenn ich das jetzt einmal ganz extrem formuliere.) Es hat keiner etwas weggenommen bekommen. Aber es ist für jeden etwas draufgelegt worden. Nun kann man das am Jahresende machen. Das ist (ja nun) völlig frei. Damals wurde das einmal im Jahr gemacht. Das Studienjahr wurde sozusagen abgerechnet. Aber es war ein enormes Erlebnis für jeden, der dort saß. Der Eine hat einen roten Kopf bekommen und der Andere hat gestrahlt. Sie verstehen, wie ich das meine, oder? Es hat jeder irgendwie etwas daraus gelernt. Und im Endeffekt hat das vorwärts gebracht.“ (126) „Und ich könnte mir durchaus vorstellen, daß so etwas auch heute den Effekt bringen würde wie damals. (...) Es gab ja zwei verschiedene Sachen. Es gab die sog. Leistungsstufen, wo also während des nächsten Jahres jeder zum Gehalt eine bestimmte Stufe zusätzlich anerkannt bekam und es gab die Prämie. (...) Diese einmalige Auszahlung konnte auch manchmal für eine einmalige ganz enorme Leistung gezahlt werden.“ (127)</p> <p>S favorisiert <b>grundsätzlich Einzelzulagen, in bestimmten Fällen sind aber auch Gruppenzulagen vorstellbar, z.B. bei Zielvereinbarungen</b> für Studiengänge/ Fakultäten (132): „Wenn solche Leistungszulagen in irgendeiner Art und Weise einen Anreiz haben sollen, müßten sie eigentlich (...) bis auf den Einzelnen runter gehen. Wobei ich mir durchaus auch vorstellen kann, daß es für bestimmte Personengruppen solche Anreize geben kann“ (127)</p> <p>„Gruppenprämien gab es bei uns hier in dem Bereich eigentlich sehr selten.“ (128)</p>
<p>5 Gewichtung Lehre, Forschung</p>	<p>„...das muß ein Mitarbeiter bringen. Er muß <b>mindestens 50 Prozent Forschung und Lehre</b> machen, sonst kann er die 2a nicht kriegen. Das sind ja für uns dann immer so die Prämissen.“ (138) Befragt zu Überlegungen, daß ja nun nicht jeder ein begnadeter Forscher ist und auch nicht jeder ein begnadeter Lehrer, so daß</p>

usw.	man das dann vielleicht auch personenbezogen anpassen müßte, z.B. 30 bis 70 Prozent jeweils, meint S: „Warum nicht.“, weist jedoch auf juristische Probleme hin. (139)
6 Art der Kopplung Bewertung und Mittel- vergabe	<p>„Ich gehe einmal davon aus, daß wir innerhalb von (irgend) einem Zeithorizont <b>Zielvereinbarungen</b> haben.“ (127/ 128) „Und da könnte ich mir dann auch vorstellen, daß bei solchen Sachen – das ist ja in dieser Arbeitsgruppe diskutiert worden - wenn das größere Vorhaben sind, eine Anschubfinanzierung nötig ist, weil es sonst nicht gemacht werden kann. Und das aber - ja zu welchem Prozentsatz? Also, damals ist gesagt worden: <b>Mindestens 30 Prozent Anschubfinanzierung. Und der Rest Geld dann, wenn es fertig ist.</b> Warum nicht. Sicherlich muß man das dementsprechend machen, was man erreichen will, daß dabei etwas herauskommt.“ (128)</p> <p>„Also ich glaube, hier sollte man sich nicht unbedingt auf eine Art festlegen. Der goldene Mittelweg, sozusagen eine Kombination – auch wenn Sie hier sagen eine direkte Kopplung - von allen möglichen Varianten, je nachdem, was in dem Moment gerade das Günstigste ist, sollte man dort vielleicht im Bereich des Möglichen lassen. Warum nicht <b>bestimmte Verbesserungen belohnen und trotzdem die Mittelvergabe und Leistungsbewertung koppeln.</b> Das war ja das, wo wir vorhin diskutiert haben. Also eine bestimmte Leistung, die kriegt etwas besonderes, die wird als besonders hervorgehoben und das Normale ist ansonsten die Kopplung an die Leistungsbewertung. Das kann man doch alles machen. Also, dort würde ich eigentlich darin das Optimum sehen, daß das alles in einem gewissen Sinne etwas offen ist.“ (141)</p>
7 Art der Spreizung	<p>„...das ist alles richtig. Aber die Frage 3.7 beantworte ich Ihnen im Moment nicht.“ (142) Im übrigen spricht S sich gegen formelgebundene Mittelverteilung und eher für das Prinzip Leistungsgruppen aus:</p> <p>„Die Bemerkung möchte ich vielleicht noch machen – aus der Erfahrung heraus. Ich könnte mir vorstellen, daß (es) so eine – also was Sie jetzt sagen, daß die in Großbritannien dann aufgehört haben bis auf die fünfte Stelle nach dem Komma womöglich das auszurechnen – (so etwas) überhaupt anzudenken, meines Erachtens Nonsens ist. Das bringt nichts. Ich meine, wenn ich qualitative Kennzahlen oder Kriterien dort miteinbeziehen will, kann ich das sowieso nicht. Und demzufolge bringt mir die qualitative Zahl (in Anführungszeichen) sowieso eine Ungenauigkeit, die viel größer ist als die, die ich überhaupt berechnen kann. Und da kann ich dann auch von vornherein sagen, <i>wer in dem Bereich liegt, und wer in dem, und wer in dem Bereich liegt</i>“ (unterdurchschnittlich, durchschnittlich, überdurchschnittlich - 143)</p> <p>Bereits ganz am Anfang des Interviews hatte S im Zuge der DDR-Retrospektive auch konkrete Erfahrungen mit „Leistungsstufen“ beschrieben. (127)</p>

Fragen	6. Fall: R
1 Grund- position	<p>„Also, Kostenrechnung halte ich für recht einfach machbar. Bei Leistungsrechnung ist es schwierig.“ (147) „Und die Alternative wäre, daß man gar nichts macht. Und wenn man gar nichts macht, keinerlei Leistungsmessung oder keinerlei quantitative oder qualitative Messung von Lehrqualitäten und das Übersetzen in monetäre Einheiten, ob dann nicht die Qualität der Ausbildung und Lehre oder der Institution Hochschule als Ganzes darunter leidet. Das wäre jetzt (erst einmal) so mein allgemeiner Standpunkt dazu. Also, ich sehe <b>Vorteile</b>. Ich sehe aber <b>auch große Probleme...</b>“ (148) „<b>Im Prinzip ist so ein leistungsorientiertes System gerechter</b> und für die Ziele der Hochschule (, das sind Forschung und Lehre) <b>angemessener.</b>“ (171)</p>
2 geeignete Kriterien	<p>„Also, <b>was man bestimmt machen kann, ist qualitative und quantitative Indikatoren von Lehrqualität zu bestimmen.</b> Also, so und soviel Prozent erreichen erfolgreich eine bestimmte <b>formale Ausbildung</b>, also eine Qualifikation. (...) Und dann müßte man noch einen objektiven Leistungstest haben. <b>Was die, die erfolgreich abgeschlossen haben, dann tatsächlich wissen.</b> (...) Und ein weiteres Problem ist, wenn man die Qualität der Lehre nicht symmetrisch abdeckt, also komplett dieses Phänomen gleich abdeckt, sondern nur bestimmte Indikatoren heranzieht. Dann könnte es sein, daß, wenn die Leistung in diesen Indikatoren Rückwirkung hat auf die Steuerung der Hochschulen, (dann) die Studiengänge oder die Institute oder die Professoren versuchen, diese Indikatoren zu maximieren. Nehmen wir einmal an, die Anzahl der Promotionen, pro Promotion würde man Geld bekommen an der Hochschule. Dann würde man die Promotionsquote eventuell auf Kosten der Qualität der Promotion erhöhen. Eine Ebene darunter, der normalen Studienabschlüsse. Also wäre zumindest denkbar. Man müßte dann <b>nicht nur die Anzahl messen, also nicht nur einen Indikator heranziehen, sondern wirklich die Qualität symmetrisch abdecken</b>, also die verschiedenen Indikatoren. Z.B. messen, wie gut diese Leute dann sind. Und das ist schwierig.“ (147/148) „<b>Studiendauer</b> würde ich schon als ein relevantes Kriterium heranziehen für die Qualität der Lehre. (...) hängen natürlich nicht nur von Qualitätsmerkmalen (ab), sondern auch von <b>Umfeldbedingungen</b> ab. Also, Studienzeiten: Frauen mit Kindern werden einfach länger brauchen zum Studieren. (...) Also, so etwas müsste man zumindest mitbeachten als Gewichtungsfaktor. <i>Anzahl der Kinder</i> müsste dann einen Bonus rechtfertigen können. (...) <b>Absolventenquote</b>, halte ich auch nicht für schlecht. Wobei man natürlich aufpassen muss, daß diese Absolventen, die ausgebildet werden auch eine gute inhaltliche Ausbildung oder ein gutes Qualifikationsniveau erreicht haben. Das ist aber ein zusätzlicher Aufwand. Da müsste man messen, wie gut sie sind, was sie an beruflichen Chancen<sup>275</sup>, wobei aber berufliche Chancen auch wieder abhängen vom Arbeitsmarkt usw. Die <b>Zahl der Studierenden</b> ist natürlich ein ganz wichtiges Kriterium. (...) je mehr Studierende, um so höher soll natürlich auch die Ausstattung sein für das Institut.“ (149)</p>

<sup>275</sup> nachträglicher Einschub: haben

3 Akzeptanz von Stud.- und Absolven- tenbefrag- ungen	<p>„Jetzt kann man natürlich (noch) qualitative Indikatoren heranziehen oder Indikatoren, die irgendwie Qualität messen. Und ein wichtiges Kriterium wäre natürlich die <b>Qualität von Lehrveranstaltungen</b>. Das sollte man auch heranziehen für Gehaltsverhandlungen, bzw. Gehaltsbemessungen, wie gut aus Sicht der Studierenden z.B. Strukturierung oder Lehrkompetenz oder der selbsteingeschätzte Lernerfolg durch eine Veranstaltung ist. (...) das <b>Studierendenurteil wird auf jeden Fall das wichtigste sein</b>. (...) Ein anderer Indikator wäre natürlich die Qualität der Absolventen. Dafür jedoch würde man einen objektiven Leistungstest benötigen. Den aber wird man nicht für alle möglichen Studiengänge entwickeln können. (...) Deshalb wird man hier nur Teile der Qualität abdecken können.“ (152/ 153)</p> <p>Auf die Frage, was R von Verbleibsstudien hält, wird formuliert: „Also, da würde man wahrscheinlich als Kriterium heranziehen, die <b>Chancen einen Beruf zu finden. Das ist natürlich ein weiteres Kriterium</b>. Das ist jeweils auch abhängig von Merkmalen des Arbeitsmarktes oder von Merkmalen der Personen, die dort ausgebildet werden. (Also, ) Im Psychologiestudium z.B. haben wir jetzt zwischen <math>\frac{2}{3}</math> und <math>\frac{4}{5}</math> Frauen. Und wenn die Kinder haben oder Familie gründen wollen, dann werden die wahrscheinlich nicht in diejenigen Berufe gehen, bei denen es nicht möglich ist, das mit der Familie zu vereinbaren. (...) Also, ich fände es gut, wenn man die Absolventen 1 Jahr nach Abschluss ihres Studiums befragen würde, 1.) <i>ob sie berufstätig sind</i> und 2.) <i>wie sie retrospektiv die Qualität der Ausbildung beurteilen</i>. Und vielleicht noch 3.) <i>inwieweit sie Inhalte dessen, was sie gelernt haben jetzt in ihrem beruflichen Feld anwenden können</i>. Also, da hätten wir drei Indikatoren, die für aussagekräftig erachtet werden können“ (154)</p>
4 Gewichtung Leistungs- anteil	<p>„Und weil man nur Teile<sup>276</sup> abdecken kann, sollte man auf jeden Fall diese flexiblen Gehaltszulagen nicht zu umfangreich bemessen im Vergleich zum Grundgehalt. Und weil ja auch jegliche Form von Indikatoren fehlerbehaftet ist, (deshalb) darf es nicht zu umfangreich oder ein zu großer Bestandteil des Gehalts sein.“ (153) „Diese flexible Komponente dürfte nicht zu groß sein. Also, das Grundgehalt: 90 oder 80 Prozent und das flexible dann <b>10 oder 20 %</b>.“ (152/ Begründung: 159) Ebene der Studiengänge: „<b>50/ 50 oder <math>\frac{2}{3}</math> zu <math>\frac{1}{3}</math></b>.“ (162)</p>
5 Gewicht- ung Lehre, Forschung	<p>R explizierte diesen Aspekt im Interview nicht, sieht jedoch grundsätzlich Lehre und Forschung als eng zusammenhängende Aufgaben: „...Ziele der Hochschule (, das sind Forschung und Lehre)...“ (171)</p>
6 Art der Kopplung Bewertung und Mittel- vergabe	<p>„Also die Möglichkeit mit den Zwischenberichten, wo erst am Ende über Sanktionen entschieden wird, (...) damit die Leute die Möglichkeit haben, sich zu verbessern. Das ist eigentlich das optimale Modell. Die Leute werden evaluiert, dann bekommen sie eine Rückmeldung, dann bekommen sie eine Beratung und schließlich wird geschaut, ob sich etwas verbessert hat oder, wenn es gut war, ob es gleich geblieben ist. Und dann wird entschieden.“ (163)</p>
7 Art der Spreizung	<p>R <b>favorisiert Leistungsgruppen</b>, und zwar: „3 oder 4. Mehr, würde ich sagen, sind überdifferenziert. (Also, bisher haben wir ja 2: C3 und C4. Vielleicht noch C2. (...) 4 Gruppen, denke ich, müßten reichen...“ (167)</p>

<sup>276</sup> Hier ist gemeint: des gesamten Spektrums der Qualität (vgl. ausführliche Interview-Transkription S. 153)

Fragen	7. Fall: A
1 Grund- position	<p>„Ja, ich finde das <b>prinzipiell gut</b>, daß man das macht. Es ist ja auch so, daß es jetzt nicht 50 oder mehr Prozent dessen ist, wahrscheinlich, was ein Professor verdient. (...) ich meine, es kann natürlich solche Fälle geben, daß jemand Leistungszulagen bekommt, die auch sein Grundgehalt übersteigen. Aber das ist ja nicht der Regelfall und mit solchen Mitteln zu arbeiten, das finde ich sinnvoll, finde ich auch gut. Ich finde es bedauerlich, daß der Vergaberahmen in der Weise gedeckelt ist, wie es eben der Bundesgesetzgeber da vorgesehen hat, denn das heißt ja schon, daß nicht die selben Leute wie heute, aber viele Leute im Prinzip mit weniger als im heutigen Durchschnitt nach Hause gehen. Und das ist für mich nicht so günstig. Prinzipiell finde ich das aber eine sinnvolle Sache. Also <b>auf einer Skala von 0 bis 10 mit 10 "Begeisterung", 0 "furchtbar", da ordne ich mich locker bei 7 oder 8 ein.</b>“ (173) „Damit sind wir schon bei einem anderen Problem: Ich denke, die ganze <b>Frage der Verbeamtung</b>. Also es ist ja ziemlich ungewöhnlich, daß man sagt, man schafft ein System der leistungsorientierten Besoldung und behält trotzdem den Beamtenstatus bei. Das, finde ich, ist eine sehr interessante Frage, wie sich diese Dinge in den nächsten Jahren entwickeln werden. Und ob an dieser Geschichte nochmal gerüttelt wird oder nicht.“ (174)</p>
2 geeignete Kriterien	<p>„...wünschen, daß möglichst viele Entscheidungen an die Hochschulen gehen. Gerade auch, was Kriterien betrifft.“ (175)</p> <p>„<b>Anzahl der Absolventen, Studiendauer. Studiendauer ist so ziemlich problematisch</b>, muß ich eigentlich sagen, weil ich mich frage, inwieweit sind tatsächlich die Gründe für eine lange Studiendauer tatsächlich mit der Leistung des einzelnen Professors zu verknüpfen? Also ich meine, prinzipiell ist es doch so, es gibt <b>externe Rahmenbedingungen</b>, die sind an den Hochschulen teilweise nicht besonders günstig. Aber die liegen sehr wenig, glaube ich, oder höchstens teilweise in dem, was ein einzelner Professor beeinflussen kann. Also müßte man schon mal die Voraussetzung haben, daß die Bedingungen, unter denen diese Leute arbeiten, für alle, die man sich da anschaut, gleich sind. Aber dann gibt es ja auch, sagen wir mal <b>individuelle Gründe</b>, die dazu führen, daß man vielleicht sein Studium verlängert.“ (176)</p> <p>Auf die Anmerkung, wenn man nur quantitative Kriterien als Fakten hat, muß man die <i>Qualitätsentscheidung</i> völlig subjektiv machen in der Kommission, meint A: „Das muß ja nicht so sein, man kann ja Gutachten hinzuziehen, da muß man natürlich abwägen zwischen dem, was wünschenswert und dem, was bezahlbar und vom Aufwand her vertretbar ist. Aber man kann ja zum Teil auf vorhandene Gutachten zurückgreifen. (...) und da ist eigentlich schon davon auszugehen, daß die ihre Schäfchen ein bißchen kennen. Und was jetzt den Lobbyismus in der Kommission betrifft, ich glaube, <i>ganz</i> wird man das nie rausnehmen können aus diesen Dingen( ...) Aber man kann sich genau überlegen, wer in so eine Kommission hinein soll, daß man also auch versucht, ein Gleichgewicht der Kräfte herzustellen. Oder daß man vielleicht auch zum Teil, aber das sind die Hochschulen natürlich, fühlen die sich nicht so wohl mit dem Gedanken, so daß man zum Teil externe Leute reinnimmt.“ (179)</p>
3 Akzeptanz	<p>„Frage, ob Lehrevaluation mit einbezogen werden soll, das denke ich schon. Ich denke <b>"Ja"</b>, man sollte das mit einbeziehen, aber <b>als einen Punkt unter</b></p>

von Stud.- und Absolven- tenbefrag- ungen	<b>anderen.</b> (...) Es gibt in Baden-Württemberg ein Papier, im Prinzip eine Art Vorstufe des, was vielleicht mal Verordnungsentwurf wird, und da ist es so, daß man im Prinzip sowas wie Parameter vorgibt oder Themen. Man sagt also: Leistungen beziehen sich auf und dann ist gesagt Lehre, Forschung, Nachwuchsförderung usw. Und da (...) ist Lehrevaluation auch explizit genannt.“ (184)
4 Gewichtung Leistungs- anteil	„in meinem kleinen Rechenbeispiel, was ja in dem Arbeitspapier mit drin ist, habe ich ja mal mit 1200 Euro pro Person gerechnet und ich denke, in diesem Rahmen wird sich das dann in etwa bewegen und das heißt dann de facto, daß die variablen Zulagen im Durchschnitt vielleicht <b>ein Viertel bis ein Fünftel des Gesamtgehaltes</b> sind.“ (187)
5 Gewicht- ung Lehre, Forschung und Gremien- tätigkeit	„Wenn ich jetzt eine Hochschule hätte [Lachen], würde ich sicherlich all diese <b>drei, sagen wir mal Kriteriengruppen oder Parameter einbeziehen</b> wollen. (...) das wird in der Praxis auch meistens so sein und die <b>Gewichtung wird dann auseinandergehen</b> . Die Universität Göttingen oder Heidelberg wird das nicht in derselben Weise behandeln wie die Fachhochschule Köthen. Und das ist auch richtig und gut so.“ (190)
6 Art der Kopplung Bewertung und Mittel- vergabe	„...im Grunde könnte man sich vorstellen, wenn man so ein <b>formelbasiertes Modell</b> hat, dann braucht man niemanden mehr, der entscheidet. Dann haben sie <b>so eine Art Maschine</b> , oben kippen sie rein die Anzahl der Publikationen usw. und unten kommt raus: Herr Müller bekommt 430 Euro. Das ist natürlich schön und das entlastet einen und diese Entlastung wird eigentlich gewünscht in den Hochschulen, das kann ich auch verstehen. <b>Aber es funktioniert nicht</b> . Sie müssen jemanden haben, der am Ende vernünftige, sinnvolle Entscheidungen trifft und diese Entscheidungen verantwortet.“ (178)  „...eine <b>gute Sache, daß man das zum Teil an Zielvereinbarungen koppelt</b> . Das ist aber sicherlich nicht das, was in den meisten Fällen stattfinden wird. (...) Aber im Prinzip wird das doch so sein, wir schlagen ja im Prinzip vor, daß sich die Leute selber bewerben sollen. Also daß sie selber erstmal einschätzen sollen: "Wie ist meine Leistung", auch um diesen gigantischen Aufwand ein bißchen zu vermindern. Und daß man eventuell auch ein Vorschlagsrecht vom Dekan oder so etwas hat. Daß, wenn jemand schüchtern ist, daß man sagen kann: "Wie sieht's denn aus?" Und dann wird im Prinzip die Leistung in irgendeiner Form bewertet und man sagt ja oder nein und dann kriegt der das Geld. [Lachen] Also im Prinzip das, was Sie als direkte Koppelung bezeichnen. und dann sehe ich im Prinzip diese zusätzliche Möglichkeit, daß man sagt, man koppelt es eben an Zielvereinbarungen, manchmal kann man vielleicht auch Leistungen auch noch nicht so richtig abschätzen.“ (197)
7 Art der Spreizung	„Und ansonsten gibt es ja im Prinzip vom CHE einen Modellvorschlag, also ein <b>Stufenmodell</b> und ich glaube, daß das viele Probleme oder potenzielle Probleme mit einer Klappe erschlägt. Das ist vielleicht auch nicht das Nonplusultra, vielleicht kommen ja da auch noch raffiniertere Lösungsvorschläge, aber ich glaube, das ist eigentlich ein <b>guter, machbarer Ansatz</b> , daß man sagt, man hat also kein, es ist nicht alles free-floating, sondern man definiert bestimmte Level und kann im Prinzip darauf zurückgreifen, sowohl in Berufungs- und Bleibeverhandlungen, als auch was die besonderen Leistungsbezüge betrifft.“ (175)

Fragen	8. Fall: L
1 Grund- position	<p>L betont, daß er <b>mehr Leistungsanreize für sehr sinnvoll und richtig</b> hält. Die Betrachtung der Effizienz als Leistungskriterium sollte jedoch eine Betrachtung der Leistungsqualität mit einbeziehen. Wichtig ist auch, daß Hochschulleistungen nicht zu eng definiert werden. (210)</p> <p>L führte aus, daß er eher Experte für die Mittelverteilung auf der Institutionenebene ist. Es wird besprochen, den Schwerpunkt des Interviews darauf zu legen. Dort könne man die Verfahren grob einteilen in:</p> <ul style="list-style-type: none"> <li>a) formelgeb. Mittelverteilungsverfahren (anhand quantit. Indikatoren)</li> <li>b) Zielvereinbarungen (z.B. anhand von Qualitätszielen)</li> <li>c) diskretionäre Verfahren (Einzelfallorientierte Entscheidungen)</li> </ul> <p>Anschließend wird die Einschätzung der Vor- und Nachteile der jeweiligen Verfahren und dabei die Art der Kopplung thematisiert. (210)</p>
2 geeignete Kriterien	<p>Auf die zunächst offene Frage, welche Kriterien/ Indikatoren stärker den bereits weiter vorn angesprochenen Aspekt der Qualität von Leistungen erfassen könnten, werden <b>Ländermodelle</b> genannt (Niedersachsen, Baden-Württemberg), <b>in denen die Qualität der Leistungen bereits eine Rolle spielt</b>, dann aber v.a. die Eignung der <b>Studiendauer</b> besprochen. (Diese wird als <i>ein</i> Kriterium im Rahmen eines Kriterienbündels für geeignet gehalten, ist jedoch nicht als alleiniges oder zentrales Kriterium geeignet.) Anschließend wird durch den Interviewer die Auffassung eingebracht, daß die bisher diskutierten Kriterien keine guten Beurteilungskriterien für Qualität seien, da sie rein quant. „Output“ messen und damit auch eher Quantität als Qualität fördern würden und daß es daher Überlegungen in die Richtung gibt, wie qualitative Aspekte integriert werden könnten in Leistungsanreizsysteme, z.B. in Form subjektiver Bewertungen der Lehrleistungen/ Lehrqualität durch Studenten, Absolventen und Lehrende. (212)</p>
3 Akzeptanz von Stud.- und Absolven- tenbefrag- ungen	<p>L schätzt <b>Absolventenbefragungen als eine sehr gute Ergänzung zu den bisher verwendeten Indikatoren</b> ein, da sie den Bezug der Studienangebote zum Arbeitsmarkt herstellen könnten. Sie hätten jedoch den Nachteil, daß sie sehr aufwendig sind und es sehr lange dauert, bis Ergebnisse vorliegen, die dann auf eine Situation bezogen seien, die sehr lange zurückliegt. Daher seien zur Einschätzung der aktuellen Situation Studentenbefragungen unverzichtbar. Von den drei Befragten seien sie sogar am besten geeignet. Die Befragung von Lehrenden zur Einschätzung der Lehrqualität wird als weniger geeignet eingeschätzt. Lehrende (bzw. „peers“ bei Evaluationen) könnten lediglich Aufbau und Grundstruktur von Studienangeboten sowie den Fächerkanon, jedoch kaum die Lehrqualität einschätzen. (212)</p> <p><b>Etwas problematischer seien studentische Bewertungen</b>, wenn sie Selbsteinschätzungen darstellen, z.B. bezüglich der sogen. sozialen Kompetenzen, die erst bei Absolventen mit den Anforderungen im Arbeitsleben kontrastiert werden können. Daher wäre es evtl. lohnend, über Ansätze ähnlich denen der PISA-Studie nachzudenken (wo z.B. TextVERSTÄNDNIS und ZUSAMMENHANGSwissen erfaßt wurde).</p> <p>Bei den subjektiven Bewertungen seien außerdem die Ausgangsbedingungen zu berücksichtigen, mit denen diese korrelieren könnten (z.B. studentische Gesamteinschätzungen mit Betreuungsrelationen). Wie in der weiteren Diskussion deutlich wurde, gilt dies letztlich aber auch für andere</p>

	Erfolgsindikatoren zum Studium, wie z.B. die Studiendauer oder Abschlußnoten, die z.B. mit der Auswahl von Studienanfängern in NC-Studiengängen zusammenhängen können. Gleiches gilt für in Zukunft evtl. stärker zu erwartende Auswahlgespräche. Im weiteren Sinne können aber auch Praktikazeiten, die Berufseinstiegssituation und natürlich unterschiedliche Regelstudienzeiten die Studiendauer beeinflussen. (213)
4 Gewichtung Leistungs- anteil	Dies hänge stark vom jeweiligen Modell ab. Jedenfalls sollte es einen Grundanteil geben, der Stabilität sichert. Erfahrungen gibt es bisher nur mit eher geringen Anteilen leistungsbezogener Finanzierung an Gesamtfinanzierung, die in der Größenordnung von 10% liegen. Im nächsten Schritt anzustreben wäre ein höherer Anteil von etwa 20%. (213)
5 Gewichtung Lehre, For- schung usw.	Diese Frage wurde nicht explizit thematisiert. Aus den Äußerungen insgesamt ließe sich mit der gebotenen Vorsicht eventuell ableiten, daß eine gleich starke Gewichtung von Lehre und Forschung jedenfalls auf der hier besprochenen Institutionenebene nicht zwingend vorausgesetzt wird. (210 ff.)
6 Art der Kopplung Bewertung und Mittel- vergabe	Als Vorteile des indikatorgestützten, <b>formelgebundenen Mittelverteilungsverfahrens</b> werden v.a. dessen gute Handhabbarkeit und Transparenz genannt. Als Nachteil bzw. als Gefahr wird die Tendenz gesehen, nur sehr wenige Indikatoren zu verwenden, was starke ungewollte „Nebenwirkungen“ haben kann. Als Beispiel wurde das Mittelverteilungsmodell für die Niedersächsischen Fachhochschulen genannt, welches lediglich 3 Indikatoren einbezog: <b>Studienplätze (zu 60%), Anteil Studierender in der Regelstudienzeit (20%) und Anzahl der Absolventen (20%)</b> . Einige Hochschulen verfolgten dazu die Strategie, ihre Einnahmen lediglich über Maximierung der Studienplätze zu optimieren, wobei die Studiendauer und die Absolventenzahl ins Hintertreffen gerieten. Als Schlußfolgerung wird abgeleitet, solche Nebenwirkungen durch parallel zur Erhebung der Leistungsindikatoren durchgeführte Evaluationen der Studienqualität zu verhindern. (211) Bei <b>Zielvereinbarungen</b> lägen die Vorteile v.a. darin, daß es keinen Automatismus wie beim formelgebundenen Mittelverteilungsmodell gäbe. Außerdem fördern sie besser die Zielklärung und die Bewußtseinsbildung für die Notwendigkeit übergreifender Ziele der Hochschule. Allerdings müßten bei Abschluß von Zielvereinbarungen auch die Sanktionsmöglichkeiten im Falle des Nichterreichens der Ziele geklärt werden. Zudem sollten Zielvereinbarungen wegen der Erfahrung einer z.T. zu starken Diversifizierung der Ziele stärker standardisiert werden, so daß sie z.B. alle Tätigkeitsbereiche einer Hochschule (Forschung, Lehre, Weiterbildung, Dienstleistungen usw.) abdecken. (211/ 212)
7 Art der Spreizung	Dem Prinzip der Einteilung in Leistungsgruppen steht L. aufgeschlossen gegenüber. Dies konnte jedoch aus Zeitmangel nicht genauer erörtert werden. (214)

Fragen	9. Fall: P
1 Grund- position	<p>P vertritt weithin die Ansicht: <i>Die Grundrichtung, wie sie jetzt verfolgt wird, ist korrekt. Aber, an der, der und der Stelle muss man schon noch nachbessern.</i> Aber: „Unter Berücksichtigung bestimmter Prämissen. Die Prämisse ist, daß es einen sehr viel stärkeren Legitimationszwang für Hochschulleistungen gibt. Ich bin nicht der Auffassung, daß dieser Legitimationszwang nun in jeder seiner Facette gerechtfertigt ist. Aber, das ist sozusagen jetzt einmal die Handlungsbedingung. Das werden wir jetzt nicht ändern - unter den gegebenen Haushaltsbedingungen insbesondere.“ (215)</p> <p>P schätzt eine in seinem Bundesland existierende <b>formelgebundene Mittelzuweisung als geeignet</b> ein: „Diese Formel versucht das (Gerechtigkeits-)Problem zu bewältigen. Das macht sie (auch) sogar ganz gut, habe ich mir sagen lassen. Das Problem zu bewältigen, das insbesondere zwischen großen und kleinen Hochschulen besteht.“ (217)</p>
2 geeignete Kriterien	<p>„Auf der Individualebene scheint mir ein sehr <b>ungelöstes Problem</b> zu sein, wie man damit umgeht, ob man die Talente, die der Einzelne hat oder den Grad der Erreichung bestimmter Standards (beispielsweise) bei der <b>Durchführung von Lehrveranstaltungen oder der Abhaltung von Prüfungen oder der Organisation von Studienbetreuung, von studienbegleitender Betreuung</b> (oder was auch immer) - daß dort alle daran gemessen werden, ist, glaube ich, im Augenblick wohl das gängige Modell, ob sie einen bestimmten Standard erreichen. Und dann sind sie entweder unterdurchschnittlich oder sie sind durchschnittlich oder überdurchschnittlich. Und, daß hier in der Regel nicht beachtet wird, daß der typische Fall ist, daß wir es mit unterschiedlichen Menschen zu tun haben, die unterschiedliche Talente haben und wo auch eine <b>Leistungssubstitution</b> stattfindet. Daß man in dem einem Bereich schlechter ist, aber dafür im anderen besser.“ (218)</p> <p>„...im Prinzip in einigen Bundesländern schon soweit, daß eben diese leistungsbezogenen Mittelverteilungen sich <b>immer auf Ressourcen beziehen</b>. Und da meine ich – also, das finde ich sehr gut. Und das sollte auf die individuelle Ebene dann eben auch heruntergebrochen werden. Und dann glaube ich, daß es auch sehr viel weniger Streit über die ganze Sache geben wird.“ (219)</p> <p>„...Bereich, den ich für wichtig halte, der sowohl die institutionelle als auch die individuelle Ebene betrifft. Das sind unterschiedliche <b>Wirkungsintervalle</b>. Bestimmte Dinge wie <b>Absolventenquoten</b> (oder so), die kann man innerhalb eines Jahres oder nach fünf Jahren, wenn man verlässlichere Daten haben möchte, so erheben und vergleichen und in das Verhältnis setzen zwischen verschiedenen Fachbereichen, aber auch innerhalb eines Fachbereiches, um zu sehen, wenn man das mithilfe eines Zeitstrahles macht, ob es da Entwicklungen gibt. (...) Also, insbesondere solche Sachen wie <i>der Erfolg der Absolventen auf dem Arbeitsmarkt</i>.“ (221)</p> <p>„...wenn man jetzt <b>Studiengänge entwickelt, Studiengänge erfolgreich akkreditieren lässt</b>, dort dafür sorgt, daß die nachgefragt werden, daß tatsächlich ein Bedarf damit getroffen wird, weil die Absolventen dann auch in irgendeiner Weise etwas damit anfangen können. Das wäre so ein Punkt, (...) der irgendwie mit angekoppelt werden müsste.“ (223)</p>

	<p>Weitere Krit.: „Wer macht die Prüfungen, wer betreut die Diplomarbeiten, wie werden Diplomarbeiten betreut. Also, ich fände es z.B. gut, wenn das möglich wäre, Kontaktdichte zwischen Lehrenden und Studierenden zu erfassen. (...) Machen die Leute eine Besprechung der Hausarbeiten oder nicht. Das ist ein ganz hartes relativ einsichtiges Kriterium. Das müsste man eben nur erfassen.“ (227)</p>
<p>3 Akzeptanz von Stud.- und Absolven- tenbefrag- ungen</p>	<p>„Studentenbefragungen halte ich für sehr wichtig, wenn man zugleich ihre Grenzen in Rechnung stellt. Studentenbefragungen sind in meiner Sicht Zufriedenheitsmessungen – in erster Linie. Studenten werden als „Experten ihrer Situation“ befragt, nicht als Experten für das Fach (...). Und insofern ist es eine Expertise, die da rauskommt. Und sie bewertet die Lehre nicht danach, ob die wirklich gut ist oder nicht, sondern sie bewerten sie danach, ob sie für mich als Studierender gut ist oder nicht. Und das hat viel damit zu tun, welche Erwartungen der oder die einzelne an das Studium stellt. (...) deshalb, denke ich, haben, bzw. müssen Lehrbewertungen zwei Komponenten haben. Also, die können über Punkte u.ä. durchaus in Bewertungen einfließen, aber viel stärker sollte bei studentischen Lehrbewertungen im Vordergrund stehen, daß die Bewertungen eine solche Befragungsqualität haben, daß die Lehrenden daraus selbst etwas für die Entwicklung der Qualität der Lehre ziehen können.“ (224)</p> <p>Absolventenbefragungen: „Da bin ich <b>gespalten</b>. Also, <b>einerseits finde ich es gut</b>, daß man Lehrinhalte daraufhin abfragen sollte, inwiefern der einzelne das Gefühl hat, die haben etwas gebracht. Ich bin aber sehr <b>skeptisch in Bezug auf die Antworten</b>, die dann regelmäßig kommen: <i>Ja, unsere Ausbildung, die war nicht praxisbezogen genug</i>. ... Der Magisterstudiengang soll mich in die Lage versetzen, mich in einer vergleichsweise kurzen Zeit in ein beliebiges Gebiet zum Spezialisten machen zu können. Und dieses könnte man ja eigentlich, wenn überhaupt, nur zuverlässig beurteilen, wenn man einmal fünf Berufsfelder ausprobiert hat. Also, <b>zum Generalisten und Spezialisten meine ich nur, daß das didaktische Modell so sein müsste</b>, daß man eine generalistische Ausbildung erhält und daneben eine Spezialausbildung. Denn nur, wenn man einmal eine Tiefenbohrung gemacht hat, weiß man, wie man dann eine Tiefenbohrung im Berufsleben anstellt.“ (228)</p>
<p>4 Gewichtung Leistungs- anteil</p>	<p>„Institutionell muss man die Grundentscheidung treffen, ob man in die Personalkosten hineingehen möchte. Die betragen je nach Hochschule zwischen siebzig und achtzig Prozent. (...) Wenn wir sagen der Personalkostenbereich ist geschützt, dann sind wir sozusagen bei <b>maximal zwanzig Prozent</b>, die man ausgeben kann, die man in so ein Leistungsbudget einstellen könnte.“ (232)</p> <p>„Wir setzen jetzt einmal den Fall, daß es allgemeine Akzeptanz findet, daß es leistungsorientierte Besoldungsbestandteile geben soll, und hinterfragen dies nicht weiter. Und dann würde ich sagen, daß die Summe dessen, was da erreichbar ist, dann so hoch sein muss, daß sie einen eigentlich intrinsisch und weniger finanziell orientierten Wissenschaftler motivieren kann, aufgrund der Bezahlung mehr zu tun oder sich zu entwickeln. Also, ich vermute einmal, daß das <b>mit zehn Prozent nicht getan</b> ist. Also, wenn ich als C4-Professor siebentausend Mark netto oder achttausend Mark netto hätte und könnte dann 700 oder 800 Mark draufbekommen oder weniger, ob ich mir deswegen ein Bein ausreiße, weiß ich nicht.“ (233)</p> <p>„Und um jetzt richtig finanziell motiviert zu sein, müsste man dann also schon</p>

	<p><b>in den Bereich von fünfundzwanzig Prozent gehen</b>, würde ich einmal schätzen. Wo dann die Leute auch wirklich sagen: <i>Also gut, wenn mir jetzt plötzlich fünfundzwanzig Prozent fehlen, das merke ich dann schon. Und dann habe ich also schon gewisse Einschränkungen in meiner Lebensführung vorzunehmen. Oder, wenn ich eben fünfundzwanzig Prozent mehr hätte, dann würde sich das schon lohnen.</i> Also, ich würde schon sagen, bei dieser Größenordnung müsste man eigentlich schon ansetzen. (...) Die Masse der Professoren, die hätte sich auch auf die Professur beworben, wenn es nur fünfundsiebzig Prozent des jetzigen C4- oder C3-Gehaltes gewesen wären.“ (234)</p> <p>„Aber, da muss man <b>vielleicht generell über die Besoldungsstruktur reden</b>, also nicht nur an Hochschulen, sondern im gesamten öffentlichen Dienst, also auch im Lehramtsbereich beispielsweise.“ (235)</p>
5 Gewichtung Lehre, Forschung usw.	<p>„Wir haben ja auch ein interessantes Modell gerade in der Diskussion von Niedersachsen in die Debatte geworfen, ob denn dieser Wert des Lehrdeputats wirklich so fix sein muss. Also, ob es nicht Hochschullehrer/ Hochschul- lehrerinnen geben kann, die zehn Stunden lehren statt acht oder zwölf statt acht und andere dann eben nur sechs oder vier.“ (236) „Und, daß dann auch gesagt wird, <b>Okay, der macht ja mehr Lehre.</b> Und in der Punktbewertung wird das angerechnet. Und der kann nicht so viel publizieren und so viele Drittmittel einwerben. Der kann ja dadurch wieder <b>substituieren. Das finde ich ganz o.k.</b> Ich würde das jetzt nicht unter dem Titel Lehrprofessur und Forschungs- professur<sup>277</sup>. Ich denke schon, daß die Lehre sich aus der Forschung speisen sollte oder aus der Beteiligung an der Forschung. So etwas kann man freiwillig machen. Man kann aber die <b>Lehrdeputate auch wettbewerblich verteilen.</b> Man kann sagen, der Durchschnitt ist acht Stunden. Aber, ob man da jetzt genau acht Stunden macht oder mehr oder weniger, das entscheidet sich danach...“ (236)</p>
6 Art der Kopplung Bewertung und Mittel- vergabe	<p>„Da ist natürlich ein systematischer Haken drin in dieser ganzen Sache. Der besteht an dem Punkt, wo man sagt, <i>Jemand, der ganz schlechte Lehre macht, wird dadurch ganz schlecht bewertet.</i> Was sagt man nun zu dem? Sagt man zu dem, <i>Weil Du schlechte Lehre machst, musst Du jetzt mehr in der Forschung tun?</i> (...) Und jetzt jemanden, der sehr gute Lehre macht, dann sozusagen dadurch zu bestrafen, daß man sagt, <i>Du machst so gute Lehre. Du machst mehr Lehre, weil Du so gute Lehre machst.</i> Man müsste auch darüber nachdenken, ob man ein <b>Kreuzanreizsystem</b> macht. Daß man sagt, <i>Jemand, der in der Forschung schwach ist, den darf man nicht zusätzlich in der Forschung schwächen, indem man sagt, Du bekommst keinen Hiwi mehr, sondern dem müsste man eigentlich noch einen Hiwi mehr geben, um zu sagen, Du bekommst jetzt einmal zeitweilig eine bessere Ausstattung, aber da wollen wir dann auch sehen, daß Du Dich verbesserst.</i>“ (237) Auch Modellkombinationen werden befürwortet, aber nicht zu kompliziert (238)</p>
7 Art der Spreizung	<p>„<b>Mehr als drei Gruppen finde ich schon gut.</b> – Wieder wegen der Motivationswirkung. Es ist einfach besser, wenn man innerhalb kleinerer Schritte merkt, daß man vorankommt.“ (243) „hochschulübergreifende Fachbereichsbewertung finde ich angemessener als Fachbereiche“(244)</p>

<sup>277</sup> Einschub: fassen

## Anhang B:

### Leitfaden für Experteninterviews

1. Einleitung: kurze persönliche Vorstellung und des Promotionsvorhabens (vgl. Exposé)

2. Ausgangslage („Rahmen“):

- erklärter Wille der Politik, Alterzulagen in Leistungszulagen umzuwandeln

**Frage ist wohl nicht mehr ob, sondern wie** dies geschehen soll

- hierfür auch bereits Kriterien genannt (vgl. HRK- EntschlieÙung 2. 11. '98)

vgl. Bericht der Expertenkommission Reform des Hochschuldienstrechts '00

**überwiegend quantitative Kriterien** (relat. Hörerzahl, Anz. abgen. Prüfg.)

BMBF diskutiert auch Ergebnisse stud. Lehrbewertungen (als Basis für

Fachgutachter), jedoch seien „Konzepte, wie dies genau geschehen soll,

erst noch zu entwickeln“ (vgl. Konzept bmb+f 2000)

- Leistungsbewertungsverfahren **auch für Studiengänge**, Fak., HS diskutiert, ebenfalls überwiegend quant. Krit. (z.B. Studiendauer, Absolventenquote)

- Hierzu gibt es ja **sehr unterschiedliche Ansichten**:

einige sind der Ansicht, **Vorschläge sind Unfug**, es sollte bleiben wie bisher

andere sind der Ansicht, bmb+f-Vorschläge zielen in die **richtige Richtung**

wiederum andere vertreten die Auffassung, wenn wie geplant die Lehr“leistung“ bzw. die Studienqualität bewertet werden soll, dann sind die bisher diskutierten Kriterien keine guten Beurteilungskriterien, da sie rein quant. „Output“ messen und damit auch **eher Quantität als Qualität** fördern könnten. Daher gibt es Überlegungen in die Richtung, **wie qualitative Aspekte integriert** werden könnten in Leistungsanreizsystem.

3. Fragen:

3.1 Frage: Was halten **Sie** von den bislang diskutierten Verfahren/Leistungskriterien?

- (auf der Ebene von Besoldungszulagen für einzelne Hochschullehrer)

- auf Ebene von Studiengängen, Fakultätsebene, Hochschulebene

3.2 Frage: **Welche Verfahren/ Kriterien** halten Sie für geeignet?

- auf Individualebene

- auf Ebene von Studiengängen, Fakultätsebene, Hochschulebene

3.3 Frage: Inwieweit ist die **Einbeziehung von Stud.-/Absolv.-befragg.** akzeptabel?

- auf Individualebene

- auf Ebene von Studiengängen, Fakultätsebene, Hochschulebene

3.4 Frage: Welche **Gewichtung Grundgehalt/ Leistungsanteil** favorisieren Sie?

- auf Individualebene
- auf Ebene von Studiengängen, Fakultätsebene, Hochschulebene

3.5 Frage: Wie Anteile **gewichten bzgl. Forschung/ Lehre/ Gremientätigkeit?**

- auf Individualebene
- auf Ebene von Studiengängen, Fakultätsebene, Hochschulebene

3.6 Frage: Welche **Art der Kopplung** Leistungsbewertung und –anreize?

- direkte Kopplung von Mittelvergabe und Leistungsbewertung,
- Belohnung von Verbesserungen (oder befristet: des Bemühens darum),
- zeitlich verzögerte Kopplung von Evaluation und Mittelfreigabe (wie NL)?

3.7 Frage: Wie groß und welche Art der „**Spreizung**“ der Zulagen?

- Prinzip **formelgebundener Kennzahlen** als Orientierung an festgestelltem gewichteten Mittel der Leistungskennzahlen (Pseudogenauigkeit?) oder
- Prinzip Einteilung in **Leistungsgruppen** mittels zuverlässiger stat. Verfahren wie Signifikanzanalysen (z.B. Spitzengruppe: +30%, Indifferente: +15% und Schlussgruppe: 0%)
- oder andere Verfahren (Welche?)

4. Abschließend: Welche weiteren Aspekte sind Ihnen im Rahmen dieses Themengebietes wichtig, die ich bisher evtl. noch nicht angesprochen habe?

Vielen Dank für das Gespräch.

## Anhang C:

### Hypothesen zur Analyse der Bestimmungsgründe beruflichen Erfolges anhand der Dresdner Absolventenstudien (vgl. Krempkow/ Pastohr 2004):

- H 1: Eine überdurchschnittliche Abiturnote erhöht die Berufschancen.
- H 2: Eine Berufsausbildung vor dem Studium erhöht die Berufschancen.
- H 3: Ein niedriges Studieneintrittsalter steigert die Berufschancen.
- H 4: Männer haben bessere Berufschancen als Frauen.<sup>278</sup>
- H 5: Eine „bessere“ soziale Herkunft der Absolventen verbessert deren Berufschancen.<sup>279</sup>
- H 6: Die Zugehörigkeit zur mathematisch-naturwissenschaftlichen oder ingenieurwissenschaftlichen Fächergruppe<sup>280</sup> verbessert die Berufschancen.<sup>281</sup>
- H 7: Praktische Erfahrungen erhöhen die Berufschancen.<sup>282</sup>
- H 8: Fachwechsel im Studium verringern die Berufschancen (da sie einen so genannten „Bruch“ im Lebenslauf darstellen).
- H 9: Auslandsaufenthalte erhöhen die Berufschancen.
- H 10a: Nebentätigkeiten während des Studiums und/oder in der Vorlesungszeit verringern die Berufschancen (da sie vom Studium ablenken).
- H 10b: Tätigkeiten als studentische Hilfskraft erhöhen die Berufschancen (da sie auf eine künftige wissenschaftliche Tätigkeit abzielen und dafür förderlich wirken).

---

<sup>278</sup> So haben Männer nach Müller (2002, S. 36) und Schomburg/ Teichler (1998: 148) ein höheres Durchschnittseinkommen. Außerdem sind Akademikerinnen häufiger von Arbeitslosigkeit, unsicherer und zunehmend inadäquater Beschäftigung betroffen als Akademiker (vgl. Schreyer 2001: 2225 – 2226, 2230 – 2231). Männliche Akademiker besetzen zudem häufiger eine Aufstiegsposition (vgl. auch Hartmann/ Kopp 2001: 449). Akademikerinnen sind hingegen – nicht zuletzt aufgrund der Familiensituation – öfter teilzeitbeschäftigt (vgl. Schomburg/ Teichler 1998: 147 – 148).

<sup>279</sup> Vgl. hierzu Müller 2002: 36.

<sup>280</sup> In der vorangegangenen Studie zu den Bestimmungsgründen beruflichen Erfolges von Krempkow/ Popp (2003) konnten zwar potenziell alle Fächergruppen, jedoch innerhalb der Fächergruppen nicht alle Fachrichtungen bzw. Fakultäten einbezogen werden. Die Daten der Fächergruppen waren folglich nicht mit allen vorhandenen Fachrichtungen bzw. Fakultäten belegt. Somit erschien es 2003 noch nicht sinnvoll, die Fächergruppen separat einzubeziehen. Dadurch, daß 2004 Ergebnisse zu allen Fächern bzw. Fakultäten vorliegen, ist dies nun möglich.

<sup>281</sup> Zwar lag die Arbeitslosenquote von Ingenieuren Mitte der 1990er Jahre aufgrund einer massiven Entlassungswelle überdurchschnittlich hoch, und auch in aktuellen Arbeitsmarktstatistiken gelten Ingenieure als größte Gruppe arbeitsloser Akademiker, jedoch ist die Zahl jüngerer erwerbsloser Ingenieure (unter 35 Jahren) niedrig und die „Ingenieurücke“ (und übrigens auch die von Mathematikern und Naturwissenschaftlern) in Deutschland paradoxerweise groß (vgl. Schreyer 2001, S. 2223 – 2224, Pastohr/ Wolter 2003: 124ff.). Ein für den höheren beruflichen Erfolg von Ingenieuren und Naturwissenschaftlern sprechendes Argument ist außerdem deren hoher Anteil an Führungspositionen und deren zu erwartendes (hohes) Bruttomonatsgehalt (vgl. Pfennig/ Renn/ Mack 2002:36ff., HIS 2000:100).

<sup>282</sup> Vgl. hierzu Teichler/ Schomburg 2001: 78.

- H 11: Sehr gute Studienbedingungen wirken sich positiv auf die Berufschancen aus (da sie die Erfolgswahrscheinlichkeit des Studiums erhöhen).
- H 12: Sehr gute Rahmenbedingungen des Studiums wirken sich positiv auf die Berufschancen aus (da sie die Studierenden im Studienleben entlasten und dadurch die Effizienz des Studiums erhöhen).
- H 13: Eine sehr gute Lehrqualität im Studium wirkt sich positiv auf die Berufschancen aus (da sie die Entwicklung der Fähigkeiten und Fertigkeiten der Studierenden begünstigt).
- H 14: Ein Zeitverlust im Studium aufgrund organisatorischer Regelungen vermindert die Berufschancen (da eine unangemessen über der Regelstudienzeit liegende Studiendauer mangelnde Kontinuität suggeriert).
- H 15: Wurden im Studium bestimmte Studienangebote vermisst, wirkt sich dies negativ auf die Berufschancen aus (da Kompetenzdefizite nachträglich beseitigt werden muss(t)en).
- H 16: Eine kurze Studiendauer erhöht die Berufschancen.<sup>283</sup>
- H 17: Ein niedriges Berufseintrittsalter erhöht die Berufschancen.
- H 18: Eine überdurchschnittliche Abschlussnote erhöht die Berufschancen.<sup>284</sup>
- H 19: Zusatzqualifikationen wie EDV- und Sprachkenntnisse erhöhen die Berufschancen.<sup>285</sup>
- H 20: Eine hohe Ausprägung beruflicher und sozialer Kompetenzen erhöht die Berufschancen.<sup>286</sup>
- H 21: Je höher die Anforderungen im Beruf sind, desto höher liegen die Berufschancen.
- H 22: Der erfolgreiche Abschluß eines Aufbau- und/oder Zweitstudiums/ einer Promotion sowie der Besuch von Weiterbildungsveranstaltungen verbessern die Berufschancen.
- H 23: Persönliche Beziehungen und Kontakte helfen beim Berufseinstieg.<sup>287</sup>
- H 24: Schwierigkeiten bei der Stellensuche verschlechtern den Berufseintritt und somit die Karrierechancen.<sup>288</sup>
- H 25: Berufsanfänger mit Partner sind erfolgreicher.
- H 26: Kinder sind hinderlich für Berufsanfänger.
- H 27: Berufliche und private Orientierungen sowie erfolgsorientierte Zukunftspläne beeinflussen die Berufschancen (z.B. sind karriereorientierte Absolventen erfolgreicher).<sup>289</sup>

---

<sup>283</sup> Vgl. hierzu Daniel 1995: 492 – 499.

<sup>284</sup> Vgl. hierzu Müller 2002: 93 – 95.

<sup>285</sup> Vgl. hierzu Teichler/ Schomburg 2001: 116.

<sup>286</sup> Vgl. hierzu Teichler/ Schomburg 2001: 116.

<sup>287</sup> Vgl. hierzu Haug/ Kropp 2002: 17.

## Anhang D:

### Ausgaben für zentrale Ergebnisse der Regresionsanalysen zu Determinanten beruflicher Zufriedenheit

#### Regressionsmodell nach Schaeper (1995)

##### Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,272 <sup>a</sup>	,074	,068	,76

a. Einflußvariablen : (Konstante), Besfristung

##### Koeffizienten<sup>b</sup>

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	2,967	,192		15,478	,000
	Besfristung	-,427	,118	-,272	-3,606	,000

a. Abhängige Variable: f5

##### Ausgeschlossene Variablen<sup>b</sup>

Modell		Beta In	T	Signifikanz	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	FACHGRP	-,024 <sup>a</sup>	-,317	,751	-,025	,998
	fstat	,044 <sup>a</sup>	,495	,621	,039	,732

a. Einflußvariablen im Modell: (Konstante), Besfristung

b. Abhängige Variable: f5

<sup>288</sup> Vgl. hierzu Schomburg/ Teichler 1998: 149.

<sup>289</sup> Nach Schomburg/ Teichler (1998: 153) ist beruflicher Erfolg nicht allein „von außen“ zu betrachten, sondern muß im Zusammenhang mit beruflichen Orientierungen und Zukunftsplänen betrachtet werden.

## Erweitertes Regressionsmodell in Anlehnung an Westermann u.a. (1996)

### Modellzusammenfassung<sup>b</sup>

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Durbin-Watson-Statistik
1	,656 <sup>a</sup>	,431	,367	,65	2,102

a. Einflußvariablen : (Konstante), Befristung, f6i, f6m, f6a, f2f, f6f, f2i, f2h, f6b, f6j, f6e, f6h, f6n, f6g, f6d

b. Abhängige Variable: f5

### Koeffizienten<sup>a</sup>

Modell	Nicht standardisierte Koeffizienten	Standardfehler	Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik		
						B	Beta	Toleranz
1	(Konstante)	-,219	,334					
	f2f	3,759E-02	,077	,035	,487	,627	,808	1,237
	f2h	4,752E-02	,102	,034	,467	,641	,815	1,226
	f2i	4,973E-02	,071	,052	,700	,485	,765	1,307
	f6a	,202	,092	,159	2,192	,030	,806	1,241
	f6b	8,670E-02	,071	,091	1,224	,223	,777	1,287
	f6d	5,996E-02	,056	,098	1,067	,288	,505	1,981
	f6e	2,335E-02	,056	,034	,416	,678	,649	1,541
	f6f	,160	,065	,202	2,457	,015	,629	1,591
	f6g	,197	,076	,220	2,587	,011	,587	1,704
	f6h	8,145E-03	,062	,011	,131	,896	,644	1,553
	f6i	6,283E-02	,057	,088	1,097	,275	,658	1,519
	f6j	2,487E-02	,068	,030	,367	,714	,651	1,535
	f6m	5,705E-02	,073	,066	,784	,435	,592	1,689
	f6n	2,529E-02	,069	,031	,364	,716	,600	1,667
	Befristung	,216	,145	,132	1,486	,140	,536	1,866

a. Abhängige Variable: f5

## Anhang E: Dokumentation der Fragebögen

1. Fragebogen zur Evaluation von Lehrveranstaltungen  
(TU Dresden, Standardversion 2000)
2. Fragebogen zum Studium und zum Übergang in den Beruf  
(TU Dresden, Dresdner Absolvent/innenstudie Maschinenwesen 2003)
3. Fragebogen für Befragung von Lehrenden zur Situation von Lehre und Studium  
(TU Dresden, Universität Frankfurt/ Oder und Hochschule Zittau-Görlitz 2001-2003)

# Fragebogen zur Evaluation von Lehrveranstaltungen unter Einbeziehung von Rahmenbedingungen

Philosophische  
Fakultät

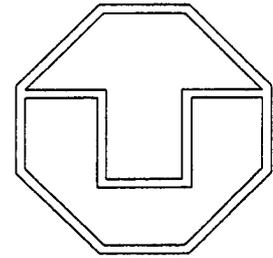
Lehrveranstaltung:..... WS/SS: .....

Bitte die jeweils zutreffende Antwort markieren!

Antwort-Felder bitte so ausmalen:



So bitte nicht:



Dieser Fragebogen dient der Rückmeldung an die Lehrenden über die Qualität der Lehrveranstaltung. Bitte beantworten Sie möglichst alle Fragen vollständig. Wenn Sie wirklich eine Frage nicht beantworten können, lassen Sie sie bitte frei.

## INHALTE UND PRÄSENTATION DER LEHRVERANSTALTUNG

	trifft zu ...		... trifft nicht zu		
	(1)	(2)	(3)	(4)	(5)
1. Die Ziele der Lehrveranstaltung sind klar erkennbar.	<input type="radio"/>				
2. Der Aufbau der Lehrveranstaltung ist gut nachvollziehbar.	<input type="radio"/>				
3. Die Relevanz der behandelten Themen für das Studium wird nahegelegt.	<input type="radio"/>				
4. Durch die Lehrveranstaltung habe ich viel gelernt.	<input type="radio"/>				
5. Mein Interesse an den Inhalten der Veranstaltung wurde gestärkt.	<input type="radio"/>				
6. Ich wurde zum Mitdenken motiviert.	<input type="radio"/>				
7. Der/die Lehrende wirkt immer gut vorbereitet.	<input type="radio"/>				
8. Der Vortragsstil des/der Lehrenden ist anregend.	<input type="radio"/>				
9. Die verwendeten Beispiele sind hilfreich.	<input type="radio"/>				
10. Der/die Lehrende kann komplizierte Sachverhalte verständlich machen.	<input type="radio"/>				
11. Der/die Lehrende steht für Rückfragen zur Verfügung.	<input type="radio"/>				
12. Der/die Lehrende nimmt die Lehre wichtig.	<input type="radio"/>				

Bitte beachten Sie: Die folgenden Fragen haben andere Antwortvorgaben!

	zu hoch	optimal		zu niedrig	
	(1)	(2)	(3)	(4)	(5)
13. Die Stoffmenge der Lehrveranstaltung ist:	<input type="radio"/>				
14. Die Anforderungen in der Veranstaltung sind:	<input type="radio"/>				
15. Das Tempo der/des Lehrenden ist:	<input type="radio"/>				

## STUDIERVERHALTEN UND MOTIVATION DER STUDIERENDEN

16. Wie oft haben Sie bisher in diesem Semester in dieser Lehrveranstaltung gefehlt?	keinmal	1x	2x	3x	mehr
	<input type="radio"/>				
17. Versuchen Sie bitte abzuschätzen, wie viele Stunden pro Woche an Vor- und Nacharbeit von Ihnen für diese Lehrveranstaltung aufzuwenden waren:	Null	bis 1	bis 2	bis 3	mehr
a) Wie viele Stunden haben Sie pro Woche im Schnitt investiert?	<input type="radio"/>				
b) Wie viele Stunden wären nach Ihrer Schätzung "eigentlich notwendig"?	<input type="radio"/>				
18. Mit welcher Note würden Sie Ihren gegenwärtigen Leistungsstand im Stoffgebiet dieser Lehrveranstaltung einschätzen?	1	2	3	4	5
	<input type="radio"/>				
19. Wie schätzen Sie das Studierverhalten Ihrer Kommilitonen ein?	trifft zu ...		... trifft nicht zu		
	(1)	(2)	(3)	(4)	(5)
a) Sie sind zu Lehrveranstaltungen pünktlich.	<input type="radio"/>				
b) Sie sind in Lehrveranstaltungen aufmerksam.	<input type="radio"/>				
c) Sie sind an Lehrveranstaltungen interessiert.	<input type="radio"/>				
d) Sie nehmen aktiv an Lehrveranstaltungen teil.	<input type="radio"/>				

20. Aus welchen Gründen besuchen Sie diese Lehrveranstaltung (LVA)?  
 (Mehrfachantworten möglich - Bitte alle Angaben ankreuzen, die für Sie zutreffen)

- |                                    |                       |   |                       |
|------------------------------------|-----------------------|---|-----------------------|
| a) Pflichtveranstaltung            | <input type="radio"/> | g) allgemeine Auffrischung/ Wiederholung        | <input type="radio"/> |
| b) Wahlpflichtveranstaltung        | <input type="radio"/> | h) aus persönlichem Interesse am Stoff          | <input type="radio"/> |
| c) studium generale                | <input type="radio"/> | i) wegen der Lehrperson in dieser LVA           | <input type="radio"/> |
| d) wegen Leistungsnachweis/Klausur | <input type="radio"/> | j) Veranstaltung ist zeitlich/ räumlich günstig | <input type="radio"/> |
| e) Stoff ist prüfungsrelevant      | <input type="radio"/> | k) andere Besuchsgründe                         | <input type="radio"/> |
| f) allgemein wichtig fürs Studium  | <input type="radio"/> | Welche? .....                                   |                       |

**RÄUMLICHE UND ORGANISATORISCHE RAHMENBEDINGUNGEN**

	trifft zu ...		... trifft nicht zu		
	(1)	(2)	(3)	(4)	(5)
21. Die Raumgröße ist der Teilnehmerzahl angemessen.	<input type="radio"/>				
22. Die Akustik in diesem Raum ist günstig.	<input type="radio"/>				
23. Die Sichtbedingungen in diesem Raum sind günstig.	<input type="radio"/>				
24. Der/die Lehrende bietet ausreichend Sprechstunden an.	<input type="radio"/>				
25. Der/die Lehrende ist zu den Sprechstunden anwesend.	<input type="radio"/>				
26. Ich kann mich auch außerhalb der Sprechzeiten beraten lassen.	<input type="radio"/>				
27. Ich bin mit der Beratung u. Betreuung der Lehrenden insgesamt zufrieden.	<input type="radio"/>				
28. Die für die Lehrveranst. notwendige Literatur ist in der Bibliothek verfügbar.	<input type="radio"/>				
29. Bei Bedarf bekomme ich in der Bibliothek einen Platz zum Arbeiten.	<input type="radio"/>				
30. Bei Bedarf bekomme ich im PC-Pool einen Platz zum Arbeiten.	<input type="radio"/>				
31. Ich kann im Studiengang ausreichend eigene Interessenschwerpkt. setzen.	<input type="radio"/>				
32. In meinem Studiengang werden genügend Lehrveranstaltungen angeboten.	<input type="radio"/>				

**KOMMENTAR/ GESAMTEINSCHÄTZUNG:** Was empfanden Sie als besonders gut bzw. schlecht?  
 Welche Verbesserungsvorschläge hätten Sie?

.....

.....

.....

.....

.....

**ANGABEN ZUR PERSON:** In welchem Studiengang (Hauptfach) studieren Sie? .....

Ihr Geschlecht: weiblich  männlich    

Fachsemester: 1. und 2.  3. und 4.  5. und 6.  7. und 8.  9. und 10.  über 10.

Welchen Abschluß streben Sie an? Diplom  Staatsexamen  Magister  Bachelor  Sonstiges

# DRESDNER ABSOLVENT/INNENSTUDIE

Fragebogen zum Studium und zum Übergang in den Beruf

## Hinweise zum Ausfüllen des Fragebogens

- Bitte beantworten Sie die Fragen in der vorgegebenen Reihenfolge und kreuzen Sie die entsprechende Antwort an bzw. füllen Sie möglichst leserlich die vorgesehenen Textfelder aus. Verwenden Sie hierzu bitte nur Kugelschreiber oder Füllfederhalterstift!
- In den meisten Fällen ist ein Kästchen anzukreuzen. Falls mehrere Kästchen angekreuzt werden können, ersehen Sie dies aus dem Hinweis: "Mehrfachnennungen möglich"!
- Falls Sie versehentlich das falsche Kästchen angekreuzt haben: schwärzen Sie bitte die falsch Markierung und markieren Sie das richtige Kästchen! Kommentare zum Fragebogen sind uns willkommen! Bitte verwenden Sie hierfür ein gesondertes Blatt
- Um bei Wiederholungsbefragungen langfristige Entwicklungen nachvollziehen zu können, ist es notwendig, Ihre Angaben Ihrer Person zuordnen zu können, wobei trotzdem die Anonymität gewahrt werden muss. Deshalb bitten wir Sie, den Fragebogen mit einer persönlichen Verschlüsselung zu versehen. Hierfür tragen Sie bitte den ersten Buchstaben des Vornamens Ihrer Mutter, dann Ihr Geburtsjahr, danach den zweiten Buchstaben des Vornamens Ihres Vaters und schließlich sein Geburtsjahr in die nachfolgenden Kästchen ein. (Wenn Mutter oder Vater unbekannt, verwenden Sie die Buchstaben Ihres Namens und Ihr Geburtsjahr.)

## A - Vor dem Studium

### A - 1 Persönliche, anonyme Verschlüsselung des Fragebogens

Beispiel: Mutter: 

D
---

4	4
---	---

 Vater: 

E
---

4	3
---	---

nun Ihre Angaben: Mutter: 

--

--	--

 Vater: 

--

--	--

### A - 2 Welche Durchschnittsnote hatten Sie in dem Schulzeugnis, mit dem Sie Ihre Studienberechtigung erworben haben?

Durchschnittsnote: 

--

 , 

--

 (Bitte eintragen!)

### A - 3 Haben Sie vor dem Studium eine berufliche Ausbildung begonnen?

- Ja, und abgeschlossen  
 Ja, aber nicht abgeschlossen  
 Nein

## B - Verlauf des Studiums

### B - 1 Bitte informieren Sie uns über Ihr an der TU Dresden zuletzt abgeschlossenes, grundständiges Studium!

Beginn/Ende (Monat, Jahr)	Studiengang/ Studienfächer	Art des höchsten Abschlusses	Art der abgelegten Prüfung/ Note								
von Monat <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td><td> </td></tr></table> Jahr <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td><td> </td></tr></table>					Fach ----- ggf. Studienrichtung/ Vertiefung	<input type="checkbox"/> Bachelor <input type="checkbox"/> Magister/ Master <input type="checkbox"/> Diplom <input type="checkbox"/> Staatsexamen <input type="checkbox"/> Anderer Abschluss z. B. Promotion	<u>Zwischenprüfung/ im Hauptfach Vordiplom</u>  Note: <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td></tr></table> , <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td></tr></table> <u>Abschlussprüfung</u>  Gesamtnote: <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td></tr></table> , <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td></tr></table>				
bis Monat <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td><td> </td></tr></table> Jahr <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td><td> </td></tr></table>					ggf. Studienrichtung/ Vertiefung						

--	--	--	--

**B - 2** Wieviele Semester haben Sie insgesamt an Hochschulen studiert? (Hochschulsemester) Wieviele Semester haben Sie das (Haupt-) Fach studiert, das Sie an der TU Dresden abgeschlossen haben? (Fachsemester)

Hochschulsemester:  (Bitte eintragen!); Fachsemester:  (Bitte eintragen!)  
 (Wie lange insgesamt an Hochschulen?) (Fachstudiendauer laut letztem Studentenausweis/Zeugnis)

**B - 2a** Haben Sie in Ihrem Studium Zeit verloren, weil Sie wegen organisatorischer Regelungen (z.B. Lehrangebot) Leistungsnachweise nicht erwerben oder nicht bestandene Klausuren, Prüfungen o. ä. erst später wiederholen konnten?

nein  ja, 1 Semester  ja, 2 Semester  ja, mehr als 2 Semester

**B - 3** Haben Sie nach der Erstimmatrikulation an der Hochschule Ihr/ en Studiengang/ -fach gewechselt?

Ich habe nicht gewechselt .....**Bitte weiter mit B - 5!**  Ich habe gewechselt .....**Bitte weiter mit B - 4!**  
 Woher?.....

**B - 4** Inwieweit haben die folgenden Gründe den Wechsel bewirkt? (Kreuzen Sie bitte jede Zeile an!)

	sehr stark				gar nicht
	1	2	3	4	5
Schwierigkeiten bei der Bewältigung des Lehrstoffes/ von Prüfungen	<input type="checkbox"/>				
Schlechte Betreuung durch Lehrende	<input type="checkbox"/>				
Schlechte Lehrqualität	<input type="checkbox"/>				
Eingeschränktes/ uninteressantes Lehrangebot	<input type="checkbox"/>				
Schlechte Berufsaussichten	<input type="checkbox"/>				
Falsche Vorstellungen/ mangelnde Informationen über d. Studiengang	<input type="checkbox"/>				
Aufnahme des ursprünglich gewünschten Studiums	<input type="checkbox"/>				
Änderung meines Berufsziels	<input type="checkbox"/>				
Andere Gründe (Welche: ..... )	<input type="checkbox"/>				

**B - 5** Waren Sie aus studienbezogenen Gründen im Ausland?

Ja  Nein.....**Bitte weiter mit B - 7!** insgesamt:  Monate (Bitte eintragen!)

**B - 6** Welchen Zweck hatte Ihr Aufenthalt/ hatten Ihre Aufenthalte? (Mehrfachnennungen möglich!)

Sprachkurs  Auslandssemester  Praktikum  Sonstiges (Was?.....)

**B - 7** Haben Sie neben Ihrem Studium ein Praktikum bzw. mehrere Praktika absolviert (keine Laborpraktika bzw. Lehrveranstaltungen) ? (Bitte Anzahl eintragen!)

Nein, ich habe kein Praktikum absolviert.....**Bitte weiter mit B - 9!**  
 Ja, ich habe  Pflichtpraktika absolviert. Gesamtdauer?  Monate  
 Ja, ich habe  **freiwillige** Praktika absolviert. Gesamtdauer?  Monate

**B - 8 Bei welcher Einrichtung/ welchen Einrichtungen haben Sie Ihr Praktikum/ Ihre Praktika absolviert?**  
(Mehrfachnennungen möglich!)

- |  |  |
|--|--|
| <input type="checkbox"/> Universität, Fachhochschule           | <input type="checkbox"/> Medien (Rundfunk, TV, Presse) |
| <input type="checkbox"/> Forschungsinstitut                    | <input type="checkbox"/> Kultur, Bibliothek, Archiv    |
| <input type="checkbox"/> Öffentliche Verwaltung, Behörde       | <input type="checkbox"/> Schule/ Bildungseinrichtung   |
| <input type="checkbox"/> Museum                                | <input type="checkbox"/> Gesundheitswesen              |
| <input type="checkbox"/> Partei, Verband, Kirche, Verein       | <input type="checkbox"/> Sozialarbeit/Jugendhilfe      |
| <input type="checkbox"/> Gewerbe, Industrie, Handel            | <input type="checkbox"/> Sonstige Dienstleistungen     |
| <input type="checkbox"/> Bank, Versicherung                    |  |
| <input type="checkbox"/> Sonstige Einrichtung, und zwar: ..... | (Ggf. bitte eintragen!)                                |

**B - 9 Wie häufig haben Sie in den Semesterferien und im laufenden Semester gearbeitet?**

	nie	nur gelegentlich	häufig	immer
In den Semesterferien	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In der Vorlesungszeit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**B - 10 Waren Sie im Verlauf Ihres Studiums als studentische Hilfskraft tätig?**

- |   |  |
|---|--|
| <input type="checkbox"/> Nein                     | <input type="checkbox"/> Ja, zwischen 1 und 2 Jahren |
| <input type="checkbox"/> Ja, weniger als ein Jahr | <input type="checkbox"/> Ja, länger als 2 Jahre      |

**B - 11 Wie fundiert waren Ihre Sprachkenntnisse zum Zeitpunkt des Studienabschlusses?**  
(Bitte tragen Sie die entsprechende Fremdsprache ein und kreuzen Sie jede Zeile an!)

	keine Kenntnisse	Grundkenntnisse	gute Kenntnisse	sehr gute Kenntnisse
Russisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Englisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Französisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Spanisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bitte eintragen!

**B - 12 Wie fundiert waren Ihre EDV-Kenntnisse beim Studienabschluss? (Kreuzen Sie bitte jede Zeile an!)**

	keine Kenntnisse	Grundkenntnisse	gute Kenntnisse	sehr gute Kenntnisse
Textverarbeitung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Internetnutzung ("Browser", E-Mail etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Programmiersprachen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Andere, fachspezifische Programme (Welche: ..... )	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**B - 13 Haben Sie sich vor dem Studium an der TU Dresden auch an anderen Universitäten oder Fachhochschulen um einen Studienplatz beworben?**

- ja, an welcher/n? .....  nein

**B - 14 Warum haben Sie sich für Dresden entschieden? (Bitte eintragen!)**

-----

## C - Beurteilung des Studiums

**C - 1** Wie beurteilen Sie die Studienbedingungen in dem (Haupt-) Fach, in dem Sie Ihren Studienabschluss gemacht haben, hinsichtlich der folgenden Aspekte? (Kreuzen Sie bitte jede Zeile an!)

	sehr gut			sehr schlecht	
	1	2	3	4	5
Aufbau und Struktur Ihres Studienganges	<input type="checkbox"/>				
Transparenz der Prüfungsanforderungen	<input type="checkbox"/>				
Breite des Lehrangebots	<input type="checkbox"/>				
Fachliche Spezialisierungs- und Vertiefungsmöglichkeiten	<input type="checkbox"/>				
Zeitliche Koordination des Lehrangebotes	<input type="checkbox"/>				
Forschungsbezug der Lehre	<input type="checkbox"/>				
Praxisbezug der Lehre	<input type="checkbox"/>				
Zugang zu/ Vermittlung von Praktika	<input type="checkbox"/>				
Beziehungen zwischen Studierenden und Lehrenden	<input type="checkbox"/>				
Fachliche Beratung und Betreuung durch die Lehrenden	<input type="checkbox"/>				
Ausreichende Anzahl von Plätzen in Lehrveranstaltungen	<input type="checkbox"/>				

**C - 2** Wie beurteilen Sie die folgenden Rahmenbedingungen Ihres Studiums an der TU Dresden? (Kreuzen Sie bitte jede Zeile an!)

	sehr gut			sehr schlecht	
	1	2	3	4	5
Qualität der Verwaltung (Imma-Amt, Raumplanung usw.)	<input type="checkbox"/>				
Qualität der Bibliothek	<input type="checkbox"/>				
Zugang zu EDV-Diensten (Internet, E-Mail, Datenbanken)	<input type="checkbox"/>				
Einflussmöglichkeiten der Studierenden an der Hochschule	<input type="checkbox"/>				
Raumsituation in Ihrem Fach	<input type="checkbox"/>				
Technische Ausstattung in Ihrem Fach	<input type="checkbox"/>				
Klima unter den Studierenden	<input type="checkbox"/>				
Angebot der Mensen und Cafeterias	<input type="checkbox"/>				
Kulturelles Angebot der Stadt Dresden	<input type="checkbox"/>				
Wohnsituation in Dresden	<input type="checkbox"/>				
Öffentliche Verkehrsmittel in Dresden	<input type="checkbox"/>				

**C - 3** Wie beurteilen Sie die folgenden Aspekte der Lehre in dem Fach, in dem Sie Ihren Studienabschluss gemacht haben? (Kreuzen Sie bitte jede Zeile an!)

	sehr gut			sehr schlecht	
	1	2	3	4	5
Vermittlung von Fachwissen	<input type="checkbox"/>				
Erlernen sozialer/kommunikativer Fähigkeiten	<input type="checkbox"/>				
Einübung in selbständiges Lernen/Arbeiten	<input type="checkbox"/>				
Einübung in wissenschaftliche Arbeitsweisen	<input type="checkbox"/>				
Einübung in mündliche Präsentation	<input type="checkbox"/>				
Einübung in schriftliche Ausdrucksfähigkeit	<input type="checkbox"/>				

**C - 4** Haben Sie in Ihrem Studium in Dresden bestimmte Studienangebote vermisst?

ja                       nein

Wenn ja, welche? .....

**C - 5 Wie beurteilen Sie rückblickend den Wert Ihres Studiums hinsichtlich...**

(Kreuzen Sie bitte jede Zeile an!)

	sehr gut				sehr schlecht
	1	2	3	4	5
... der Nützlichkeit des Hochschulabschlusses für berufliche Karriere?	<input type="checkbox"/>				
... der Verwendbarkeit der Studieninhalte im Beruf?	<input type="checkbox"/>				
... der Möglichkeit, sich persönlich zu entwickeln?	<input type="checkbox"/>				
... der Möglichkeit, den studentischen Freiraum zu genießen?	<input type="checkbox"/>				

**C - 6 Wie würden Sie sich hinsichtlich Ihres bisherigen Werdegangs aus heutiger Sicht verhalten?**

- Ich würde kein Abitur mehr machen.....**Bitte weiter mit D!**
- Ich würde das Abitur machen, aber nicht mehr studieren.....**Bitte weiter mit D!**
- Ich würde nach dem Abitur zunächst eine Berufsausbildung absolvieren und dann studieren.
- Ich würde gleich nach dem Abitur studieren.
- Ich würde vor dem Abitur eine Berufsausbildung machen.
- Ich würde eine Berufsausbildung mit Abitur machen.

**C - 7 Welches Fach würden Sie dann voraussichtlich studieren?**

- Ein anderes Fach an einer anderen Hochschule
- Ein anderes Fach an derselben Hochschule
- Das gleiche Fach an einer anderen Hochschule
- Das gleiche Fach an derselben Hochschule

**D - Nach dem Studium**

**D - 1 Bitte tragen Sie in der folgenden Tabelle Ihre erste und Ihre derzeit aktuelle Beschäftigung seit Ihrem Studienabschluss ein. (Nennen Sie ggf. bitte auch Honorararbeit, Weiterqualifizierung, Umschulung oder Fortbildung, Zeiten der Erwerbslosigkeit (z.B. Hausfrau/-mann, auch Weltreise o.ä.).)**

**ERSTE BESCHÄFTIGUNG**

<p><b>Beginn/Ende</b></p> <p>von Monat <input type="text"/> <input type="text"/></p> <p>Jahr <input type="text"/> <input type="text"/></p> <p>bis Monat <input type="text"/> <input type="text"/></p> <p>Jahr <input type="text"/> <input type="text"/></p>	<p><b>Beschäftigungsbezeichnung:</b></p> <p>-----</p> <p><b>Beschäftigungsstatus:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p> <p><b>Beschäftigungsform:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p> <p><b>Beschäftigungsbereich:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p>	<p><b>tatsächliche Wochenarbeitszeit:</b></p> <p>in Stunden <input type="text"/> <input type="text"/></p> <p><b>Bruttomonatseinkommen:</b> incl. Zulagen in € (unbezahlt = 0 €) <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/></p> <p><b>Befristung:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein</p>
---	--	--

**AKTUELLE BZW. LETZTE BESCHÄFTIGUNG (wenn andere als erste Beschäftigung)**

<p><b>Beginn/Ende</b></p> <p>von Monat <input type="text"/> <input type="text"/></p> <p>Jahr <input type="text"/> <input type="text"/></p> <p>bis Monat <input type="text"/> <input type="text"/></p> <p>Jahr <input type="text"/> <input type="text"/></p>	<p><b>Beschäftigungsbezeichnung:</b></p> <p>-----</p> <p><b>Beschäftigungsstatus:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p> <p><b>Beschäftigungsform:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p> <p><b>Beschäftigungsbereich:</b> <input type="checkbox"/> (Code-Nr. siehe Seite 6)</p>	<p><b>tatsächliche Wochenarbeitszeit:</b></p> <p>in Stunden <input type="text"/> <input type="text"/></p> <p><b>Bruttomonatseinkommen:</b> incl. Zulagen in € (unbezahlt = 0 €) <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/></p> <p><b>Befristung:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein</p>
---	--	--

**D - 2 Wenn Sie zwischen der ersten und der letzten Tätigkeit noch andere Tätigkeiten ausgeübt haben, tragen Sie diese bitte in die folgende Tabelle ein! (Bei keiner Veränderung bitte freilassen!)**

Monate nach Abschluss	Beschäftigungsbezeichnung (Bei mehreren Tätigkeiten die überwiegend Ausgeübte!)	Beschäftigungsstatus (Code-Nr. siehe unten)	Beschäftigungsform (Code-Nr. siehe unten)	Beschäftigungsber. (Code-Nr. siehe unten)	Wochenarbeitszeit (Stunden)	Bruttomonats-einkommen incl. Zulagen in € (unbezahlt = 0 €)	Befristung
3	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
6	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
12	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
24	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
36	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
48	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein
60	-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> ja <input type="checkbox"/> nein

**Liste des Beschäftigungsstatus:**

a = Ungelernter/ Angelernter Arbeiter  
 b = Facharbeiter, Vorarbeiter, Meister  
 c = Angestellter mit einfacher Tätigkeit (z.B. Verkäufer)  
 d = Angestellter mit qualifizierter Tät. (z.B. Sachbearbeiter)  
 e = Angestellter mit hochqualifizierter Tät. (z.B. Prokurist)  
 f = Angestellter mit Führungsaufgaben (z.B. Geschäftsführer)  
 g = Beamter  
 h = Selbständiger/ Freiberufler  
 i = Hausfrau/ Hausmann  
 k = Anderes (z.B. arbeitslos, Jobben, ...)

**Liste der Beschäftigungsformen:**

1 = Reguläre Erwerbstätigkeit  
 2 = Honorartätigkeit  
 3 = Referendariat u.ä.  
 4 = Aus-, Weiterbildung, Umschulung usw.  
 5 = Weiterstudium oder Promotion  
 6 = Praktikum  
 7 = Jobben (ohne feste Anstellung)  
 8 = arbeitslos/ arbeitssuchend  
 9 = Bund/ Zivi, Erziehungsurlaub  
 10 = Sonstiges (z.B. Reise, länger krank)

**Liste der Beschäftigungsbereiche:**

A = Hochschule und Forschungsinstitute  
 B = Öffentliche Verwaltung, Behörde  
 C = (Privat-) Wirtschaft  
 D = Organisationen ohne Erwerbscharakter (Kirchen, Vereine, Verbände)  
 E = arbeite nicht  
 F = sonstige Bereiche

**D - 3 Wieviele Mitarbeiter hat(te) das Unternehmen bzw. die Einrichtung, in der Sie in Ihrer ersten Tätigkeit arbeit(et)en am Standort ungefähr?**

- 1 - 10       11 - 100       101 - 1000       1001 - 10000       über 10000

**E - Zusatzqualifikation/ Weiterbildung**

**E - 1 Haben Sie nach Ihrem Studienabschluss ein weiteres Studium oder eine Promotion begonnen?**

(Mehrfachnennungen möglich!)

Nein

**Begonnen:**

- Ja, eine Promotion begonnen  
 Ja, ein Aufbaustudium begonnen  
 Ja, ein Zweitstudium begonnen  
 Ja, sonstiges: -----

**Abgeschlossen:**

- Ja, eine Promotion abgeschlossen  
 Ja, ein Aufbaustudium abgeschlossen  
 Ja, ein Zweitstudium abgeschlossen

(Ggf. bitte eintragen!)

**E - 2 Haben Sie seit Ihrem Studienabschluss an einer oder mehreren Weiterbildungsveranstaltung/en (Kurse, Workshops etc.) - bezogen auf Ihre derzeitige oder angestrebte berufliche Karriere - teilgenommen?**

- Nein .... **Bitte weiter mit E - 5!**       Ja, an einer       Ja, an mehreren

**E - 3 Welche Inhalte wurden in den von Ihnen besuchten Veranstaltungen vermittelt?**  
(Mehrfachnennungen möglich!)

- |  |  |
|--|--|
| <input type="checkbox"/> Neues Wissen aus meinem Fachgebiet                        | <input type="checkbox"/> Betriebswirtschaftliches Wissen             |
| <input type="checkbox"/> Disziplinübergreifendes Fachwissen verschiedener Bereiche | <input type="checkbox"/> Management-/Führungskompetenzen             |
| <input type="checkbox"/> Methodische Kompetenzen                                   | <input type="checkbox"/> Rechtliche Kenntnisse                       |
| <input type="checkbox"/> Fremdsprachenkompetenzen                                  | <input type="checkbox"/> Kommunikations- oder Präsentationstechniken |
| <input type="checkbox"/> Computerfertigkeiten                                      | <input type="checkbox"/> Umgang mit Kunden/Klienten                  |
|  | <input type="checkbox"/> Anderes, und zwar: (Ggf. bitte eintragen!)  |
- .....

**E - 4 Inwieweit hat Ihnen die Weiterbildung tatsächlich geholfen ...**  
(Bitte jede Zeile ankreuzen!)

	sehr stark					gar nicht
	1	2	3	4	5	
... eine Erwerbstätigkeit zu finden?	<input type="checkbox"/>					
... mit den Arbeitsanforderungen zurecht zu kommen?	<input type="checkbox"/>					
... den Rahmen Ihrer Tätigkeiten zu erweitern?	<input type="checkbox"/>					
... Ihre berufliche Position zu verbessern?	<input type="checkbox"/>					
... Sonstiges: .....	<input type="checkbox"/>					

**E - 5 Kennen Sie den Weiterbildungskatalog der TU Dresden?**

- ja  nein

**E - 5a Haben Sie in diesem Rahmen schon einmal an einer Weiterbildung an der TU Dresden teilgenommen?**

- nein  ja, einmal  ja, mehrfach

**E - 6 Inwieweit sind Sie an einem weiteren Kontakt zur TU Dresden interessiert?**

- sehr stark     gar nicht

**E - 7 Bitte nehmen Sie zu folgenden Aussagen Stellung!**

	trifft zu				trifft nicht zu
Ich würde einem Absolventenverein meines Fachbereiches beitreten.	<input type="checkbox"/>				
Ich würde einem Absolventenverein der TU Dresden (fächerübergreifend) beitreten.	<input type="checkbox"/>				
Ich würde eine kostenlose Absolventenzeitung abonnieren.	<input type="checkbox"/>				
Sonstiges (Bitte eintragen!): .....	<input type="checkbox"/>				

**F - Stellensuche**

**F - 1 Ab wann haben Sie versucht, eine Beschäftigung zu finden?**

- Noch gar nicht
- Bereits frühzeitig während des Studiums
- Während der Zeit meines Studienabschlusses
- Ca.   Monate nach meinem Studienabschluss (Ggf. bitte Anzahl eintragen!)

**F - 2 Was haben Sie unternommen, um nach dem Studium eine Beschäftigung zu finden?**  
(Mehrfachnennungen möglich!)

- (01)  Bewerbung auf Stellenausschreibungen
- (02)  Initiativbewerbung/ Blindbewerbung
- (03)  Suche über Arbeitsamt
- (04)  Stellensuche über Internet
- (05)  Nutzung der Kontakte zu Lehrenden der TU Dresden
- (06)  Nutzung der Kontakte während eines Jobs/ Praktikums im Studium
- (07)  Nutzung der Kontakte während einer Ausbildung/ Tätigkeit vor Studium
- (08)  Versuch, mich selbständig zu machen
- (09)  Nutzung persönlicher Kontakte (Eltern, Freunde, Verwandte, etc.)
- (10)  Absolventenmessen (HS-Kontaktbörse)
- (11)  Private Vermittlungsagentur
- (12)  Sonstiges: ----- (Ggf. bitte eintragen!)

**F - 3 Welche Vorgehensweise führte zu Ihrer ersten Beschäftigung?**  
(Bitte tragen Sie die entsprechende Nummer aus Frage F - 2 ein!)

- Erfolgreiche Strategie: 

--	--

 (Ggf. bitte eintragen!)
- Habe noch keine Beschäftigung gefunden

**F - 4 Bei wievielen Arbeitgebern haben Sie sich beworben?**

Zahl der Bewerbungen: 

--	--

 (Bitte eintragen!)

**F - 5 Wie oft wurden Sie zu Vorstellungsgesprächen eingeladen?**

Zahl der Gespräche: 

--	--

 (Bitte eintragen!)

**F - 6 Wieviele Monate nach Ihrem Studienabschluss haben Sie aktiv gesucht, bis Sie Ihre erste Stelle fanden?**

Monate: 

--	--

 (Bitte eintragen!)

**F - 7 Welche Schwierigkeiten sind Ihnen bei Ihrer Stellensuche - unabhängig von deren Erfolg - bislang begegnet? (Mehrfachnennungen möglich!)**

- Ich habe bisher keine Probleme gehabt. ...**Bitte weiter mit G - 1!**
- Es wurden meist Absolventen mit einem anderen Schwerpunkt gesucht.
- Oft wurde ein anderer Studienabschluss verlangt (z.B. Promotion statt Diplom, FH- statt Uni-Abschluß, etc.).
  
- Die angebotene(n) Stelle(n) entsprach(en) nicht meinen Gehaltsvorstellungen.
- Die angebotene(n) Stelle(n) entsprach(en) nicht meinen Vorstellungen über Arbeitszeit und/ oder Arbeitsbedingungen.
- Es wurden überwiegend Bewerber mit Berufserfahrung gesucht.
  
- Die angebotene(n) Stelle(n) war(en) zu weit entfernt.
- Es wurden spezielle Kenntnisse verlangt, die ich nicht habe (z.B. EDV, Fremdsprachen etc.).
- Die angebotene(n) Stelle(n) entsprach(en) nicht meinen inhaltlichen Vorstellungen.
  
- Die angebotene(n) Stelle(n) ließen sich nicht mit der Familie vereinbaren.
- Andere Probleme, und zwar -----

(Ggf. bitte eintragen!)

## G - Berufsstart

Falls Sie berufstätig sind bzw. nach dem Studienabschluss schon berufstätig waren .....weiter mit G - 1  
 Falls Sie seit Ihrem Studienabschluss noch nicht berufstätig waren.....weiter mit H - 1

### G- 1 Wie wichtig waren Ihrer Meinung nach die folgenden Aspekte für Ihren Arbeitgeber, Sie zu beschäftigen? (Bitte jede Zeile ankreuzen!)

	sehr wichtig				gar nicht wichtig
	1	2	3	4	5
Mein Studiengang	<input type="checkbox"/>				
Meine fachliche Spezialisierung	<input type="checkbox"/>				
Thema meiner Abschlussarbeit (wenn zutreffend)	<input type="checkbox"/>				
Meine Examensnote	<input type="checkbox"/>				
Praxiserfahrungen	<input type="checkbox"/>				
Meine Fremdsprachenkenntnisse	<input type="checkbox"/>				
Meine persönlichen Beziehungen	<input type="checkbox"/>				
Meine Computer- Kenntnisse	<input type="checkbox"/>				
Meine sozialen Kompetenzen (sprachliche Gewandtheit, Auftreten)	<input type="checkbox"/>				
Meine Bereitschaft, zum Wohnortwechsel (auch ins Ausland)	<input type="checkbox"/>				
Ruf der Hochschule, an der ich studierte	<input type="checkbox"/>				
Meine Auslandserfahrungen	<input type="checkbox"/>				
Ein zügiges Studium	<input type="checkbox"/>				
Meine Bereitschaft zu finanziellen Abstrichen	<input type="checkbox"/>				
Empfehlungen/ Referenzen von Dritten	<input type="checkbox"/>				
Mein Geschlecht	<input type="checkbox"/>				
Meine Weltanschauung (z.B. Parteimitgliedschaft, Religion)	<input type="checkbox"/>				

### G - 2 Wo ist Ihr gegenwärtiger Arbeitsplatz bzw. wo war Ihr letzter Arbeitsplatz? (Bitte nur eine Angabe!)

- |   |  |
|---|--|
| <input type="checkbox"/> Dresden                  | <input type="checkbox"/> Alte Bundesländer |
| <input type="checkbox"/> Sachsen (ohne Dresden)   | <input type="checkbox"/> Ausland           |
| <input type="checkbox"/> Anderes neues Bundesland |  |

### G - 3 Wenn Sie Ihre Stelle schon einmal gewechselt haben, was waren Ihre Gründe, den Arbeitsplatz zu wechseln? (Mehrfachnennungen möglich!)

- |  |  |
|--|--|
| <input type="checkbox"/> Habe noch nicht gewechselt          | <input type="checkbox"/> Befristeter Vertrag                   |
| <input type="checkbox"/> Interessantere Aufgabe              | <input type="checkbox"/> Wegfall der Stelle                    |
| <input type="checkbox"/> Eigenständigeres Arbeiten           | <input type="checkbox"/> Kündigung durch den Arbeitgeber       |
| <input type="checkbox"/> Höheres Einkommen                   | <input type="checkbox"/> Zu hohe Anforderungen                 |
| <input type="checkbox"/> Bessere Aufstiegschancen            | <input type="checkbox"/> Zu geringe Anforderungen              |
| <input type="checkbox"/> Besseres Betriebsklima              | <input type="checkbox"/> Wunsch nach Ortswechsel               |
| <input type="checkbox"/> Weiterqualifikationsmöglichkeit(en) | <input type="checkbox"/> Unvereinbarkeit von Beruf und Familie |
| <input type="checkbox"/> Vorher nur Übergangslösung          | <input type="checkbox"/> Andere persönliche Gründe             |
| <input type="checkbox"/> Sonstige Gründe: .....              | (Ggf. bitte eintragen)   |

**G - 4 Inwieweit traten bei Ihrem Berufsstart folgende Probleme auf? (Bitte jede Zeile ankreuzen!)**

	sehr stark					gar nicht
	1	2	3	4	5	
Hektik im Beruf, Termindruck, Arbeitsüberlastung	<input type="checkbox"/>					
Undurchschaubarkeit betrieblicher Entscheidungsprozesse	<input type="checkbox"/>					
Gefühl der Unterforderung	<input type="checkbox"/>					
Fühlte mich nicht qualifiziert genug	<input type="checkbox"/>					
Mangel an Kooperation unter den Kolleg/innen	<input type="checkbox"/>					
Probleme mit Vorgesetzten	<input type="checkbox"/>					
Vereinbarkeit von Beruf und Familie	<input type="checkbox"/>					
Wenig Feedback über die geleistete Arbeit	<input type="checkbox"/>					
Schwierigkeiten mit bestimmten beruflichen Normen (z.B. geregelte Arbeitszeit, Kleidung, Betriebshierarchie)	<input type="checkbox"/>					

**H - Berufliche Kompetenzen, Anforderungen und Zufriedenheit**

**H - 1 Bitte geben Sie an, in welchem Maße Sie über die folgenden Kompetenzen bei Studienabschluss verfügt haben und in welchem Maße diese auf Ihrer jetzigen/letzten Stelle gefordert sind bzw. waren! (Falls Sie bisher nicht erwerbstätig waren, füllen Sie bitte nur die linke Spalte aus!)**

Kompetenzen bei Studienabschluss					jetzt geforderte Kompetenzen					
in hohem Maße		gar nicht			in hohem Maße		gar nicht			
1	2	3	4	5	1	2	3	4	5	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Fachkenntnisse.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Lernfähigkeit.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Kreativität.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Zeiteinteilung.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Kritisches Denken.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Ausdrucksfähigkeit.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Breites Allgemeinwissen.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Fächerübergreifendes Denken....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Fremdsprachen.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....EDV- Fertigkeiten.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Planen, koordinieren und organisieren.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Wirtschaftliches Denken.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Verhandlungsgeschick.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Teamarbeit.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	.....Anpassungsfähigkeit.....	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	..Andere führen/ Mitarbeiterführung..	<input type="checkbox"/>				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	----- Sonstiges	<input type="checkbox"/>				

**H - 2 Wie zufrieden sind Sie alles in allem mit Ihrer beruflichen Situation?**

sehr zufrieden	<input type="checkbox"/>	sehr unzufrieden				
----------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	------------------

**H - 2a Wie zufrieden sind Sie im einzelnen mit nachfolgend genannten Aspekten Ihrer beruflichen Situation?**

	sehr zufrieden			gar nicht zufrieden	
	1	2	3	4	5
Hohes Einkommen und Karrierechancen	<input type="checkbox"/>				
Attraktive Gegend/Stadt zum Leben	<input type="checkbox"/>				
Interessante Tätigkeit	<input type="checkbox"/>				
Arbeits- und Betriebsklima	<input type="checkbox"/>				
Vereinbarkeit von Beruf und Familie	<input type="checkbox"/>				
Weitgehend eigenständige Arbeitsgestaltung	<input type="checkbox"/>				

**H - 3 Geben Sie bitte an, wie wichtig die folgenden Aspekte für Sie beim Berufsstart waren bzw. - falls Sie bisher noch nicht berufstätig waren - heute sind! (Bitte jede Zeile ankreuzen!)**

	sehr wichtig				gar nicht wichtig
	1	2	3	4	5
Hohes Einkommen und Karrierechancen	<input type="checkbox"/>				
Attraktive Gegend/Stadt zum Leben	<input type="checkbox"/>				
Interessante Tätigkeit	<input type="checkbox"/>				
Arbeits- und Betriebsklima	<input type="checkbox"/>				
Vereinbarkeit von Beruf und Familie	<input type="checkbox"/>				
Weitgehend eigenständige Arbeitsgestaltung	<input type="checkbox"/>				

**I - Zukunft, berufliche Pläne und Orientierungen**

**I - 1 Welche beruflichen Veränderungen wollen Sie in den nächsten zwei Jahren realisieren? (Mehrfachnennungen möglich!)**

**Ich beabsichtige ...**

- |   |   |
|---|---|
| <input type="checkbox"/> erstmals eine Tätigkeit aufzunehmen.                     | <input type="checkbox"/> beruflich aufzusteigen.  |
| <input type="checkbox"/> mich auf Kindererziehung/ Familie zu konzentrieren.      | <input type="checkbox"/> eine größere Arbeitsplatzsicherheit zu erreichen.  |
| <input type="checkbox"/> die Beschäftigung/den Tätigkeitsbereich zu wechseln.     | <input type="checkbox"/> eine Beschäftigung, die mehr meinen Fähigkeiten/ meiner Qualifikation entspricht, aufzunehmen. |
| <input type="checkbox"/> ein Vollzeit- /Teilzeit-Studium zu beginnen.             | <input type="checkbox"/> mich beruflich zu qualifizieren.   |
| <input type="checkbox"/> die Anzahl der Arbeitsstunden zu erhöhen.                |   |
| <input type="checkbox"/> die Anzahl der Arbeitsstunden zu vermindern.             |   |
| <input type="checkbox"/> eine selbständige/ freiberufliche Tätigkeit aufzunehmen. |   |
| <input type="checkbox"/> Sonstiges: .....   | (Ggf. bitte eintragen!)   |
- keine Veränderung.

**I - 2 Es folgen nun einige allgemeine Aussagen über den Stellenwert des Berufs. Bitte geben Sie an, inwieweit Sie den jeweiligen Aussagen zustimmen! (Bitte jede Zeile ankreuzen!)**

	stimme völlig zu				stimme gar nicht zu
	1	2	3	4	5
Ich kann mir ein Leben ohne Beruf nur schwer vorstellen. ....	<input type="checkbox"/>				
Ich ziehe eine Tätigkeit, die mich "fesselt", vor auch wenn sie mit beruflicher Unsicherheit verbunden ist. ....	<input type="checkbox"/>				
Wenn sich berufliches Fortkommen und Kinder nicht vereinbaren lassen, ist mir mein Beruf wichtiger. ....	<input type="checkbox"/>				
Ich möchte mir schon einiges leisten können, deshalb ist das Einkommen für mich wichtig. ....	<input type="checkbox"/>				
Ich bevorzuge eine gesicherte Berufsposition, auch wenn die Arbeit weniger meinen sonstigen Erwartungen entspricht. ....	<input type="checkbox"/>				
Ich kann mich auch verwirklichen, ohne berufstätig zu sein. ....	<input type="checkbox"/>				
Ich kann auf viele materielle Dinge verzichten, wenn ich mich nur in meiner Tätigkeit wohlfühle. ....	<input type="checkbox"/>				
Der Beruf ist für mich nur ein Mittel zum Zweck. ....	<input type="checkbox"/>				

**I - 3 Welche berufliche Tätigkeit streben Sie längerfristig an?**

(Bitte eintragen!)

**I - 4 Wie schätzen Sie die Chancen ein, diese beruflichen Pläne zu realisieren?**

- sehr gut       gut       teils/teils       schlecht       sehr schlecht

**J - Angaben zur Person und zu den Eltern**

**J - 1 Ihr Geschlecht?**

- Weiblich       Männlich

**J - 2 In welchem Jahr sind Sie geboren?**

19   (Bitte eintragen!)

**J - 3 Was ist Ihr Familienstand?**

- Ledig ohne Partner/in lebend       Verheiratet       Geschieden  
 Ledig mit Partner/in lebend       Getrennt lebend       Verwitwet

**J - 4 Haben Sie Kinder?**

- Nein       Ja, ein Kind       Ja, zwei Kinder       Ja, mehr als zwei Kinder

**J - 5 Welchen allgemeinbildenden Schulabschluss haben Ihre Eltern?**

(Geben Sie bitte immer den höchsten Abschluss an!)

Vater		Mutter
<input type="checkbox"/>	Schule beendet ohne Abschluss	<input type="checkbox"/>
<input type="checkbox"/>	Volks- und Hauptschulabschluss/ Abschluss der 8. Klasse	<input type="checkbox"/>
<input type="checkbox"/>	Mittlere Reife/ Realschulabschluss/ Abschluss der 10. Klasse	<input type="checkbox"/>
<input type="checkbox"/>	Abitur/ sonstige Hochschulreife	<input type="checkbox"/>
<input type="checkbox"/>	Abitur und abgeschlossenes Hochschulstudium	<input type="checkbox"/>
<input type="checkbox"/>	Anderen Abschluss	<input type="checkbox"/>
<input type="checkbox"/>	Weiß ich nicht	<input type="checkbox"/>

**J - 6 Was ist/ war die überwiegende berufliche Stellung Ihrer Eltern?**

Vater		Mutter
<input type="checkbox"/>	Ungelernte/r / angelernte/r Arbeiter/in	<input type="checkbox"/>
<input type="checkbox"/>	Facharbeiter/in, Vorarbeiter/in, Meister/in	<input type="checkbox"/>
<input type="checkbox"/>	Angestellte/r mit einfacher Tätigkeit (z.B. Verkäufer/in)	<input type="checkbox"/>
<input type="checkbox"/>	Angestellte/r mit qualifizierter Tätigkeit (z.B. Sachbearbeiter/in)	<input type="checkbox"/>
<input type="checkbox"/>	Angestellte/r mit hochqualif. Tätigkeit/Leitungsfunktion (z.B. Prokurist/in)	<input type="checkbox"/>
<input type="checkbox"/>	Angestellte/r mit Führungsaufgaben (z.B. Geschäftsführer/in)	<input type="checkbox"/>
<input type="checkbox"/>	Beamter/ Beamtin	<input type="checkbox"/>
<input type="checkbox"/>	Selbstständige/r, Freiberufler/in	<input type="checkbox"/>
<input type="checkbox"/>	Hausfrau/Hausmann	<input type="checkbox"/>
<input type="checkbox"/>	Anderes: _____ (Bitte eintragen!)	<input type="checkbox"/>
<input type="checkbox"/>	Weiß ich nicht	<input type="checkbox"/>

**Wir danken Ihnen herzlich für Ihre Bereitschaft, unseren Fragebogen auszufüllen und unsere Untersuchung zu unterstützen!**

## Befragung von Lehrenden zur Situation von Lehre und Studium

Wenn Sie in diesem Semester keine Lehrveranstaltungen durchführen: Bitte beziehen Sie die Fragen in diesem Fragenkomplex auf Ihre Erfahrungen in den vorhergehenden Semestern!

### 1. Wie zufrieden sind Sie mit folgenden Rahmenbedingungen der Lehre in Ihren Lehrveranstaltungen? von 1 (sehr zufrieden) bis 5 (sehr unzufrieden)

	1	2	3	4	5
a) Verfügbarkeit von Räumen .....	<input type="checkbox"/>				
b) Raumgröße .....	<input type="checkbox"/>				
c) Akustik der Räume .....	<input type="checkbox"/>				
d) Ausstattung der Räume .....	<input type="checkbox"/>				
e) Verfügbarkeit der für die Lehrveranstaltung notwendigen Literatur in der Bibliothek .....	<input type="checkbox"/>				
f) Verfügbarkeit von Kommunikationsorten für Lehrende und Studierende .....	<input type="checkbox"/>				

### 2. Bitte bewerten Sie nachfolgend die Aspekte für den Studiengang/ den Fachbereich insgesamt, für den Sie in der Lehre überwiegend tätig sind!

	von 1 (sehr gut) bis 5 (sehr schlecht)				
	1	2	3	4	5
a) Aufbau und Struktur des Studiums .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Umfang des Lehrangebotes .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Möglichkeit für die Studierenden, im Studium eigene Interessenschwerpunkte zu setzen .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Forschungsbezug der Lehre .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Berufsvorbereitung/ Praxisbezug .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Klarheit der Prüfungsanforderungen .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Engagement der Lehrenden im Studiengang .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h) Fachliche Kompetenz der Lehrenden .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
i) Beratungs- und Betreuungsqualität .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
j) Gibt es Aspekte, die Ihnen wichtig sind und die hier nicht berücksichtigt wurden? (Bitte nennen!) .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 3. Wie beurteilen Sie aufgrund Ihrer Erfahrung die Situation von Lehre und Studium dieses Semesters im Vergleich zu den vergangenen Semestern?

	von 1 (besseres Semester) über 3 (gleich geblieben) bis 5 (schlechteres Semester)				
	1	2	3	4	5
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Zufriedenheit mit der beruflichen Situation

### 4. Wieviel Prozent Ihrer Arbeitszeit investieren Sie für folgende Bereiche?

Wieviel Prozent möchten Sie gerne den verschiedenen Bereichen widmen?

	Vorlesungszeit		Vorlesungsfreie Zeit	
	Wirklichkeit	Wunsch	Wirklichkeit	Wunsch
<b>Forschung</b>				
<b>Lehre</b>				
<b>Wiss. Dienstleistung</b>				
<b>Verwaltung</b>				
<b>Sonstiges (bitte nennen!)</b> .....				
<b>Insgesamt</b>	<b>100 %</b>	<b>100 %</b>	<b>100 %</b>	<b>100 %</b>

### 5. Wie zufrieden sind Sie – alles in allem betrachtet – mit Ihrer beruflichen Situation?

von 1 (sehr zufrieden) bis 5 (sehr unzufrieden)

1      2      3      4      5  
           

### 6. Wie zufrieden sind Sie im einzelnen mit nachfolgend genannten Aspekten Ihrer beruflichen Situation?

von 1 (sehr zufrieden) bis 5 (sehr unzufrieden)

- |  | 1                        | 2                        | 3                        | 4                        | 5                        |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| a) Lehrveranstaltungen .....   | <input type="checkbox"/> |
| b) Verhältnis zu Kollegen im Fachbereich .....                                       | <input type="checkbox"/> |
| c) Kooperation mit anderen Fachbereichen ..  | <input type="checkbox"/> |
| d) Sicherheit des Arbeitsplatzes .....   | <input type="checkbox"/> |
| e) Aufstiegsmöglichkeiten .....  | <input type="checkbox"/> |
| f) Möglichkeit, eigene Ideen zu verwirklichen.                                       | <input type="checkbox"/> |
| g) Berufliche Anerkennung .....  | <input type="checkbox"/> |
| h) Höhe des Gehalts .....  | <input type="checkbox"/> |
| i) Leistungsgerechtigkeit des Gehalts im Vergleich zu Kollegen im Fachbereich .....  | <input type="checkbox"/> |
| j) Art der Verwaltung in der Hochschule .....  | <input type="checkbox"/> |
| k) Art der Verteilung finanzieller Mittel innerhalb der Hochschule .....             | <input type="checkbox"/> |
| l) Art der Verteilung finanzieller Mittel innerhalb Fakultät/ ggf. Fachbereich ..... | <input type="checkbox"/> |
| m) Mitbestimmungsmöglichkeiten .....   | <input type="checkbox"/> |
| n) Informationen über Entscheidungen der Hochschulleitung .....                      | <input type="checkbox"/> |
| o) Weitere Aspekte (Bitte nennen!) .....   | <input type="checkbox"/> |



## Themen der Hochschulpolitik

**14. In der Hochschulpolitik wird in jüngster Zeit sehr viel über u. g. Themen diskutiert. Wie bewerten Sie persönlich die nachfolgenden Statements?**

von 1 (stimme zu) bis 4 (lehne ich ab)

- |  | 1                        | 2                        | 3                        | 4                        |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| a) Neuorganisation des Studiums nach angelsächsischem Vorbild .....                        | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| b) Profilbildung der Universitäten:  |                          |                          |                          |                          |
| - fachlich-inhaltlich .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| - qualitativ .....   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| c) Das Hochschulsystem braucht mehr marktwirtschaftliche Elemente .....                    | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| d) Veröffentlichung von Leistungskennzahlen  |                          |                          |                          |                          |
| - hochschulintern .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| - sachsenweit .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| e) leistungsorientierte Mittelzuweisung  |                          |                          |                          |                          |
| - durch Studenten .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| - durch Absolventen .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| - durch Experten .....   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| f) Einbeziehung studentischer Bewertungen in leistungsorientierte Mittelvergabe .....      | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| g) Einbeziehung der Bewertung von Absolventen in leistungsorientierte Mittelvergabe .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| h) Einbeziehung der Bewertung von Experten in leistungsorientierte Mittelvergabe .....     | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| i) Bildungsgutscheine zur teilweisen Finanzierung des Hochschuletats durch Studenten ..... | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| j) mehr private Stiftungslehrstühle .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| k) mehr private Hochschulen .....  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

**15. Wenn Sie Ihre Meinung zu einem weiteren Thema sagen wollen: Bitte notieren Sie dies hier kurz oder fügen Sie ein gesondertes Blatt bei!**

.....

.....

.....

.....

### Angaben zur Person

**Fakultät:** .....

**Status:** Professor/in bzw. Hochschullehrer/in

Wissenschaftliches Personal

**Befristet:** ja/ nein



(Wenn die letzte Angabe aus Ihrer Sicht die Anonymität beeinträchtigt, können Sie diese auch freilassen.)

**Vielen Dank für Ihre Unterstützung!**

## Lebenslauf

**Persönliche Daten:** René Krempkow  
Altfränkener Str. 5, 01159 Dresden  
geb. am 18. 02. 1970 in Ludwigslust (Meckl.)

**Schulbildung:**  
1976-1986 Allgemeinbildende Polytechnische Oberschule  
in Ludwigslust, Abschluß mit "Sehr gut"

**Berufsausbildung:**  
09/86-04/88 Ausbildung zum Instandhaltungsmechaniker (Betriebsschlosser)  
im Bahn-Ausbesserungswerk Wittenberge, Abschluß mit "Sehr gut"

**Wehrdienst:**  
05/88-04/90 Wehrdienst als Funktechniker in Schwerin

**Hochschulreife:**  
09/90-07/91 Vorkurs zum Erwerb der Hochschulreife an der Hochschule für  
Verkehrswesen (HfV) Dresden, Abschluß mit "Gut"

**Studium & Praktika:**  
10/91-09/93 Grundstudium Fahrzeugtechnik an HfV (ab 10/92 Auslaufstudiengang)  
07/93-07/96 Mitarbeit im Lehrforschungsprojekt "Studentisches Wohnen in Sachsen"  
am neu gegründeten Institut für Soziologie der TU Dresden (TUD)  
10/93-03/98 Studium der Soziologie/ Kommunikationswissenschaft/ Psychologie an  
der TU Dresden  
09/94 Praktikum im Referat Bürgerinformation der Sächsischen Staatskanzlei  
09/95 Praktikum bei der Schweriner Volkszeitung  
07/96-08/96 Praktikum bei der Sächsischen Zeitung  
03/97-09/97 Diplomarbeit: "Ist gute Lehre meßbar? - Untersuchungen zur Validität,  
Zuverlässigkeit und Vergleichbarkeit studentischer Lehrbewertungen";  
10/97-03/98 Konzeption des Pilotprojektes für Studentenbefragungen zur Lehrqualität  
an der TUD (als Studentische Hilfskraft zeitgleich zu Diplomprüfungen)  
Studienabschluß mit der Note "Gut" (1,7), Diplomarbeit „Sehr gut“ (1,3)

**Studienbegleitend:**  
10/91-09/93 Mitglied der Fachschaftsvertretung Verkehrswissenschaften,  
Sozialberatung im Studentenrat der TU Dresden und  
Organisation des hochschulübergreifenden Projektes "Semesterticket in  
Dresden" (insbes. Akzeptanzuntersuchung und Öffentlichkeitsarbeit)  
10/92-03/98 Stipendiat des Begabtenförderungswerkes "Hans-Böckler-Stiftung"  
12/92-09/93 Geschäftsführer des Studentenrates der TU Dresden  
seit 12/92 Freie Mitarbeit bei Hochschulzeitung "ad rem" und "Sächsische Zeitung"  
seit 10/96 Freie Mitarbeit bei Universitätsjournal/ Absolventenmagazin der TUD

**Berufstätigkeit:**  
05/90-08/90 Instandhaltungsmechaniker im Bahnbetriebswerk Schwerin  
06/95-08/95 Projektarbeit/ befristete Tätigkeit im Referat Öffentlichkeitsarbeit des  
Sächsischen Staatsministeriums für Wissenschaft und Kunst  
03/98-12/00 Projektleiter Evaluation der Lehre an der TU Dresden (1/2 Stelle) und  
04/99-09/00 Mitarbeiter im Studentischen Evaluationsbüro Sachsen (1/2 Stelle)  
04/99-09/05 Koordinator Projektgruppe Absolventenbefragungen TUD  
10/00-09/04 Promotionsstipendiat der Hans-Böckler-Stiftung  
seit 11/04 Leitung des Projektes „Hochschulbericht Sachsen“