

Qualität in der Wissenschaft

Zeitschrift für Qualitätsentwicklung in
Forschung, Studium und Administration

■ Fakultätsübergreifendes Monitoring
der Veranstaltungsqualität am
Karlsruher Institut für Technologie (KIT)

■ Das Bachelor-Studium aus systemischer Sicht

■ Internationale Vergleichbarkeit von Noten im
Hochschulbereich? Problematik der Notenvergabe,
Referenzgrößen und der Verwendung
der Gaußschen Normalverteilung

■ ECTS-Notenzuweisung:
Pragmatische Kohortenbildung

1
—
2010

Herausgeberkreis

Doris Carstensen, Dr. Mag., Vizerektorin für Qualitätsmanagement, Gender Mainstreaming und Personalentwicklung an der KunstUniversität Graz

Hans-Dieter Daniel, Prof. Dr., Professur für Sozialpsychologie und Hochschulforschung, ETH Zürich (CH), Leiter der Evaluationsstelle der Universität Zürich

Michael Heger, Dr., Evaluationsbeauftragter der Fachhochschule Aachen, Leiter des Bereichs Hochschuldidaktik und Evaluation in der zentralen Qualitätsentwicklung ZQE

Achim Hopbach, Dr., Geschäftsführer der Stiftung zur Akkreditierung von Studiengängen in Deutschland, Bonn

Stefan Hornbostel, Prof. Dr., Leiter des Instituts für Forschungsinformation und Qualitätssicherung (IFQ), Bonn, Professur für Soziologie (Wissenschaftsforschung), Institut für Sozialwissenschaften der Humboldt-Universität zu Berlin

René Krempkow, Dr., Institut für Forschungsinformation und Qualitätssicherung (IFQ), Bonn

Sandra Mittag, Dr., Referatsleiterin Qualitätsmanagement, TU Darmstadt

Philipp Pohlenz, Dr., Leiter der Servicestelle für Lehrevaluation, Universität Potsdam

Martina Röbbecke, Dr., acatech (Deutsche Akademie der Technikwissenschaften), München

Uwe Schmidt, Dr., Leiter des Zentrums für Qualitätssicherung und -entwicklung der Universität Mainz, Geschäftsführer des Hochschulevaluationsverbundes Südwest

Wolff-Dietrich Webler, Prof. Dr., Leiter des Instituts für Wissenschafts- und Bildungsforschung Bielefeld (IWBB), Professor of Higher Education, University of Bergen (Norway); Staatliche Pädagogische Universität Jaroslavl/Wolga

Don Westerheijden, Dr., Center for Higher Education Policy Studies (CHEPS), University of Twente, Netherlands

Lothar Zechlin, Prof. Dr., ehem. Rektor der Universität Duisburg-Essen, Vertreter der Agenturen im Akkreditierungsrat

Hinweise für die Autoren

In dieser Zeitschrift werden i.d.R. nur Originalbeiträge publiziert. Sie werden doppelt begutachtet. Die Autor/innen versichern, den Beitrag nicht zu gleicher Zeit an anderer Stelle zur Publikation angeboten zu haben. Beiträge werden nur dann angenommen, wenn die Autor/innen den Gegenstand nicht in vergleichbarer Weise in einem anderen Medium behandeln. Senden Sie bitte zwei Exemplare des Manuskripts in Papierform sowie einmal in Dateiform (kann als Daten-CD der Papierform beigelegt oder per

E-Mail zugeschickt werden) an die Redaktion (Adresse siehe Impressum).

Wichtige Vorgaben zu Textformatierungen und beigefügten Fotos, Zeichnungen sowie Abbildungen erhalten Sie in den „Autorenhinweisen“ auf unserer Verlags-Homepage: „www.universitaetsverlagwebler.de“.

Ausführliche Informationen zu den in diesem Heft aufgeführten Verlagsprodukten erhalten Sie ebenfalls auf der zuvor genannten Verlags-Homepage.

Impressum

Anschrift Verlag, Redaktion, Abonnentenverwaltung

UVW UniversitätsVerlagWebler
Der Fachverlag für Hochschulthemen
Bünder Straße 1-3 (Hofgebäude)
33613 Bielefeld
Tel.: 0521 - 92 36 10-12, Fax: 0521 - 92 36 10-22

Satz: K. Gerber, info@universitaetsverlagwebler.de

Anzeigen:

Die Zeitschrift „Qualität in der Wissenschaft“ veröffentlicht Verlagsanzeigen, Ausschreibungen und Stellenanzeigen. Aufträge sind an den Verlag zu richten. Die jeweils gültigen Anzeigenpreise sind der Homepage erhalten Sie auf Anfrage beim Verlag.

Erscheinungsweise: 4mal jährlich

Redaktionsschluss dieser Ausgabe: 20.04.2010

Umschlagsgestaltung:

Wolff-Dietrich Webler, Bielefeld
Gesetzt in der Linotype Syntax Regular

Druck:

Sievert Druck & Service GmbH,
Potsdamer Str. 190, 33719 Bielefeld

Abonnement/Bezugspreis:

Jahresabonnement: 59 Euro zzgl. Versandkosten
Einzelpreis dieser Doppelausgabe: 29 Euro zzgl. Versandkosten

Abbestellungen und die Bestellungen von Einzelheften sind unterschrieben per Post, E-Mail oder Fax an den Verlag zu richten.

Das Jahresabonnement verlängert sich automatisch um ein Jahr, wenn es nicht 6 Wochen vor Jahresende gekündigt wird.

Copyright: Uvw UniversitätsVerlagWebler

Die mit Verfassernamen gekennzeichneten Beiträge geben nicht in jedem Falle die Auffassung der Herausgeber bzw. Redaktion wieder. Für unverlangt eingesandte Manuskripte/Rezenzionsexemplare wird keine Verpflichtung zur Veröffentlichung/Besprechung übernommen. Sie können nur zurückgegeben werden, wenn dies ausdrücklich gewünscht wird und ausreichendes Rückporto beigelegt ist. Die Urheberrechte der hier veröffentlichten Artikel, Fotos und Anzeigen bleiben bei der Redaktion. Der Nachdruck ist nur mit schriftlicher Genehmigung des Verlages gestattet.

Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Qualität in der Wissenschaft

Zeitschrift für Qualitätsentwicklung in
Forschung, Studium und Administration

Einführung des geschäftsführenden Herausgebers

1

Qualitätspolitik/ Qualitätsentwicklung

Michael Craanen

Fakultätsübergreifendes Monitoring der Veranstaltungs-
qualität am Karlsruher Institut für Technologie (KIT) **2**

E. W. Udo Küppers

Das Bachelor-Studium aus systemischer Sicht **12**

Diskussionsforum

Wolff-Dietrich Webler

Internationale Vergleichbarkeit von Noten im
Hochschulbereich? Problematik der Notenvergabe,
Referenzgrößen und der Verwendung der
Gaußschen Normalverteilung **20**

Jonas Kunze & Andreas Geyer-Schulz

ECTS-Notenzuweisung:
Pragmatische Kohortenbildung **24**

Tagungsankündigung

Wie kann Qualitätsmanagement zur Entwicklung
der Hochschullehre beitragen?

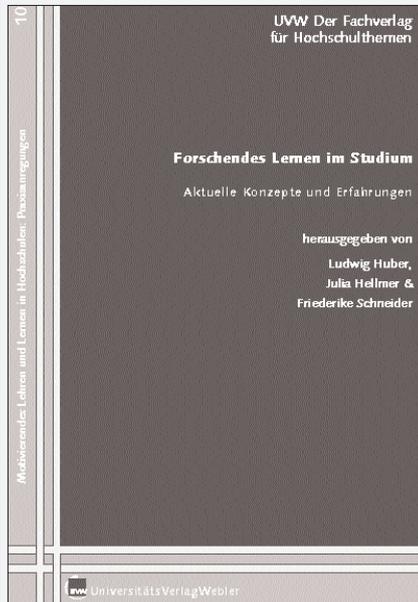
Tagung der Internationalen Bodensee-Hochschule **III**

Seitenblick auf die Schwesterzeitschriften

Hauptbeiträge der aktuellen Hefte
Fo, HSW, HM, P-OE und ZBS **IV**

im UniversitätsVerlagWebler erhältlich:

**Ludwig Huber, Julia Hellmer & Friederike Schneider (Hg.):
Forschendes Lernen im Studium. Aktuelle Konzepte und Erfahrungen**



ISBN 3-937026-66-5, Bielefeld 2009,
227 Seiten, 29.60 Euro

Das Konzept des Forschenden Lernens, das vor 40 Jahren von der Bundesassistentenkonferenz ausgearbeitet wurde und weithin großes Echo fand, gewinnt gegenwärtig erneut an Aktualität. Im Zusammenhang mit dem „Bologna-Prozess“ werden Anforderungen an die Entwicklung allgemeiner Kompetenzen der Studierenden gestellt, zu deren Erfüllung viel größeres Gewicht auf aktives, problemorientiertes, selbstständiges und kooperatives Arbeiten gelegt werden muss; Forschendes Lernen bietet dafür die einem wissenschaftlichen Studium gemäße Form.

Lehrenden und Studierenden aller Fächer und Hochschularten, die Forschendes Lernen in ihren Veranstaltungen oder Modulen verwirklichen wollen, soll dieser Band dienen. Er bietet im ersten Teil Antworten auf grundsätzliche Fragen nach der hochschuldidaktischen Berechtigung und den lerntheoretischen Gründen für Forschendes Lernen auch schon im Bachelor-Studium. Im zweiten Teil wird über praktische Versuche und Erfahrungen aus Projekten Forschenden Lernens großenteils aus Hamburger Hochschulen berichtet. In ihnen sind die wichtigsten Typen und alle großen Fächerbereiche der Hochschulen durch Beispiele repräsentiert. Die Projekte lassen in ihrer Verschiedenartigkeit die unterschiedlichen Formen und Ausprägungsgrade erkennen, die Forschendes Lernen je nach Fach annehmen kann (und auch muss); zugleich zeigen sie die reizvolle Vielfalt möglicher Themen und Formen. Im dritten Teil werden in einer übergreifenden Betrachtung von Projekten zum Forschenden Lernen Prozesse, Gelingensbedingungen, Schwierigkeiten und Chancen systematisch zusammengeführt.

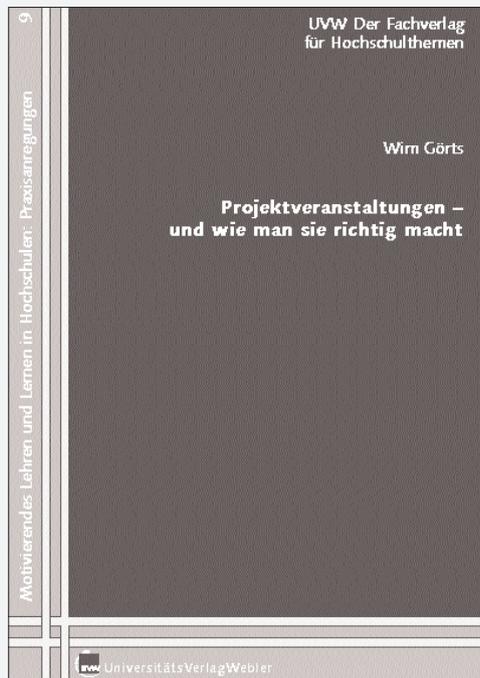
Insgesamt soll und kann dieses Buch zu immer weiteren und immer vielfältigeren Versuchen mit Forschendem Lernen anregen, ermutigen und helfen.

Reihe Motivierendes Lehren und Lernen
in Hochschulen: Praxisanregungen

Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

**Wim Görts
Projektveranstaltungen – und wie man sie richtig macht**

Wim Görts hat hier seinen bisherigen beiden Bänden zu Studienprojekten in diesem Verlag eine weitere Anleitung von Projekten hinzugefügt. Ein variationsreiches Spektrum von Beispielen ermutigt zu deren Durchführung. Das Buch bietet Lehrenden und Studierenden zahlreiche Anregungen in einem höchst befriedigenden Bereich ihrer Tätigkeit. Die Verstärkung des Praxisbezuges der Lehre bzw. der Handlungskompetenz bei Studierenden ist eine häufig erhobene Forderung. Projekte gehören - wenn sie gut gewählt sind - zu den praxsnächsten Studienformen. Mit ihrer ganzheitlichen Anlage kommen sie der großen Mehrheit der Studierenden, den holistischen Lernern, sehr entgegen. Die Realisierung von Projekten fördert Motivation, Lernen und Handlungsfähigkeit der Studierenden erheblich und vermittelt dadurch auch besondere Erfolgserlebnisse für die Lehrenden bei der Realisierung der einer Hochschule angemessenen, anspruchsvollen Lehrziele. Die Frage zum Studienabschluss, in welcher Veranstaltung Studierende am meisten über ihr Fach gelernt haben, wurde in der Vergangenheit häufig mit einem Projekt (z.B. einer Lehrforschung) beantwortet, viel seltener mit einer konventionellen Fachveranstaltung. Insofern sollten Studienprojekte gefördert werden, wo immer es geht. Die Didaktik der Anleitung von Projekten stellt eine „Königsdisziplin“ der Hochschuldidaktik dar. Projekte gehören zum anspruchsvollsten Bereich von Lehre und Studium. Nur eine begrenzte Zeit steht für einen offenen Erkenntnis- und Entwicklungsprozess zur Verfügung. Insofern ist auf die Wahl sowie den Zuschnitt des Themas und die Projektplanung besondere Sorgfalt zu verwenden. Auch soll es der Grundidee nach ein Projekt der Studierenden sein, bei dem die Lehrperson den Studierenden über die Schulter schaut. Die Organisationsfähigkeit und Selbstdisziplin der Studierenden sollen gerade im Projekt weiter entwickelt werden. Der vorliegende Band bietet auch hierzu zahlreiche Anregungen.



ISBN 3-937026-60-6, Bielefeld 2009,
138 Seiten, 19.80 Euro

Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Reihe Motivierendes Lehren und Lernen
in Hochschulen: Praxisanregungen

Um ihre Themenfelder sowohl in der Forschung als auch in der Praxis noch besser zu repräsentieren, an speziellen Kompetenzen noch hinzuzugewinnen und weitere anregende Aufsätze einzuwerben, hat die Zeitschrift QiW den Kreis ihrer Herausgeber weiter gezogen. Der Verlag und die bisherigen Herausgeber begrüßen als neue Kolleginnen und Kollegen in unseren Reihen:

Achim Hopbach, Dr., Geschäftsführer der Stiftung zur Akkreditierung von Studiengängen in Deutschland, Bonn;

René Kremppkow, Dr., Institut für Forschungsinformation und Qualitätssicherung (IFQ), Bonn;

Sandra Mittag, Dr., Referatsleiterin Qualitätsmanagement, TU Darmstadt;

Philipp Pohlentz, Dr., Leiter der Servicestelle für Lehrevaluation, Universität Potsdam;

Lothar Zechlin, Prof. Dr., ehem. Rektor der Universität Duisburg-Essen, Vertreter der Agenturen im Akkreditierungsrat.

Die neuen Mitglieder im Herausgeberkreis werden in der Ausgabe 2-2010 vorgestellt.

Zwar sind studentische Veranstaltungsbewertungen inzwischen an den meisten Hochschulen eingeführt. Aber - von großen Unterschieden in der methodischen Qualität der verwendeten Fragebögen abgesehen - werden oft zahlreiche Daten erhoben, ohne sie systematisch auszuwerten („erklären sich von selbst“) oder regelmäßig Konsequenzen daraus zu ziehen. So werden umfangreiche Datenfriedhöfe erzeugt. Die Unlust, Fragebögen zu beantworten, wächst auf allen Seiten, weil diese Prozesse folgenlos scheinen. Das Wort von der „Evaluitis“ geht um. In dieser Situation hat *Michael Craanen* ein folgenreiches Modell entwickelt. Er stellt es in seinem Aufsatz **Fakultätsübergreifendes Monitoring der Veranstaltungsqualität am Karlsruher Institut für Technologie (KIT)** vor. Das Monitoring zieht z.Z. viel Aufmerksamkeit auf sich und wird an anderen Hochschulen sehr interessiert, aber auch sehr kontrovers diskutiert. Die Firma ElectricPaper arbeitet an einer Software als Ergänzung zu EvaSys, die das Monitoring und die Follow-up-Gruppeneinteilung für das Berichtswesen automatisiert. Die Software soll auf der nächsten EvaSys-Anwendertagung in Karlsruhe am KIT im September der Fachöffentlichkeit vorgestellt und dann international beworben werden.

Seite 2

E.W. Udo Küppers führt in seinem Aufsatz **Das Bachelor-Studium aus systemischer Sicht** vor, wie systemisches Denken dazu beitragen kann, Zusammenhänge in einer vernetzten dynamischen Umwelt zu analysieren und zu bewerten. Der Autor legt ein Modell vor, wie komplexe Systeme - hier beispielhaft das Bachelor-Studium - mit Hilfe von Wirkungsnetzen als Werkzeuge systemischen Bildungsdenkens analysiert und einer Lösung zugeführt werden können. Damit behandelt er ein in der Debatte um die Reform der Bachelor-Studiengänge wichtiges Thema, das in dieser Zeitschrift publiziert werden sollte. Die hier vorgelegten Gedanken sind bisher in dieser Weise noch nicht

öffentlich diskutiert worden. Unter den Möglichkeiten, neue Erkenntnisse in diesem Rahmen zu publizieren, wahlweise oder kombiniert durch 1. theoretische Analyse der Realität; 2. empirische Analyse; 3. Visionen/Entwürfe neuer Realität; 4. Präsentation neuer oder Weiterentwicklung bekannter Methoden bzw. Verfahren, die die vorstehenden drei Varianten befördern könnten, ist der vorliegende Aufsatz zu Punkt 4 zu zählen.

Seite 12

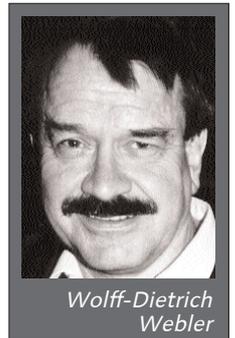
Ein gemeinsamer europäischer Hochschulraum und europäischer Arbeitsmarkt erfordern vergleichbare Bewerbungsunterlagen. Damit kann Chancengleichheit gefördert, die Gefahr von Diskriminierungen durch Klischees, Vorurteile usw. gemindert werden. Zu vergleichbaren Bewerbungen gehören vergleichbare Noten. Als eine Voraussetzung hat die EU die ECTS-Noten eingeführt. Ist damit Vergleichbarkeit hergestellt? *Wolff-Dietrich Webler* bestreitet das. In seinem Artikel: **Internationale Vergleichbarkeit von Noten im Hochschulbereich? Problematik der Notenvergabe, Referenzgrößen und die Verwendung der Gaußschen Normalverteilung** durchleuchtet er die Gründe, warum selbst gleiche Noten noch keine Vergleichbarkeit der zugrunde liegenden Leistungsbilder bedeuten. Dazu fehlen viele Voraussetzungen, z.B. Einigkeit über Studienziele, Leistungsniveaus, die den Noten zugrunde liegen, u.v.m. Noten, die nach vergleichbaren Kriterien vergeben werden, so weist der Autor nach, sind wohl noch lange nur Wunschvorstellung.

Seite 20

Jonas Kunze und Andreas Geyer-Schulz gehen in ihrem Beitrag **Die grundlegende Problematik der ECTS-Notenzuweisung: Ein Musterverfahren** das Thema der ECTS-Noten von anderer Seite an. Die EU-Kommission fördert die ECTS-Notenzuweisung, auf nationaler Ebene gibt es Empfehlungen und Vorgaben (z.B. HRK) - also sind dies zunächst einmal Fakten, auch wenn die EU-Kommission feststellt, von diesem Verfahren werde noch ziemlich wenig Gebrauch gemacht. Die Autoren lassen sich also realistischerweise auf das Verfahren ein. Die konkrete Berechnung von ECTS-Noten stellt nach wie vor eine Herausforderung an die Hochschulen dar, bei der viele Fragen offen sind. Seine Analyse fördert ein immanentes, gravierendes methodisches Problem zu Tage. In ihrem Beitrag stellen Jonas Kunze und Andreas Geyer-Schulz die Verwendung einer absoluten Kohorte/Referenzgruppe in Frage, indem anhand von Rechenbeispielen unerwünschte Effekte nachgewiesen werden. Sie entwickeln dann eine Lösung, die die bisherigen immanenten methodischen Probleme vermeidet. Die Autoren empfehlen ein kontinuierliches Verfahren, welches (auf operativer und rechtlicher Ebene) entscheidende Vorteile bzgl. Robustheit, Nachvollziehbarkeit und Einfachheit bietet.

Seite 24

W.W.



Wolff-Dietrich Webler

Michael Craanen

Fakultätsübergreifendes Monitoring der Veranstaltungsqualität am Karlsruher Institut für Technologie (KIT)



Michael Craanen

Veranstaltungsevaluationen finden heute nahezu an jeder deutschen Hochschule statt, doch trotz ihrer hervorgehobenen Rolle als eines der wichtigsten Instrumente der Sicherung und Entwicklung der Prozessqualität von Studium und Lehre sind ihre Veränderungspotentiale in Hinblick auf mehr „Qualität“ auf der Veranstaltungsebene bisher begrenzt. Der vorliegende Aufsatz stellt ein neues empirisches Verfahren vor, das die klassisch durchgeführten Veranstaltungsevaluationen als Paper & Pencil bzw. Online-Befragungen mit Rückmeldung der empirischen Ergebnisse an die Lehrenden ergänzt mit dem Ziel, empirisch messbares *Follow-Up*¹ zu erzeugen.

Das hier vorgestellte Monitoring der Veranstaltungsqualität mit verknüpftem *Follow-Up* auf unterschiedlichen Ebenen ist im Rahmen des Aufbaus der Qualitätssicherung für Studium und Lehre an der Universität Karlsruhe (TH) seit Anfang 2007 entwickelt worden. Mit der Fusion der Universität Karlsruhe (TH) mit dem Forschungszentrum Karlsruhe zum Karlsruher Institut für Technologie (KIT) im Oktober 2009 wird das Konzept nun im KIT fortgeführt.

Seit Sommersemester 2008 liefert das Monitoring vergleichbare Daten für alle 11 Fakultäten und zentralen Einrichtungen mit Lehraufgaben, die sowohl die Wirkungen der vielfältigen Maßnahmen der Qualitätssicherung dokumentieren² als auch konkret aufzeigen, in welchen Veranstaltungen jedes Semester gezielt *Follow-Up* erzeugt werden soll.

Seit der Einführung des Monitoring hat sich allmählich eine Qualitätskultur im Bereich der Lehre bzw. des Lernens auf allen Ebenen entwickelt. Sowohl Studierende und deren Fachschaften, alle Fakultätsleitungen und der Bereichsvorstand Lehre des KIT³ tragen das Modell aktiv mit. Lediglich auf Seiten der Lehrenden gibt es gelegentlich Vorbehalte Einzelner, die sich vor allem dadurch erklären, dass zunehmend von den Studiendekanen und den Studienkommissionen und ggf. auch vom Bereichsvorstand Lehre eingegriffen wird, wenn eine Veranstaltung die Mindestqualitätsanforderungen des KIT nicht erfüllt.

1. Zuständigkeiten, Durchführung und Rechtsgrundlage am KIT

Mit der Durchführung der fakultätsübergreifenden Lehrveranstaltungsevaluationen ist die Evaluationsstelle beauftragt, die organisatorisch in Planung und Controlling des KIT angesiedelt ist. Getragen von einem Team aus 6-7 stu-

dentischen und wissenschaftlichen Hilfskräften sowie einer wissenschaftlichen Mitarbeiterin und der Leitung der Evaluationsstelle werden jedes Semester etwa 1.500 Veranstaltungen an allen Fakultäten und zentralen Einrichtungen evaluiert. Der Rücklauf ausgefüllter Fragebögen beträgt zwischen 45.000 und 55.000 Fragebögen pro Semester. Dies entspricht einer Rücklaufquote von ca. 95%. Zur Vermeidung des bekannten Stichprobenbias von Online-Evaluationen werden am KIT ausschließlich „Paper & Pencil-Befragungen“ durchgeführt. Die Auswertung und der Ergebnisversand an die Lehrenden geschieht mit dem automatisierten Lehrevaluationssystem „EvaSys“ der ElectricPaper GmbH. Aus den Einzelergebnissen werden für alle Lehrereinheiten Berichte zur Veranstaltungsqualität erstellt, die den Fakultätsleitungen und dem Bereichsvorstand Lehre des KIT als Grundlage zur Verbesserung der Studienbedingungen dienen.

Die Verpflichtung zur Durchführung der Evaluation der Lehre und als Teil davon der von Lehrveranstaltungen ergibt sich aus dem Landeshochschulgesetz Baden-Württemberg (LHG). Dieses enthält in § 5 Abs. 2 eine Satzungsermächtigung, nach der die Hochschulen weitere Regelungen – insbesondere in datenschutzrechtlicher Hinsicht – zu treffen haben. Die Teilnahme von Lehrenden und Studierenden ist verpflichtend. Sie wird durch § 5 Abs. 2 Satz 2 LHG bestimmt: „Die betroffenen Mitglieder und Angehörigen der Hochschule sind gegenüber ihrer Hochschule zur Mitwirkung und zur Angabe auch personenbezogener Daten verpflichtet.“

Für das KIT ist die wichtigste Regelungsgrundlage die Evaluationsordnung für Lehre, Studium und Weiterbildung, die die Lehrveranstaltungsevaluation und ihren organisatorischen Ablauf verbindlich regelt.

¹ Follow-Up wird hier definiert als alle Folgemaßnahmen der Veranstaltungsevaluationen, die zu einer Verbesserung der Einzelqualität einer Veranstaltung oder der Veranstaltungen einer Lehrereinheit insgesamt führen.

² Die „Dokumentation“ der Wirkungen der Maßnahmen der Qualitätssicherung ist ein wichtiger Bestandteil der Systemakkreditierung, gefordert im vierten Kriterium „Berichtssystem und Datenerhebung“: „Die Hochschule verfügt über ein internes Berichtssystem, das die Strukturen und Prozesse in der Entwicklung und Durchführung von Studiengängen sowie die Strukturen, Prozesse und Maßnahmen der Qualitätssicherung, ihre Ergebnisse und Wirkungen dokumentiert“ (Drs. AR 93/2009: 25).

³ Der „Bereichsvorstand Lehre des KIT“ entspricht der Funktion des „Prorektor bzw. Vizepräsident für Studium und Lehre“.

2. „Veranstaltungsqualität“ am KIT

„Veranstaltungsqualität“ wird in der allgemeinen Diskussion oft mit Lehrqualität und der in der Hochschuldidaktik seit vielen Jahrzehnten andauernden Diskussion über die Frage „Was ist gute Lehre?“ verbunden. Beim Ausfüllen eines Veranstaltungsfragebogens beantworten die Studierenden die Frage nach der Lehrqualität einer Veranstaltung jedoch mitnichten, obgleich sich natürlich hinter der Auswahl der Fragen in einem typischen Fragebogen zur Veranstaltungsbewertung immer auch ein Modell guter Lehre aus der Sicht der Fragebogenkonstrukteure verbirgt. Studierende bewerten die Fragen in Hinblick auf ihre eigene persönliche Studiensituation (Anzahl gleichzeitig besuchter Veranstaltungen, Anzahl Prüfungen, zeitliche Passung in den Studienplan etc.), in Bezug auf ihre persönliche Situation (Jobben vs. Arbeitsaufwand fürs Studium etc.) und ihres persönlichen Lerntyps.

Diese Unterschiedlichkeit der persönlichen Lernbedürfnisse und -bedingungen führt auch immer zu einer gewissen Varianz bei den Antworten, die sich jedoch bei zunehmender Teilnehmerzahl einer Veranstaltungsevaluation wieder relativiert. Veranstaltungen erfüllen nur in seltenen Fällen die Lernbedürfnisse aller teilnehmenden Studierenden gleichermaßen. Auf der Grundlage dieser empirisch nachprüfbarer Erfahrung wurde am KIT ein eigener studierendenzentrierter Qualitätsbegriff für Veranstaltungen entwickelt, der auch als theoretische Grundlage bei der Operationalisierung des fakultätsübergreifenden Monitoring dient:

Veranstaltungsqualität ist die Erfüllung der Lernbedürfnisse der Studierenden in einer Veranstaltung unter inhaltlicher Berücksichtigung der gesellschaftlichen Anforderungen an den Kompetenzerwerb für Forschung, Lehre und Wirtschaft.

Je mehr es in einer Veranstaltung gelingt, die naturgemäß unterschiedlichen Lernbedürfnisse der Studierenden zu erfüllen, desto höher ist auch die allgemeine Qualität dieser Veranstaltung aus der Sicht aller teilnehmenden Studierenden. Beim Monitoring werden die Lernbedürfnisse der Studierenden in den Kern der Qualitätsbetrachtung von Studium und Lehre gerückt. Das Verfahren trägt damit dem im Bologna-Prozess geforderten Paradigmenwechsel Rechnung, dem zufolge die Hochschulen in erster Linie als „lernorientierte“ und weniger als „lehrorientierte“ Bildungseinrichtungen auszugestalten sind.

Die Frage nach der allgemeinen Qualität der Lernbedingungen inklusive der Lehrqualität, die ursächlich für den Veranstaltungserfolg sind, spielt beim Monitoring selbst noch keine Rolle, sondern erst beim zeitlich dem Monitoring folgenden Follow-Up.

3. Das KIT-Modell der fakultätsübergreifenden Veranstaltungsevaluation

Veranstaltungsevaluationen bedienen Ziele, die je nach Akteur sehr unterschiedlich sind. Auf einer ganz allgemeinen Ebene ist ihre Funktion die Transparenz und Rechenschaftslegung der Qualität von Studium und Lehre gegenüber Staat und Öffentlichkeit. Auf der Ebene der Fakultätsleitun-

gen und der für die Lehre verantwortlichen Gremien und Personen dienen die Ergebnisse als empirische Grundlage für die Sicherung der Studierbarkeit der Studiengänge. Auf der Ebene der Lehrenden stellen sie den Lehrenden empirische Daten zur Verfügung, die sie zur Selbstvergewisserung über die Qualität des eigenen Hochschulunterrichts nutzen können. Für Studierende sind sie das wichtigste Feedback an die Lehrenden über den eigenen Lernerfolg.

Auf der Grundlage dieser komplexen unterschiedlichen Ansprüche an die Ergebnisse der Veranstaltungsevaluationen sowie der Notwendigkeit, die datenschutzrechtlichen Vorgaben in das Gesamtkonzept der fakultätsübergreifenden Lehrveranstaltungsevaluation zu implementieren und zusätzlich unter strenger Berücksichtigung des zentralen Ziels jeder Evaluation, zu einer Qualitätsentwicklung des Untersuchungsgegenstandes beizutragen, wurde ein anspruchsvolles empirisches und dazu sehr praxisnahes *KIT-Modell der fakultätsübergreifenden Veranstaltungsevaluation* entwickelt.

Kern des Modells sind ein schlankes Berichtswesen, das bei überschaubarem Personaleinsatz semesterweise ein Maximum an empirischen Informationen zur Verfügung stellt sowie eine konsequente *Follow-Up*-Orientierung. Ergänzt wird das Modell durch einen sehr engen informellen Austausch zwischen Fakultätsleitungen, für die Lehre verantwortlichen Gremien und Personen der Fakultäten, dem Bereichsvorstand Lehre des KIT und der Evaluationsstelle sowie durch die Integration der fachspezifischen Evaluationenkulturen in die zentrale Evaluation.

Die wichtigsten Funktionen des Monitoring für die Qualitätssicherung des KIT als zentraler Teil der fakultätsübergreifenden Veranstaltungsevaluation sind dabei je nach Zielgruppe unterschiedlich:

Studierende:

Die Studierenden am KIT nehmen schon seit einiger Zeit die positiven Veränderungen wahr, die durch das Monitoring initiiert wurden. Über die Veranstaltungsevaluationen teilen sie der Fakultätsleitung systematisch jedes Semester mit, in welchen Veranstaltungen aus ihrer Perspektive der dringendste Handlungsbedarf besteht. Aus den Erfahrungen, dass die Evaluationen tatsächlich auch zu Veränderungen führen, hat sich eine studentische Evaluationskultur herausgebildet, die sich u.a. auch in einer starken Zustimmung aller Fachschaften am KIT zugunsten der zentralen Evaluation äußert.

Lehrende:

Anhand weniger „visueller“ Signale wird vermittelt, ob die eigene Veranstaltung die Mindestqualitätsstandards des KIT erfüllt.

Fakultätsvorstände:

Filterung der Einzelergebnisse der Massenevaluationen zur Erzeugung von gezieltem *Follow-Up* in den Lehreinheiten und Überprüfung der Wirkungen der Maßnahmen der Qualitätssicherung im zeitlichen Verlauf.

Bereichsvorstand Lehre und Senatskommission Studium und Lehre:

Darstellung der Veranstaltungsqualität aller Studiengänge anhand weniger Kennzahlen und Überprüfung der Wirkun-

gen der Maßnahmen der Qualitätssicherung im zeitlichen Verlauf.

Staat und Gesellschaft:

Teil der öffentlichen Rechenschaftslegung über die Qualität der wissenschaftlichen Ausbildung am KIT.

4. Der Lehrqualitätsindex (LQI)

Wie eingangs schon erwähnt, ist das wichtigste Merkmal der zentralen Veranstaltungsevaluation die konsequente Follow-Up-Fokussierung auf der Ebene der Einzelveranstaltung und der Lehreinheit.

Die traditionell aggregierte Darstellung der Ergebnisse von Veranstaltungsevaluationen in Häufigkeitsdiagrammen, Mittelwerten sowie Standardabweichungen in den Endberichten zur Situation der Lehre erfüllt diesen Anspruch naturgemäß nicht. Die negativen Ergebnisse einer Minderheit von Veranstaltungen, deren Qualität dringend verbesserungswürdig ist, sind bei der Anwendung aggregierender und deskriptivstatistischer Verfahren aufgrund der großen Mehrheit von positiv bewerteten Veranstaltungen nicht darstellbar. Diese Berichte haben folglich keine Aussagekraft für gezielte Verbesserungen in Studium und Lehre. Aggregierte Berichte können lediglich als sinnvolle Ergänzung zu den Einzelergebnissen für die Lehrenden herangezogen werden, um die eigene Bewertung im Vergleich mit den Gesamtergebnissen besser einordnen und interpretieren zu können.

Die Einführung des Monitoring in das Berichtswesen der Veranstaltungsevaluation löst das Problem der fehlenden Aussagekraft aggregierter Berichte für die Qualitätssicherung und -entwicklung der Lehre. Das Monitoring setzt nicht auf die aggregierte Ergebnisdarstellung der Einzelfragen aus den Fragebogen, sondern es stellt das Erreichen eines Mindestqualitätsstandards für jede Einzelveranstaltung dar und ermöglicht so gezieltes *Follow-Up* durch die Fakultätsleitungen und den Bereichsvorstand Lehre des KIT. Die wichtigste Voraussetzung für das Monitoring ist, dass alle Fach- und Veranstaltungsspezifischen Fragebögen einen einheitlichen Frageblock zur Berechnung von vergleichbaren Kennwerten enthalten.

Am KIT hat sich als einheitlicher Kennwert für das Monitoring der Begriff *Lehrqualitätsindex (LQI)* ausgeprägt, dessen Berechnung im folgenden dargestellt werden soll. Der Begriff *Lehrqualitätsindex* ist an sich etwas unglücklich gewählt. Er hat sich schon in der ersten Erprobungsphase des Monitoring umgangssprachlich an allen Fakultäten am KIT eingebürgert, so dass wir es dabei belassen haben. Auch eine einfache Umbenennung in „Lern“qualitätsindex würde es nicht genau treffen, denn er bildet anhand weniger Variablen die Erfüllung der Lernbedürfnisse der Studierenden durch die Veranstaltung ab. Im übertragenen Sinne ist er auch ein Maß der Zufriedenheit der Studierenden mit einer Veranstaltung.

Der LQI als zentraler Kennwert für das Monitoring beruht auf die einfache in der Praxis der fakultätsübergreifenden Veranstaltungsevaluation gemachte Beobachtung, dass die Freitextkommentare der Studierenden in den Veranstaltungsbögen immer kritischer werden, je niedriger der berechnete LQI für diese Veranstaltung ist. Ein testtheoretisches Verfahren zur Überprüfung der Genauigkeit dieser

Korrelation zwischen den qualitativen Aussagen der Studierenden und einem statistischen Indikator wie dem LQI, der auf einer rein „quantitativen“ Bewertung beruht, gibt es bisher nicht. Das Fehlen einer testtheoretischen Überprüfung der Wirksamkeit als Indikator für das fakultätsübergreifende Monitoring wird jedoch von mir nicht als Problem eingestuft, da mir

1. mittlerweile ausreichend Daten für einen Längsschnittvergleich vorliegen. Die Auswertung von knapp 190.000 Fragebögen aus gut 5.000 Veranstaltungen aus vier Semestern zeigt eine sehr hohe Validität des LQI, mit nur geringen Schwankungen im Dezimalbereich, die sich auch alle kausal aus den Ergebnissen der Evaluationen erklären lassen, und
2. allein auf der Grundlage des LQI keine handlungsrelevanten Schlüsse gezogen werden. Der LQI verweist nur auf bestimmte Veranstaltungen, deren umfangreiche Einzelergebnisse von Fakultätsvorständen, Studiendekanen sowie die Veranstaltung verantwortenden Lehrenden einer genaueren Prüfung unterzogen werden sollen. Er trennt gewissermaßen die Spreu vom Weizen und hilft die große Masse an Veranstaltungsevaluationen, die jedes Semester am KIT durchgeführt werden, in verschiedene Follow-Up-Gruppen zu sortieren und die Komplexität der empirischen Ergebnisse auf ein überschaubares Maß zu reduzieren. Erst in der nachfolgenden Follow-Up-Phase werden in einem „qualitativen“ Bewertungsprozess Maßnahmen festgelegt und durchgeführt.

4.1 Berechnung des LQI

Die Berechnung des Lehrqualitätsindex nutzt eine integrierte Auswertungsfunktion der bei der Veranstaltungsevaluation am KIT eingesetzten Software EVASYS der Firma ElectricPaper GmbH.

Der LQI wird aus den Mittelwerten von fünf intervallskalierten unterschiedlich gewichteten Kernfragen berechnet.⁴

Abbildung 1: Lehrqualitätsindex (LOI): verwendete Fragen und Einzelgewichtungen

Qualitätsrichtlinie	Kernfrage im Fragebogen	Gewichtung im LQI
QR1: Gesamte Note	Bitte benoten Sie die Lehrveranstaltung insgesamt (sehr gut/ mangelhaft)	50%
QR2: Notwendiger Arbeitsaufwand	Der notwendige Arbeitsaufwand für diese Lehrveranstaltung ist... (angemessen/ unangemessen)	12,5%
QR3: Struktur der Lehrveranstaltung	Wie ist die Lehrveranstaltung strukturiert (sehr gut/ mangelhaft)	12,5%
QR4: Engagement und Motivation des Dozenten	Wirkt der Dozent (bzw. Übungsleiter) engagiert und motiviert bei der Durchführung der Veranstaltung (sehr stark/ überhaupt nicht)	12,5%
QR5: Eingehen des Dozenten auf Fragen und Belange der Studierenden	Geht der Dozent (bzw. Übungsleiter) auf Fragen und Belange der Studierenden ein? (sehr stark/ überhaupt nicht)	12,5%

Die Auswahl der Fragen, die Anzahl und die Gewichtung im LQI sind für das Monitoring veränderbar. Entscheidend bei der Zusammenstellung der LQI-Fragen ist jedoch, dass recht unterschiedliche und für alle Veranstaltungen vergleichbare Aspekte abgebildet werden. Fragen, die keine Gültigkeit für alle evaluierten Veranstaltungen haben, wie

⁴ Als „Kernfragen“ werden die Fragen bezeichnet, die gemäß der Evaluationsordnung für Lehre, Studium und Weiterbildung, § 5, Abs. 2, Satz 1 obligatorisch für alle Veranstaltungsfragebögen im KIT vorgegeben sind, um eine fakultätsübergreifende Vergleichbarkeit bei den Ergebnissen herstellen zu können.

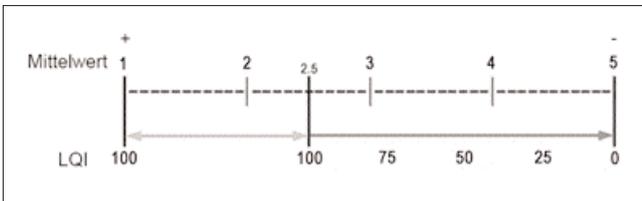
z.B. die Transparenz von Leistungs- und Prüfungsanforderungen, können nicht verwendet werden. Ein hoher Grad an Unterschiedlichkeit der ausgewählten Fragen wirkt sich positiv auf die Höhe der Korrelation zur Gesamtbewertung der Veranstaltung aus.

Diese oder ähnliche Fragen findet man in fast jedem Veranstaltungsfragebogen beliebiger Hochschulen. Dadurch lässt sich das Monitoring unkompliziert in anderen Kontexten anwenden, auch rückwirkend, wenn eine Intervallskalierung vorliegt.

Dass wir bei den Fragen zum Monitoring den Begriff „Qualitätsrichtlinie“ verwenden, wurde in der Einführungsphase des Monitoring öfter kritisiert. Insbesondere der empfundene „Notwendige Arbeitsaufwand“ sei keine Qualitätsrichtlinie aus der Sicht der Lehrenden. An dieser Stelle möchte ich daher noch einmal auf unseren Qualitätsbegriff hinweisen, der nicht die Sicht der Lehrenden abbildet, sondern u.a. die Studierbarkeit aus der Perspektive der Studierenden.

Zur Berechnung des LQI wird für die Fragen zuerst ein „Mindestqualitätsstandard“ festgelegt, der einem mindestens zu erreichenden arithmetischen Mittelwert für jede Einzelfrage entspricht. Am KIT wurde hierfür der arithmetische Mittelwert 2,5 bei einer 5er-Intervallskala festgelegt. Die positive Merkmalsausprägung liegt bei allen Fragen bei 1, die negative bei 5.

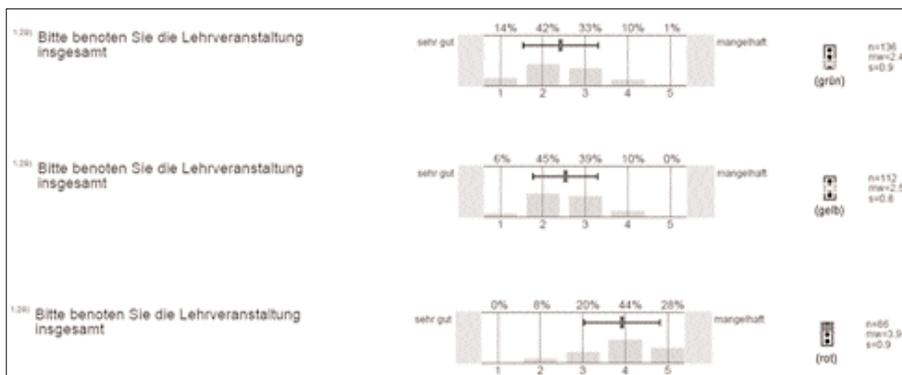
Abbildung 2: Berechnung des LQI



Die Festlegung des Mindestqualitätsstandards auf den Mittelwert 2,5 ist natürlich veränderbar, für unser Berichtswesen macht eine Anpassung zur Zeit jedoch keinen Sinn, dazu aber in der abschließenden Diskussion (vgl. Abschnitt 6) mehr.

Allen Einzelfragen mit einem arithmetischen Mittelwert besser als 2,5 ($< 2,5$) wird ein LQI von „100“ zugeordnet, ist der Mittelwert schlechter als 2,5 ($= 2,5 = 5$) sinkt der LQI linear zum erreichten Mittelwert bis „0“ (vgl. Abbildung 2).

Abbildung 3: Darstellung des LQI in den Einzelergebnissen



Der LQI der Einzelveranstaltung wird dann aus den 5 „LQI“-Fragen mit der in Abbildung 1 dargestellten Gewichtung berechnet.

4.2 Der LQI in den Einzelergebnissen für die Lehrenden

In der Auswertung ihrer Veranstaltungsbewertung erkennen die Lehrenden die 5 LQI-Fragen an einer kleinen Ampel neben der Grafik mit den Ergebnissen der Frage. Die 5er-Intervallskalen werden zur Darstellung der Ampel in 3 Unterintervalle eingeteilt.

- von 1 bis kleiner 2.5
- 2.5 bis kleiner 3.5
- 3.5 bis 5

Liegt der Mittelwert zwischen $1 < 2.5$ ist die Ampel grün. Liegt der Mittelwert zwischen $2.5 < 3.5$ ist die Ampel gelb und bei einem Mittelwert ab 3.5 wird die Ampel rot:

4.3 Die Zuordnung der Veranstaltungen in „Follow-Up Gruppen“ und „Follow-Up-Bereiche“

Die Erzeugung von *Follow-Up* ist das wichtigste Ziel für die Qualitätssicherung und -entwicklung. Da aber die große Mehrheit der Veranstaltungen unkritisch ist und von den Studierenden auch häufig hervorragend bewertet werden, macht ein flächendeckender Aufruf zur Qualitätssicherung aller Einzelveranstaltungen wenig Sinn. Das KIT-Modell der fakultätsübergreifenden Veranstaltungsevaluation will *Follow-Up* gezielt da erzeugen, wo es anhand der Evaluationsergebnisse wirklich notwendig ist.

Evaluationsberichte enthalten grundsätzlich Handlungsempfehlungen, die auf der Grundlage der empirischen Ergebnisse erstellt werden. Bei Veranstaltungsevaluationen sind differenzierte Handlungsempfehlungen für jede Einzelveranstaltung jedoch nicht möglich, da unter der Bedingung von Massenevaluationen der Arbeitsaufwand von den Evaluatoren nicht zu leisten ist. An die Lehrenden werden i.d.R. automatisiert erstellte deskriptive statistische Ergebnisse der Evaluation versendet, möglichst zeitnah nach der Evaluation um ggf. eine formative Nachsteuerung zu ermöglichen, jedoch ohne Interpretation und Empfehlungen.

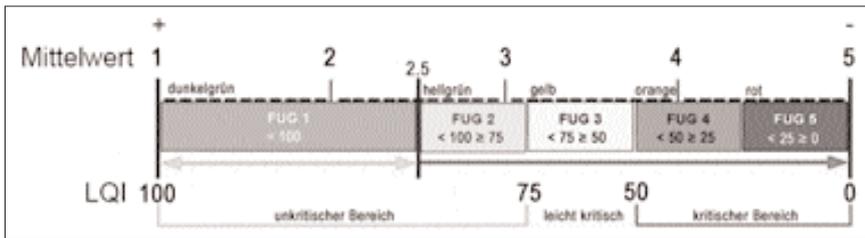
Das Monitoring ermöglicht jedoch im Berichtswesen für die Fakultätsvorstände, Studienkommissionen und Studiendekane eine Verknüpfung der Veranstaltung mit strukturiert vorgegeben Handlungsempfehlungen, die in einem Diskussionsprozess - in der Regel in der Studienkommission einer Lehrereinheit - an die Bedürfnisse der einzelnen Veranstaltung angepasst werden. Verpflichtend zu beachten sind die Handlungsempfehlungen zwar nicht, dennoch werden sie am KIT allgemein von allen Fakultätsvorständen beachtet und umgesetzt.

Im Verlauf der Einführung des Monitoring hat sich auch auf der Ebene der Fakultätsleitungen eine hohe Akzeptanz für das Verfahren herausgebildet, die einer der wichtigsten Pfeiler für die Qualitätskultur im Bereich Studium und Lehre des KIT geworden ist.

Jede Einzelveranstaltung wird anhand des erreichten LQI einer von 5 verschiedenen Follow-Up-Gruppen zugeteilt. Zu jeder Follow-Up-Gruppe

gehört eine eindeutige Farbe, mit der die Gruppen in allen Ergebnisdiagrammen in den Berichten dargestellt werden. Im Endbericht für Fakultätsvorstände und dem Bereichsvorstand Lehre des KIT werden die 5 Follow-Up-Gruppen zur Reduktion der Informationsvielfalt in die drei *Follow-Up-Bereiche* „unkritisch“, „leicht kritisch“ und „kritisch“ zusammengefasst. Für jede der fünf *Follow-Up-Gruppen* gibt es darüber hinaus differenzierte Empfehlungen, auch zu den Angeboten zur hochschuldidaktischen Weiterbildung im Hochschuldidaktischen Zentrum Baden-Württemberg (HDZ BW) (vgl. Abschnitt 5.5).

Abbildung 4: LQI: Zuordnung zu Follow-Up-Gruppen (FUG) und Follow-Up-Bereiche



Follow-Up-Empfehlungen für die *Follow-Up-Bereiche* „unkritisch“, „leicht kritisch“, „kritisch“:

- Der ersten Follow-Up-Gruppe (FUG 1) werden die Veranstaltungen zugeordnet, die bei allen 5 Kernfragen eine LQI von „100“ erreicht haben. Die FUG 2 umfasst die Veranstaltungen, deren durchschnittlicher LQI zwischen „75 und 100“ liegen. In Bezug auf die Ergebnisse bewerten wir diese beiden Gruppen als „unkritische Veranstaltungen“ (Dunkelgrün/Hellgrün).
- Bei den Veranstaltungen der FUG 3 liegt der durchschnittliche LQI zwischen „50 und 75“. In Bezug auf die Ergebnisse bewerten wir diese als „leicht kritisch“. Hier empfehlen wir eine Einsicht der Einzelergebnisse durch die Fakultätsleitungen (Gelb).
- Bei den Veranstaltungen der FUG 3 und 4 liegt der durchschnittliche LQI zwischen „50 und 0“. In Bezug auf die Ergebnisse bewerten wir diese Veranstaltungen als „kritisch“. Bei diesen Veranstaltungen empfehlen wir den Fakultätsleitungen eine ausführliche Analyse der Einzelergebnisse ggf. verbunden mit nachhaltigen Maßnahmen (Orange/Rot). Alle Veranstaltungen in der 5. Follow-Up-

Abbildung 5: Darstellung des LQI im Bericht für den Fakultätsvorstand, Studiendekane und die Studienkommission am Beispiel Vorlesungen

Vorlesung	Qualitätsrichtlinien (QR 1- QR 5)	Lehrqualitätsindex	Anzahl Fragebögen	Follow-Up Gruppe
Vorlesung 1	●●●●●	100,0	89	1
Vorlesung 2	●●●●●	100,0	10	1
Vorlesung 3	●●●●●	100,0	51	1
Vorlesung 4	●●●●●	97,2	119	2
Vorlesung 5	●●●●●	95,8	31	2
Vorlesung 6	●●●●●	89,1	125	2
Vorlesung 7	●●●●●	61,6	166	3
Vorlesung 8	●●●●●	23,5	96	5

Gruppe (Rot) werden zusätzlich dem Bereichsvorstand Lehre des KIT mitgeteilt, der sie in einem persönlichen Gespräch mit den zuständigen Studiendekanen bespricht.

4.4 Der LQI in den Berichten für den Fakultätsvorstand, Studiendekane und Studienkommission

In den Berichten für Fakultätsvorstand, Studiendekane und Studienkommission wird der LQI zunächst in einer Tabelle dargestellt. Für jeden Veranstaltungstyp (Vorlesung, Übung, Seminar etc.) gibt es eine eigene Ergebnisdarstellung. Die einzelnen Spalten enthalten den Veranstaltungstitel, die 5 Qualitätsrichtlinien, den LQI der Einzelveranstaltung, der aus den 5 LQI-Fragen berechnet wird, die Anzahl der Fragebögen und die Zuordnung zu einer Follow Up-Gruppe. Die 5 Qualitätsrichtlinien, die jeweils einer Frage aus dem Fragebogen zugeordnet sind, werden als einzelner Kreis dargestellt. Liegt der Mittelwert bei einer Frage innerhalb des 1. Unterintervalls ($1 < 2.5$) ist der gesamte Kreis Grün (hier hellgrau dargestellt). Liegt der Mittelwert außerhalb

des 1. Intervalls, wird der Kreis in vier Kreissegmente aufgeteilt. Je schlechter der erreichte Mittelwert, desto mehr Kreissegmente verfärben sich in Rot (hier dunkelgrau dargestellt). Anhand der Tabelle lässt sich „auf einen Blick“ erkennen, welche Vorlesungen die Mindestqualitätsstandards erfüllen und welche sie nicht erreichen. Über die Spalte „Follow-Up Gruppe“ werden zusätzlich die standardisierten Empfehlungen abgegeben, wie mit den Ergebnissen umgegangen werden kann (vgl. Erläuterung zu Abbildung 4). Über einen mit besonderen Rechten versehenen Zugang zu EVASYS können die Fakultätsvorstände und Studiendekane der Fakultät zudem gezielt die Einzelergebnisse der Veranstaltungsbeurteilungen sowie die Freitextkommentare der Studierenden einsehen, um sich ein differenziertes Bild zu den Einzelergebnissen als die Grundlage für die Diskussion der Ergebnisse in der Studienkommission zu verschaffen.

Bei der zusammenfassenden Ergebnisdarstellung werden die Häufigkeiten aller evaluierten Veranstaltungen zu den oben beschriebenen Follow-Up-Gruppen (FUG) dargestellt. So lässt sich für alle Lehreinheiten differenziert nach Veranstaltungsformen erkennen, wie groß die Anteile unkritischer Veranstaltungen (dunkelgrün, hellgrün), leicht kritischer Veranstaltungen (gelb) und kritischer Veranstaltung (orange, rot) insgesamt sind. Zusätzlich wird der Gesamt-LQI früherer Evaluationen für die Lehreinheit mit ausgegeben, um potentielle Veränderungstendenzen aufzuzeigen, deren Ursachen sich im direkten Vergleich mit den Ergebnissen der vorangegangenen Evaluationen differenziert analysieren lassen (vgl. Abbildung).

Bei der zusammenfassenden Ergebnisdarstellung werden die Häufigkeiten aller evaluierten Veranstaltungen zu den oben beschriebenen Follow-Up-Gruppen (FUG) dargestellt. So lässt sich für alle Lehreinheiten differenziert nach Veranstaltungsformen erkennen, wie groß die Anteile unkritischer Veranstaltungen (dunkelgrün, hellgrün), leicht kritischer Veranstaltungen (gelb) und kritischer Veranstaltung (orange, rot) insgesamt sind. Zusätzlich wird der Gesamt-LQI früherer Evaluationen für die Lehreinheit mit ausgegeben, um potentielle Veränderungstendenzen aufzuzeigen, deren Ursachen sich im direkten Vergleich mit den Ergebnissen der vorangegangenen Evaluationen differenziert analysieren lassen (vgl. Abbildung).

Bei der zusammenfassenden Ergebnisdarstellung werden die Häufigkeiten aller evaluierten Veranstaltungen zu den oben beschriebenen Follow-Up-Gruppen (FUG) dargestellt. So lässt sich für alle Lehreinheiten differenziert nach Veranstaltungsformen erkennen, wie groß die Anteile unkritischer Veranstaltungen (dunkelgrün, hellgrün), leicht kritischer Veranstaltungen (gelb) und kritischer Veranstaltung (orange, rot) insgesamt sind. Zusätzlich wird der Gesamt-LQI früherer Evaluationen für die Lehreinheit mit ausgegeben, um potentielle Veränderungstendenzen aufzuzeigen, deren Ursachen sich im direkten Vergleich mit den Ergebnissen der vorangegangenen Evaluationen differenziert analysieren lassen (vgl. Abbildung).

4.5 Der LQI im Bericht für den Bereichsvorstand Lehre und die Senatskommission Studium und Lehre

Im Bericht an die Senatskommission für Studium und Lehre und den Bereichsvor-

Abbildung 6: Aggregierte Darstellung des LQI im Bericht für den Fakultätsvorstand, Studiendekane und die Studienkommission

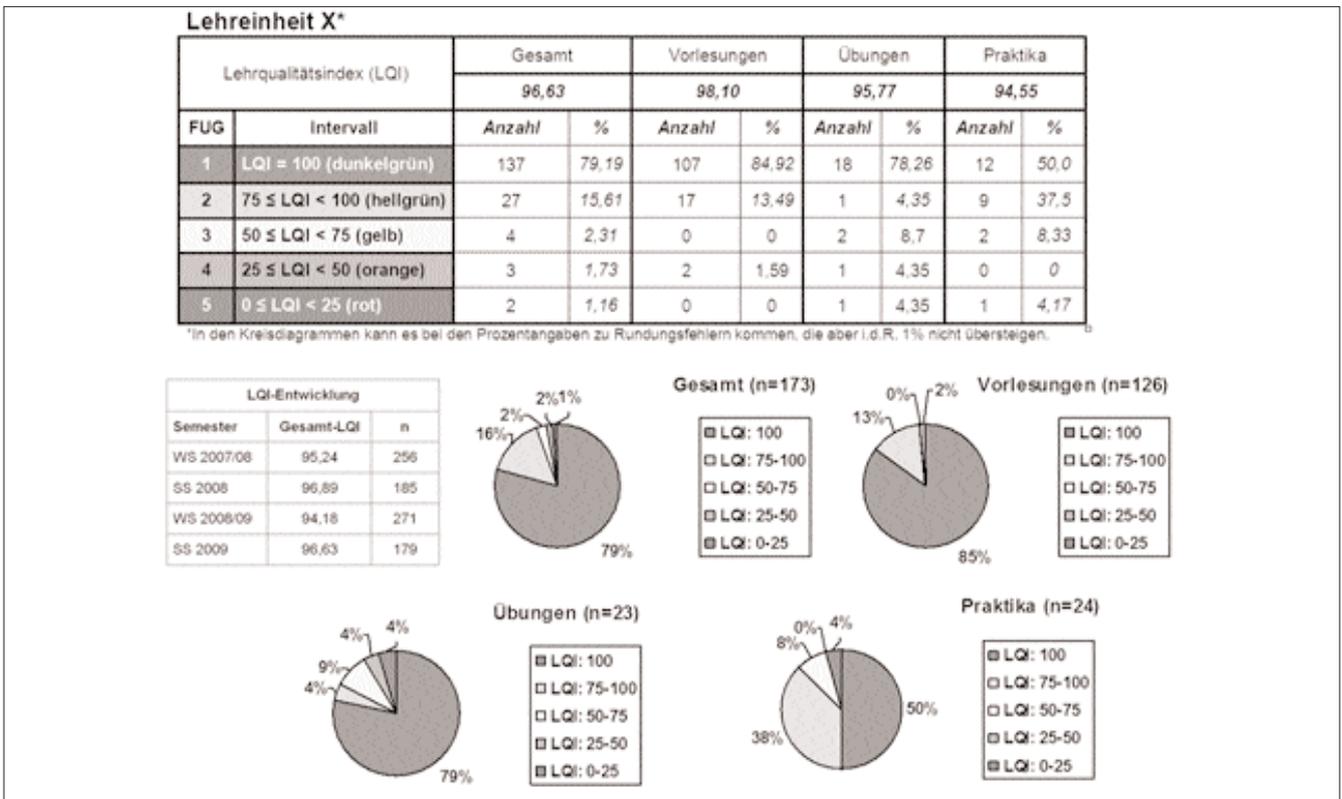
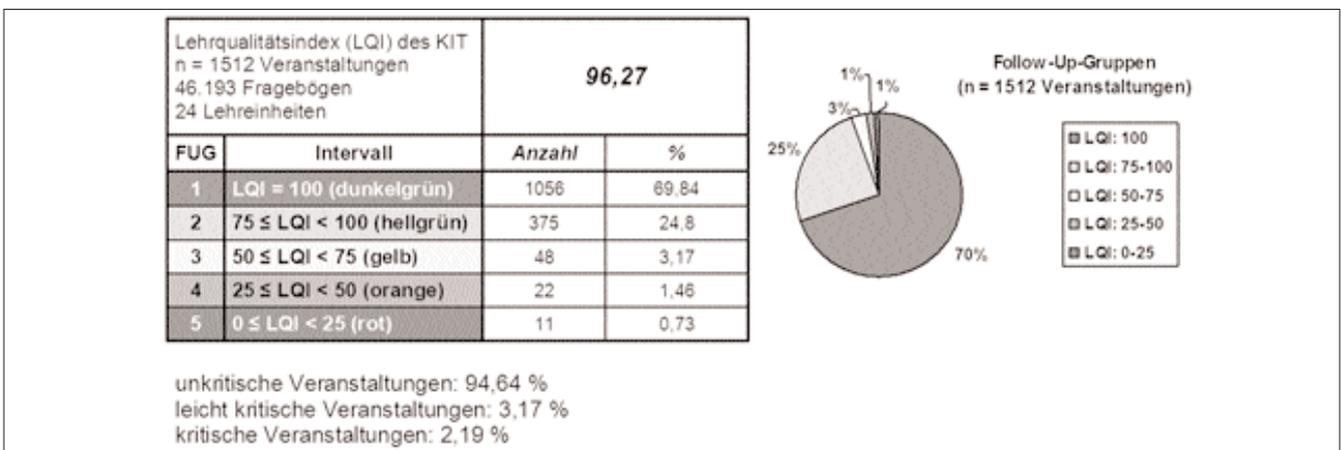


Abbildung 7: LQI des KIT - Sommersemester 2009



stand Lehre verändert der LQI seine Funktion. Fungiert er für die Fakultätsvorstände, Studiendekane und Studienkommissionen vor allem als Veranstaltungsmonitoring, das es ihnen ermöglicht, gezielt Schwachpunkte in der Lehre zu erkennen und *Follow-Up* zu erzeugen, so gibt er dem Bereichsvorstand Lehre und seiner Senatskommission jedes Semester vor allem einen allgemeinen (anonymen) Überblick über den Zustand der Veranstaltungsqualität für alle Lehreinheiten. Entsprechend enthält dieser Bericht die aggregierten Darstellungen des LQI wie in Abbildung 6 für alle Lehreinheiten des KIT. Der Bericht schließt mit einer aggregierten tabellarischen Darstellung der Verteilungen aller evaluierten Veranstaltungen eines Semesters am KIT in die *Follow-Up-Gruppen* ab: vgl. Abbildung 7.

5. Erzeugung von Follow-Up

5.1 Ebene Veranstaltungen

Die Erzeugung von *Follow-Up* auf der Veranstaltungsebene setzt die Bereitschaft der Lehrenden voraus, die eigenen Evaluationsergebnisse wahrzunehmen und sie aktiv zur Gestaltung der Lehre zu nutzen. Diese Bereitschaft hat am KIT in den letzten Semestern mit dem Aufbau der Evaluationsstelle spürbar zugenommen. So melden sich jedes Semester viele Lehrende in der Evaluationsstelle, deren Veranstaltungen gar nicht in einer zu evaluierenden Stichprobe sind, und bitten uns um die Durchführung einer Evaluation ihrer Veranstaltung. Eine Beobachtung, die auch in den vielen Gesprächen des Autors mit den Studiendekanen des KIT

bestätigt wird ist, dass das Vorhandensein der flächendeckenden und regelmäßigen Evaluationen sich positiv auf die Veranstaltungsqualitäten auswirkt. Die Lehrenden strengen sich insgesamt mehr an, um auch bei den Evaluationen gut abzuschneiden.

Sehr förderlich für die Auseinandersetzung mit den Ergebnissen erweist sich auch der Umstand, dass die Lehrenden am KIT über die *Evaluationsordnung für Lehre, Studium und Weiterbildung* zur Reflexion ihrer Veranstaltung verpflichtet sind: „§ 11 Diskussion der Ergebnisse in der Lehrveranstaltung: Das Ergebnis der Lehrveranstaltungsevaluation wird von dem Lehrenden der Lehrveranstaltung im laufenden Semester in der Lehrveranstaltung vorgestellt und diskutiert.“

5.2 Ebene Fakultätsvorstand, Studienkommission, Studiendekan

Die Evaluationsberichte werden vom Leiter der Evaluationsstelle jedes Semester in einem persönlichen Gespräch den Studiendekanen der Lehreinheiten übergeben und die Ergebnisse und mögliche Folgemaßnahmen diskutiert. Der Studiendekan informiert alle Mitglieder des Fakultätsvorstandes und die Studienkommission über die Ergebnisse und verständigt sich über die einzuleitenden Maßnahmen der Qualitätssicherung (z.B. pers. Gespräch einzelner Lehrenden mit Fakultätsvorstand, Kontrolle der Evaluationsergebnisse einzelner Veranstaltungen über mehrere Semester, individuelle Empfehlungen zur hochschuldidaktischen Weiterbildung, Nichtverlängerung von Verträgen (bei Lehrbeauftragten), Vergabe von Lehrpreisen, Umstrukturierung des Lehrangebots etc.). Über die Maßnahmen wird der Leiter der Evaluationsstelle schriftlich oder mündlich informiert. Der Leiter der Evaluationsstelle bespricht mit dem Bereichsvorstand Lehre des KIT jedes Semester den Endbericht mit den aggregierten Ergebnissen für alle Lehreinheiten.

5.3 Ebene Bereichsvorstand Lehre und Senatskommission Studium und Lehre

Über alle Veranstaltungen, die einen LQI im „rot“-kritischen Bereich erreicht haben, tauscht sich der Bereichsvorstand Lehre mit den zuständigen Studiendekanen aus und bespricht die einzuleitenden Maßnahmen. Der Endbericht für den Bereichsvorstand Lehre wird jedes Semester in einer Sitzung der Senatskommission Studium und Lehre vorgestellt und diskutiert. Die Präsentation in der Senatskommission Studium und Lehre wird zur allgemeinen Diskussion über Verbesserungen im Evaluationsverfahren sowie zu Absprachen zwischen den Studiendekanen inter fakultätiver Studiengänge und Veranstaltungen genutzt.

5.4 Strukturierte Empfehlungen für die Follow-Up-Bereiche „kritisch“ und „leicht kritisch“

In der Erläuterung zu Abbildung 4 werden für die *Follow-Up-Bereiche* „kritisch“ und „leicht kritisch“ Empfehlungen an die Fakultätsleitungen ausgesprochen, diese Veranstaltungen im Fakultätsvorstand und in der Studienkommission zu diskutieren. Dieser Diskussionsprozess ist am KIT zwar nicht formell festgeschrieben, es wird aber dennoch von allen Fakultätsleitungen so gehandhabt.

Dabei sind grundsätzlich nicht die Evaluationsergebnisse alleine Gegenstand der Diskussion, sondern es handelt sich vielmehr um einen qualitativen Findungsprozess der betei-

ligten Diskussionspartner aus Studierendenvertretung und Fakultätsleitung für das richtige bzw. angemessene *Follow-Up* für eine spezielle Veranstaltung vor dem Hintergrund

1. der Evaluationsergebnisse der Einzelveranstaltung und insbesondere auch der Freitextkommentare der Studierenden,
2. der Anzahl der teilnehmenden Studierenden und des Veranstaltungsortes,
3. des Kontextes, in dem die Veranstaltung stattgefunden hat (Pflichtveranstaltung, Wahlpflichtveranstaltung, Interessenveranstaltung, Uhrzeit, Dauer etc.),
4. der tiefen Innenkenntnis der Lehreinheit, die nur die „Insider“ haben können, sowie
5. der persönliche Kenntnis der für die Veranstaltung verantwortlichen Lehrenden.

Auf dieser Diskussions-Ebene spielen nun auch die Bewertung der Lernbedingungen und der Lehrqualität als didaktische Kompetenz eine Rolle, die das Monitoring nicht berühren kann.

5.5 Strukturierte Empfehlungen für alle Follow-Up-Gruppen

Die fakultätsübergreifende Evaluation der Veranstaltungen ist ein Instrument der individuellen Entwicklung und des Aufbaus der Lehrkompetenz ebenso wie Teil der Personalentwicklung am KIT insgesamt. Grundsätzlich sollten alle Lehrenden des KIT in regelmäßigen Abständen bezüglich Ihrer Lehrtätigkeit reflektieren, sich austauschen und weiterbilden. Hierzu bietet ihnen das KIT durch die Arbeitsstelle Hochschuldidaktik (HDZ)/kww verschiedene Möglichkeiten und ein breites Angebot.

Für alle Lehreinsteiger bietet sich eine Teilnahme an Modul 1 des HDZ (bestehend aus dem Grundlagenworkshop Fit für die Lehre I + II, Praxisberatungen und kollegialer Lehrhospitation) an. Für erfahrene Lehrpersonen wird eine regelmäßige, mindestens 1 x jährliche Teilnahme an Weiterbildungsworkshops aus Modul 2 des HDZ, sowie auch regelmäßige Lehrhospitationen und Lehrcoachings empfohlen (günstigerweise sollte auch hier im Vorfeld das Modul 1 absolviert sein). Nach Abschluss des Modul 3 kann das Baden-Württemberg Zertifikat für Hochschuldidaktik erworben werden.

Entsprechend der Lehrevaluationsergebnisse empfehlen sich auch speziellere und individuelle Maßnahmen, welche im folgenden Bereich der Follow-Up-Gruppen erläutert werden. Eine dieser spezielleren Maßnahmen ist das Angebot einer „Praxisberatung zum Umgang mit Evaluationsergebnissen“.

1. Follow-Up-Gruppe (Farbe: Dunkelgrün)

LQI = 100

Erläuterung: In der 1. Follow-Up-Gruppe befinden sich die Veranstaltungen, die bei allen 5 Kernfragen, die zur Messung des Mindestqualitätsstandards dienen („Qualitätsrichtlinien“), einen LQI von 100 erreicht haben. Die Einzelergebnisse und die Freitextkommentare sind generell so positiv, dass wir diese Veranstaltungen als gänzlich unkritisch einstufen.

Empfehlung: Um die Qualität dieser Veranstaltungen mit hervorragenden Konzeptionen, denen es gelingt, die un-

terschiedlichen Lernbedürfnisse der Studierende zu integrieren, zu erhalten, wird wie bereits eingangs im allgemeinen Teil erläutert, eine regelmäßige Reflexion und der kollegiale Austausch, wie auch die regelmäßige jährliche Teilnahme an hochschuldidaktischen Workshops aus Modul 2 empfohlen. Zielsetzung: Reflektion des Erfolgs um die Lehrkompetenz weiter auszubauen.

2. Follow-Up-Gruppe (Farbe: Hellgrün)

LQI: < 100 > 75

Erläuterung: In der 2. Follow-Up-Gruppe befinden sich die Veranstaltungen, die zwar nicht bei allen Qualitätsrichtlinien den vollen LQI von 100 erreicht haben, die aber insgesamt eine gute Bewertung bekommen. Die Einzelergebnisse und die Freitextkommentare sind auch in dieser Gruppe überwiegend positiv, so dass wir diese Veranstaltungen gleichfalls als unkritisch einstufen.

Empfehlung: Den Lehrenden wird empfohlen, die Ergebnisse ihrer Evaluation, insbesondere die Freitextkommentare und das verpflichtende Gespräch mit den befragten Studierenden über die Ergebnisse, verstärkt zu nutzen und Ihre Veranstaltung auf diese Grundlage zu reflektieren und ggf. Verbesserungen vorzunehmen.

Des Weiteren wird wie bereits eingangs im allgemeinen Teil erläutert, eine regelmäßige Reflexion und der kollegiale Austausch (Praxisberatung), wie auch die regelmäßige jährliche Teilnahme an hochschuldidaktischen Workshops aus Modul 2 empfohlen. Wichtig ist es hier seine gemachten Erfahrungen systematisch auszuwerten, zu reflektieren und die eigene Lehrkompetenz weiter auszubauen.

3. Follow-Up-Gruppe (Farbe: Gelb)

LQI: < 75 > 50

Erläuterung: In der 3. Follow-Up-Gruppe befinden sich die Veranstaltungen, die i.d.R. bei der Mehrheit der Qualitätsrichtlinien deutlich weniger als den vollen LQI von 100 erreicht haben. Die Einzelergebnisse sind in dieser Gruppe oft sehr ambivalent und bei den Freitextkommentaren sind negative und positive Kommentare in der ganzen Breite vorhanden. Diese Veranstaltungen erfüllen oft nur die Lernbedürfnisse eines Teils der Studierenden.

Empfehlung: Eine Einsicht in die Einzelergebnisse durch die Fakultätsvorstände und Studiendekane ist empfehlenswert. Anhand der Einzelergebnisse kann dann entschieden werden, ob z.B. persönliche Gespräche mit den Lehrenden über die Veranstaltung sinnvoll sind.

Den Lehrenden wird empfohlen, die Ergebnisse ihrer Evaluation, insbesondere die Freitextkommentare und das verpflichtende Gespräch mit den befragten Studierenden über die Ergebnisse, verstärkt zu nutzen und Ihre Veranstaltung auf dieser Grundlage zu verbessern. Des Weiteren wird zu einem verstärkten Austausch unter Kollegen, gerne auch mit Unterstützung durch das HDZ/kww geraten. Regelmäßiges Reflektieren und die Teilnahme am hochschuldidaktischen Weiterbildungsprogramm sollten verstärkt erfolgen. Es kann auch überlegt werden, sich eine Einzelberatung im HDZ zu buchen. Zielsetzung wäre hier die Analyse der Lehrveranstaltungen, ihrer Rahmenbedingungen und der Lehransatz des Lehrenden.

4. Follow-Up-Gruppe (Farbe: Orange)

LQI: < 50 > 25

Erläuterung: In der 4. Follow-Up-Gruppe befinden sich die Veranstaltungen, die bei fast allen Qualitätsrichtlinien eine schlechte Bewertung durch die Studierenden bekommen haben.

Die Einzelergebnisse und die Freitextkommentare sind in dieser Gruppe sehr oft negativ. In diesen Veranstaltungen wird kaum auf die Lernbedürfnisse der Teilnehmer eingegangen.

Empfehlung: Eine ausführliche Analyse der Einzelergebnisse und der Struktur des Lehrbetriebs durch die Fakultätsvorstände und Studiendekane wird dringend empfohlen. In der Fakultätsleitung sollte dann über nachhaltige Maßnahmen wie den Abschluss von Zielvereinbarungen mit den Lehrenden nachgedacht werden. Dieses Gespräch wird gerne von Seiten des HDZ/kww beratend unterstützt. Die Hinzunahme liegt im Ermessen des Studiendekans.

Den Lehrenden wird unbedingt geraten die Ergebnisse ihrer Evaluation, insbesondere die Freitextkommentare und das verpflichtende Gespräch mit den befragten Studierenden über die Ergebnisse, intensiv zu nutzen und Ihre Veranstaltung auf dieser Grundlage zu verbessern. Die verstärkte Auseinandersetzung mit der eigenen Lehre ist unabdingbar. Die regelmäßige Teilnahme an hochschuldidaktischer Weiterbildung (beginnend mit Modul 1) mit vorhergehender Beratung wird empfohlen.

5. Follow-Up-Gruppe (Farbe: Rot)

LQI: < 25 > 0

Erläuterung: In der 5. Follow-Up-Gruppe befinden sich die Veranstaltungen, die i.d.R. bei allen Qualitätsrichtlinien eine sehr schlechte Bewertung durch die Studierenden bekommen haben.

Die Einzelergebnisse und die Freitextkommentare sind in dieser Gruppe fast ausschließlich negativ. In diesen Veranstaltungen wird kaum bis gar nicht auf die Lernbedürfnisse der Teilnehmer eingegangen.

Empfehlung: Eine ausführliche Analyse der Einzelergebnisse und der Struktur des Lehrbetriebs durch die Fakultätsvorstände und Studiendekane wird dringend empfohlen. In der Fakultätsleitung sollte dann über nachhaltige Maßnahmen wie den Abschluss von Zielvereinbarungen mit den Lehrenden nachgedacht werden. Dieses Gespräch soll von Seiten des HDZ/kww beratend unterstützt werden.

Den Lehrenden wird unbedingt geraten die Ergebnisse ihrer Evaluation, insbesondere die Freitextkommentare und das verpflichtende Gespräch mit den befragten Studierenden über die Ergebnisse, intensiv zu nutzen und Ihre Veranstaltung auf dieser Grundlage zu verbessern. Hier kann die Moderation durch Experten des HDZ erfolgen. Die verstärkte Auseinandersetzung mit der eigenen Lehre ist unabdingbar. Die regelmäßige Teilnahme an hochschuldidaktischer Weiterbildung (beginnend mit Modul 1) mit vorhergehender Beratung wird nahegelegt.

6. Diskussion

Das Monitoring der Veranstaltungsqualität hat nach den ersten öffentlichen Vorträgen unter den Fachkollegen zu lebhaften Diskussionen und auch konkreten Fragen geführt, von denen ich hier die wichtigsten beantworten möchte.

Eine der am häufigsten gestellten Fragen war, ob die Organisation der Evaluationen als zentrale Evaluation, wie dies am KIT der Fall ist, nicht zu starken Vorbehalten aus den Fachbereichen geführt und wie lange es gedauert hat, bis alle Fachbereiche die zentrale Evaluation unterstützt und akzeptiert haben. Begonnen haben wir mit der Einführung im Wintersemester 07/08 und erhebliche Vorbehalte hat es natürlich gegeben. Aber durch einen sehr engen persönlichen Austausch mit den Fakultätsleitungen (der hier dauerhaft aufrechterhalten wird und eine zentrale Rolle bei der allgemeinen Akzeptanz der zentralen Evaluation spielt), konnte von mir vermittelt werden, dass die zentrale Evaluation eine wissenschaftliche Dienstleistungseinrichtung für die Fachbereiche ist und die Hauptverantwortung für die Lehre und deren Qualität in den Fachbereichen verbleibt. Nur bei den Veranstaltungen, bei denen die dezentrale Steuerung versagt, greift der Bereichsvorstand Lehre des KIT ein. Dies ist aber so gut wie nie der Fall, und wenn es in der Vergangenheit vorkam, dann auf Wunsch der Fakultäten selbst.

Eine Frage bei der allgemeinen Akzeptanz spielt natürlich auch der Arbeits- und Kostenfaktor. Am KIT organisiert die Evaluationsstelle mit wenigen Mitarbeitern alle internen Evaluationen, nicht nur die Veranstaltungsevaluationen, sondern auch die fakultätsübergreifenden Absolventenbefragungen, Studieneingangsbefragungen, Befragungen des wissenschaftlichen Nachwuchses u.a. Dies bedeutete nicht nur eine erhebliche Minimierung des Arbeitsaufwandes für die Fachbereiche, sondern auch eine erhebliche Kostenreduktion für alle, mit einem erheblichen Mehrwert an gut aufbereiteten Informationen.

Die Phase des Aufbaus der zentralen Evaluation hat 2 Jahre gedauert, doch bereits nach weniger als einem Jahr waren alle Fakultäten in die Evaluationsprozesse integriert. Das zweite Jahr diente v.a. der Verfahrensverbesserung und der Optimierung des Berichtswesens sowie der Integration der zentralen Einrichtungen mit Lehrangeboten.

Eine Kritik war, ob man den Zustand der Lehre am KIT nicht einfach schön reden würde, denn der Anteil der unkritischen *Follow-Up-Bereiche* sei doch recht hoch. Es ist auch richtig, dass ca. 95 % der Veranstaltungen aus dem strukturierten *Follow-Up* herausfallen, aber das entbindet nicht die Lehrenden mit unkritisch bewerteten Veranstaltungen von ihrer Verantwortung, ihre Evaluationsergebnisse zu reflektieren und zu diskutieren. Auch für die Bereiche der unkritischen Lehre werden vielfältige Entwicklungsmöglichkeiten zur Verfügung gestellt, z.B. über unsere umfangreichen hochschuldidaktischen Weiterbildungsmöglichkeiten, die alle *Follow-Up-Gruppen* ansprechen sollen. Darüber hinaus richten wir unsere Aufmerksamkeit und unsere Außendarstellung bei der Lehre gerade nicht auf die unkritischen Bereiche, sondern auf die Verbesserung der Kritischen. Die Sicherung und Entwicklung der Prozessqualität der Lehre ist ein dauerhafter Prozess, mit dem man versu-

chen kann, ein hohes Niveau dauerhaft zu erreichen und zu halten, einen perfekten Zustand kann es jedoch nicht geben: Jeder Lehrende hat auch mal ein schlechteres Semester, nicht jede Veranstaltungskonzeption kann sofort optimal sein und neue Lehrende müssen sich auch erst „einlehren“, nur um ein paar Beispiele zu nennen. Durch das Monitoring wissen wir am KIT aber jedes Semester genau, wo die Prozesse nicht optimal gewesen sind und steuern hier systematisch entgegen. Auch wurde in diesem Zusammenhang nachgefragt, ob der am KIT festgesetzte Mindestqualitätsstandard, der arithmetische Mittelwert bei den LQI-Fragen von 2.5, nicht zu hoch ist und hier ein anderer Mittelwert, z.B. von 2.0, nicht sinnvoller wäre. Natürlich ist hier eine Anpassung möglich, in unserer Evaluationsordnung wird bei der in Kürze anstehenden Revision festgelegt, dass der Mindestqualitätsstandard vom Bereichsvorstand Lehre jederzeit geändert werden kann. Nur macht eine Änderung derzeit keinen Sinn, denn die Trennung der Veranstaltungen in die *Follow-Up-Gruppen* funktioniert damit z.Z. reibungslos und liefert uns jedes Semester ca. 5 % von allen evaluierten Veranstaltungen, die dringend einer *Follow-Up*-Diskussion bedürfen. Die Anzahl der nachzubesprechenden Veranstaltungen ist gerade noch von den Fachbereichen zu bewältigen. Eine Verschärfung des Mindestqualitätsstandards birgt daher im Moment die Gefahr, dass das *Follow-Up*-Verfahren nicht mehr praktikabel wird. Bei einer permanenten Qualitätsentwicklung werden die kritischen Veranstaltungen allmählich verschwinden und dann soll natürlich auch der Mindestqualitätsstandard daran angepasst werden.

Eine letzte häufig an mich gestellte Frage ist die nach der Abhängigkeit des Monitoring von der am KIT eingesetzten automatisierten Evaluationssoftware EvaSys der ElectricPaper GmbH. Als empirisches Verfahren ist die Berechnung des LQI natürlich erstmal unabhängig von der Software und auch anders technisch realisierbar. Für eine Volluniversität wie das KIT mit regelmäßigen Massenevaluationen unter ca. 19.000 Studierenden sehe ich allerdings zu EvaSys aktuell keine Alternative. Daher sind eine Integration der Berechnungsmöglichkeit des LQI nach dem KIT-Modell sowie die automatisierte Einteilung der Veranstaltungen anhand des LQI in die *Follow-Up-Gruppen* in EvaSys geplant. Das KIT kooperiert diesbezüglich mit ElectricPaper. Bereits nach der kurzen Etablierungsphase von etwas mehr als 2 Jahren findet das KIT-Modell der fakultätsübergreifenden Evaluation der Veranstaltungen in der derzeitigen Form auf allem Stausebenen des KIT eine sehr hohe Akzeptanz. Die große Stärke des dazu gehörigen Monitoring liegt v.a. auch in der leichten Anpassbarkeit des Verfahrens an die Bedürfnisse anderer Bildungseinrichtungen und Fachbereiche. Hier sei abschließend noch einmal darauf hingewiesen, dass das Monitoring bestehende Evaluationsverfahren im Bereich der Lehre mit dem Ziel einer deutlichen Verbesserung des *Follow-Up* ergänzen kann, sie aber nicht ersetzen soll.

Danksagung

An dieser Stelle möchte ich mich abschließend herzlich bei allen Beteiligten bedanken, die die Entwicklung des hier vorgestellten KIT-Modells ermöglicht haben. Mein Dank gilt zunächst dem gesamten Präsidium des KIT und der ehemaligen Leiterin der Stabsabteilung Planung und Control-

ling, Frau Barbara Windscheid, für das Vertrauen und die freien Gestaltungsmöglichkeiten beim Aufbau der zentralen Evaluation am KIT. Besonderen Dank gilt auch meiner Mitarbeiterin Frau Cédrine Lussac, die mit ihrer Kreativität und Ihren tiefen Anwendungskenntnissen von EvaSys entscheidende Impulse für die Entwicklung des Monitoring eingebracht hat. Dank auch an Frau Anke Diez, die Leiterin der Personalentwicklung am KIT, die sich der hier vorgestellten Verknüpfung von Evaluation und hochschuldidaktischer Weiterbildung angenommen hat, Heinrich Schullerer und Andreas Lumpe von ZENDAS, der zentralen Datenschutz-

behörde der Baden-Württembergischen Universitäten, die mir geholfen haben, die datenschutzrechtlichen Aspekte zu integrieren und an die Fakultätsleitungen und Fachschaften am KIT, die meine Ideen tatkräftig unterstützt und erheblich zum Erfolg beigetragen haben.

■ **Dr. Michael Craanen**, Leiter der Evaluationsstelle in der Stabsabteilung Planung und Controlling, Karlsruher Institut für Technologie - KIT, E-Mail: michael.craanen@kit.edu

**Heinz W. Bachmann:
Systematische Lehrveranstaltungsbeobachtungen an einer Hochschule
Verläufe von Lehrveranstaltungen an einer schweizerischen Fachhochschule
bei Einführung der Bologna-Studiengänge – eine Fallstudie**

Seit Herbst 2006 bieten alle Fachhochschulen der Schweiz Studiengänge organisiert nach dem Bachelor-Master-System an, wie das in der Bologna-Deklaration beschlossen worden war. Einer der Haupttriebfedern des Reformprozesses, neben der akademischen Mobilität und der Vorbereitung der Hochschulabsolventen auf den europäischen Arbeitsmarkt, ist die Steigerung der Anziehungskraft der europäischen Hochschulen zur Verhinderung von brain drain und der Förderung von brain gain. Neben diesem globalen Wettbewerb wird durch die gegenseitige Anrechenbarkeit der Studienleistungen in den verschiedenen Ländern auch die Konkurrenz der Hochschulen untereinander gefördert.

Die Bologna-Reform geht von einem neuen Lehrverständnis aus von der Stoffzentrierung hin zu einer Kompetenzorientierung, begleitet von einem shift from teaching to learning. Der Fokus liegt also nicht beim Lehren, sondern auf der Optimierung von Lernprozessen. Vor dem Hintergrund neuerer Erkenntnisse aus der Lernforschung wird auch deutlich, dass das Vermitteln von Wissen im traditionellen Vorlesungsstil nur noch bedingt Gültigkeit hat. Unter Berücksichtigung der obigen Erkenntnisse müsste man heute eher vom Hochschullernen als von der Hochschullehre sprechen. Die vorliegende Studie wird zum Anlass genommen, ein Instrument vorzustellen, mit dem Lehre systematisch beobachtet werden kann. Mit dem beschriebenen Instrument wird der Frage nachgegangen, inwieweit an der untersuchten schweizerischen Pädagogischen Hochschule die oben beschriebene Neuorientierung in der Lehre schon stattgefunden hat. Mit Hilfe des VOS (VaNTH Observational System) sollen systematisch Lehrveranstaltungsbeobachtungen gemacht und festgehalten werden. Das Ziel dieser Studie ist es, Lehrveranstaltungsverläufe an der untersuchten Pädagogischen Hochschule zu erheben im Hinblick auf die Entwicklung von Kursen in Hochschuldidaktik. Die gefundenen Ergebnisse sollen mit der Schulleitung besprochen werden, vor allem auch auf dem Hintergrund des neuen Lernens an Hochschulen. Basierend auf den gewonnen empirischen Daten und den von der Schulleitung entwickelten Zielen können hochschuldidaktische Kurse geplant und umgesetzt werden. Zusätzlich besteht die Chance, bei einer Wiederholung der Studie in einigen Jahren mögliche Veränderungen in der Lehre festzustellen. Es wird davon ausgegangen, dass das Untersuchungsdesign und die erhobenen Daten nicht nur von Interesse für die betroffene Hochschule sind, sondern generell Fachhochschulen interessieren dürften, die in einem ähnlichen Prozess der Neuorientierung stecken.



ISBN 3-937026-65-7, Bielefeld 2009, 172 Seiten, 24.90 Euro

Reihe: Hochschulwesen - Wissenschaft und Praxis

Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

E. W. Udo Küppers

Das Bachelor-Studium aus systemischer Sicht



*E. W. Udo
Küppers*

Bildungssysteme sind in ihrer Komplexität äußerst schwer zu steuern. Der vorliegende Artikel will zeigen, wie das oft in seinen Wechselwirkungen unüberschaubare Bildungssystem aus neuer Perspektive transparent und steuerbar bzw. regelbar gestaltet werden kann. Dazu wird ein Modell vorgestellt, wie komplexe Systeme - hier beispielhaft das Bachelor-Studium - mit Hilfe systemischer Analyse, genauer: mit Hilfe von Wirkungsnetzen als Werkzeuge systemischen Bildungsdenkens analysiert und einer fehlertoleranten Lösung zugeführt werden kann. Damit behandelt diese Arbeit eine in der Debatte um die Reform der Bachelor-Studiengänge u.U. wichtige Thematik.

Die Notwendigkeit einer neuen Art des Denkens

Mit traumwandlerischer Sicherheit fanden sich ca. 50 Studierende verschiedener Fachdisziplinen im Sommersemester jeden Freitag spätnachmittags im fensterlosen Computerraum C der Fakultät B ein. Der spezielle Zugang war durch Absprache der Fakultät A (angestellte Lehrkraft) mit der Fakultät B (Betreiber des Computerraums C) und durch die Uni-Raumverwaltung gewährleistet.

Das ist zwar eine etwas komplizierte Kommunikation aber sie funktioniert.

Die Technik aller 40 Computer funktionierte weitgehend reibungslos, die Lüftungsanlage entzog dem Raum überschüssige Wärme und lebhafte Diskussionen zwischen Studierenden untereinander und mit der Lehrperson förderten eine entspannte Lern-Lehr-Atmosphäre. Wenn kleine Störungen im Ablauf auftraten dann wurden sie i.d.R. durch kausale – oft monokausale – Verfolgungen der Ursachen schnell und erfolgreich behoben. Alle denkbaren Voraussetzungen schienen erfüllt für einen störungsfreien Arbeitsablauf der nun seit Wochen anhielt. - Wirklich alle?

Es begann mit einem sehr heißen Freitag. Die Raumtemperatur im fensterlosen Computerraum war trotz offener Türen wenig erträglich. Die uralte Technik der Lüftungsanlage war – wie später bekannt wurde – exakt auf eine durchschnittliche Raumtemperatur, auf die Abwärme der Zahl Computer im Raum und auf die zugehörige Personenzahl abgestimmt! Auf die Überbelegung der Personenzahl im Raum bei hochsommerlichen Temperaturen war die Lüftungstechnik nicht ausgelegt. Erste Computer fallen aus. Technische Hilfe war nicht möglich, weil die zugehörigen Uni-Verwaltungsstellen Freitagnachmittags nicht mehr erreichbar waren. Der generelle Umzug in einen größeren Computerraum scheiterte an der ausschließlichen Softwareimplementierung (Eigentum der Fakultät A) im Com-

puterraum C der Fakultät B. Die Folge: Abbruch der Arbeiten wegen nicht zu behebender **unerwarteter Probleme!** Diese unerwarteten Probleme hielten noch 3 Wochen bis zum Semesterende an – und vergrößerten sich zudem noch durch Fehler im zentralen elektronischen Raum-Zugangssystem das einen Zugang verweigerte! Auch Reparatur und Kartentausch mit neuem Zugangscode blieben erfolglos. Das Semester wurde unter völlig unbefriedigenden Bedingungen halbwegs ordentlich für alle Beteiligten zu Ende geführt.

Was zeigt dieses erlebte Beispiel eines universitären Lehrbetriebs?

Trotz guten Willens werden die realen Zusammenhänge nicht hinreichend genau erkannt oder ausgeblendet. Das bedeutet jedoch noch lange nicht, dass sie wirkungslos bleiben oder sind. In vielen Fällen treten unerwartete Folgen zeitversetzt auf. Mit umso größerer Intensität dämpfen sie die durch kausales Denken und Handeln erreichten Ziele und Werte wieder oder machen sie zunichte. Die öffentlichen Medien sind voll von Beispielen aus politischen, sozialen, wirtschaftlichen und naturnahen Aktivitäten, die kurzzeitige Erfolge versprechen und zum Teil auch realisieren, aber nachhaltig katastrophale Wirkungen auslösen. Auch die Struktur des Bachelor-Studiums, an der seit nunmehr 10 Jahren mit groben und feinen „Feilen“ modelliert wird, ist von solchen Folgeproblemen nicht ausgeschlossen. Dazu später mehr. Abbildung 1 zeigt eine Lehrveranstaltung als komplexes System.

Abbildung 1 - Quelle: FS-Uni Jena Archiv



Lehrveranstaltung als komplexes System

Jedes Zusammentreffen von Menschen – hier Studierende, Lehrende, wissenschaftliche und verwaltungstechnische Mitarbeiter und andere – ob im Umfeld einer Bildungseinrichtung oder anderswo ist ein komplexes System. Weil es vielschichtig ist, sich mit der Zeit ändert und wesentliche Eigenschaften des Gesamtsystems Lehrveranstaltung nicht durch noch so detaillierte Einzelanalysen erfasst werden können. Wissenschaftlich ausgedrückt spricht man von **Emergenz**, alltagssprachlich von: **Das Ganze ist mehr als die Summe der Teile** (siehe u.a. Lewin 1996, S. 24f.; Richter et.al. 2002, S. 3)

Sind wir überhaupt noch in der Lage, die Turbulenzen im Bildungsbereich realistisch wahrzunehmen, sie vorteilhaft zu steuern geschweige denn zu regeln? Und in welche Richtung drehen wir an den sensiblen „Stellschrauben“ komplexer Bildungssysteme, um dem Ziel einer nachhaltigen und sozial ausgewogenen Bildungsentwicklung näher zu kommen? Bildungssysteme sind in ihrer Komplexität äußerst schwer zu lenken. Der vorliegende Artikel will zeigen, dass das in seinen Wechselwirkungen oft unüberschaubare Bildungssystem aus neuer Perspektive mit Hilfe systemischer Analyse transparenter, problemvorbeugender und somit lenkbarer gestaltet werden kann.

Vielfach entscheidet die **richtige Kommunikation** mit der Umwelt über die Häufigkeit und Bewältigung von Problemen und über unsere Zukunft. Durch systemisches Denken wird der so genannte „Tunnelblick“, das „Scheuklappenverhalten“, die „Silomentalität“, die *organisatorische Kurzsichtigkeit* vermieden, weil die Gegenstände der Betrachtung und Bearbeitung aus einer „übergeordneten“ Rundherumblickrichtung beleuchtet werden. Dies ist ein Erfolg versprechender Weg, komplexe Zusammenhänge realistisch zu analysieren und zu bewerten. Und dieser ist auch im Bildungsbereich unerlässlich.

Diese Überblicksicht bedeutet aber keinesfalls eine Vernachlässigung der Details. Wesentlich ist, unter dem „Mantel“ des Überblicks einzelne Einflussgrößen im Bildungsbe- reich und insbesondere ihre vernetzten Wirkungen unter- einander mit dem richtigen Gliederungsgrad und der richti- gen Wirkungsstärke zu erkennen.

Was bedeutet es systemisch zu denken?

Systemisches Denken analysiert und bewertet Zusammen- hänge in einer vernetzten dynamischen Umwelt. Dadurch besteht die Chance, komplexe Realitäten zu erkennen und zu nutzen statt sie – wie so oft – zu ignorieren. Die Natur ist durchsetzt von **komplexen Systemen**. Sie zu verstehen und sich darin nachhaltig weiterzuentwickeln haben die Orga- nismen über Jahrmilliarden perfektioniert. Die Menschen sind – nicht nur - im Bildungsbereich noch weit davon ent- fernt, vorhandene komplexe Beziehungen auch nur annäh- ernd so elegant und fehlertolerant zu nutzen.

Komplexes System oder nur kompliziertes System?

Ein **komplexes System** ist ein vielschichtiges System, dessen Einzelteilanalyse nicht zum Verständnis des Gesamtsystems beiträgt. Insbesondere die Vernetzung vermeintlicher Ein- zelteile prägt ganz entscheidend die Eigenschaften des Ges- amtsystems. Komplexe Systeme besitzen ebenso eine Dy- namik der Entwicklung. Lebewesen als komplexe Organismen adaptieren sich an verändernde Umweltbedingung

derart geschickt, dass sie ihr Überleben nachhaltig optimie- ren. Die Dynamik der Anpasstheit in technischen, wirt- schaftlichen oder allgemein gesellschaftlichen komplexen Systemen verfolgt ihre spezifischen Zielvorgaben. Ob dies immer mit der richtigen Kommunikation oder wirksamen Lösungsinstrumenten zu *nachhaltigen* Erfolgen für das Ge- meinwesen führt – darüber sind wohl Zweifel angebracht. Was ein komplexes System ist, errahnen wir. Was ist aber demgegenüber ein kompliziertes System?

Kompliziert ist etwas, was verwirrend ist, undurchschaubar ist, schwierig zu verstehen ist, umständlich gestaltet ist oder unklar ist. Nicht, weil sich ein kompliziertes System ständig verändert – was dann komplex zu nennen wäre –, nein, weil es sich nicht mit den uns bekannten und ge- wohnten Mitteln erfassen oder nutzen lässt. Die technische Bedienungsanleitung für ein „Handy“ bzw. ein mobile phone, geschrieben in Taiwan und zu entziffern von eu- ropäischen „Handy“-Nutzern, ist ein typischer Fall. Bevor wir uns beispielhaft der Komplexität im Bachelor-Stu- dium zuwenden, wollen wir kurz und prägnant auf die we- sentlichen Ursprünge eingehen, die u.a. zu der weiter unten postulierten Anwendung des mächtigen Werkzeuges **Wirkungsnetz** zur Lösungen komplexer Aufgaben innerhalb und außerhalb des Bildungsbereiches geführt haben.

Komplexität ist überall und Wirkungsnetze erfassen die Realität

Die Natur ist die Mutter aller Komplexität! Noch heute zeigt ein viel beschworener Satz worauf es beim Erkenntnis- gewinn im Umgang mit der Natur wesentlich ankommt: Wer die Nahrungs(wirkungs)netze der Natur versteht, kennt ihr wahres Gesicht.

Ernst Haeckel hat 1866 die wissenschaftliche Grundlage für die heutige Ökologie erarbeitet, indem er Darwins Begriff „Kampf ums Dasein“ auf die „... **gegenseitigen Wechselbe- ziehungen der Organismen** ...“ bezog (Begon et.al. 1998, XVII), deren System-Verstehen maßgeblich durch soge- nannte Wirkungsnetz-Analysen erfassbar ist. Natürliche Wirkungsnetze sind daher maßgebend verantwortlich für die höchst effektiven und effizienten Leistungen der Evolu- tion! Die Natur- und später die Ingenieurwissenschaften haben das sogenannte dynamische Systemdenken oder Denken in Zusammenhängen – jenseits kausalen determi- nistischen Denkens – in Bereiche der Mathematik, Physik, Informatik, Kybernetik und Regelungstechnik eingearbei- tet. In Wirtschaft, Medizin, Soziologie, Organisations- und Gesellschaftswissenschaften sind systemisches Denken und Wirkungsnetz-Analysen, nicht zuletzt auch aufgrund der Globalisierung in der menschlichen Gesellschaft, schon lange Gegenstand von Forschung und Entwicklung (s. u.a. Mainzer 2008; Vester 2007; Dieckmann 2004; Briggs et.al. 1997; Dörner 1991; Perrow 1989).

Unzählige Diskussionen und Veröffentlichungen befassen sich mit dem Bachelor-Studium im Rahmen des 1999 ge- starteten Bologna-Prozesses. Das zuständige Bundesminis- terium schreibt dazu aktuell: „In Deutschland haben wir die Chance genutzt, durch die größte Hochschulreform seit Jahrzehnten die Qualität von Studienangeboten zu verbes- sern, mehr Beschäftigungsfähigkeit zu vermitteln und die Studiendauer zu verkürzen“ (www.bmbf.de/de/3336.php). Ob dies immer aus systemisch vernetzter Sicht der Dinge geschehen ist und weiter geschieht, kann ernsthaft bezwei-

felt werden, wenn u.a. der deutsche Bildungswissenschaftler Rolf Schulmeister die Bachelorstruktur bewertet als u.a.: „kleine Module, viele Prüfungen und kein Freiraum für ein Selbststudium“ (Witzmann 2010). Die anhaltende öffentliche Kritik am Bachelor-Studium mehr als 1 Jahrzehnt nach Bologna zeigt dass der Bildungswissenschaftler mit seinem Urteil nicht alleine steht.

Trotz erheblicher Fortschritte, Komplexität – wo sie uns begegnet – zu erfassen und realistisch mit ihr umzugehen, bleibt noch viel zu lernen. Die Hochschulen mit über Jahrzehnte gewachsenen Strukturen scheinen – oft selbstverschuldet – überfordert mit der neuen internationalen Ausrichtung eines Bachelor-Master-Studiums. Birger Priddat brachte es 2006 auf den Punkt, der auch für das Bildungsdenken gilt:

„Seien wir nachsichtig mit Menschen, die meinen, Komplexität lasse sich einfach „eintüten“. Sie haben es nicht anders gelernt.“

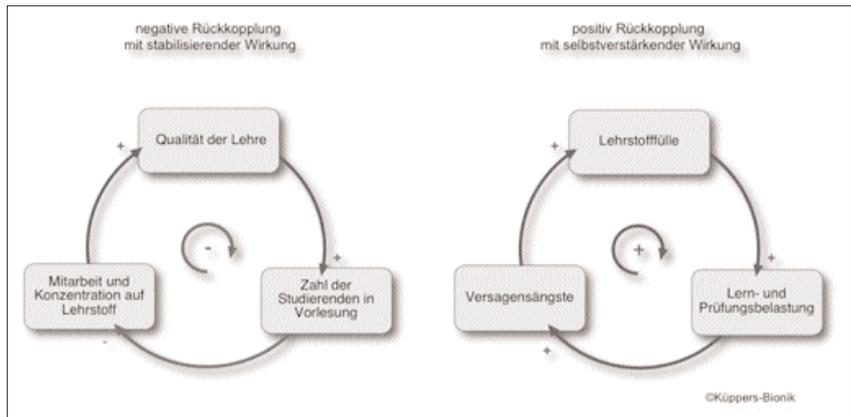
Annäherung an die Komplexität in Bildungsvorgängen

Das Wirkungsnetz – ein mächtiges Werkzeug für systemisches Bildungsdenken.

Das sogenannte Wirkungsnetz zur realistischen Abbildung einer zu lösenden Aufgabe im komplexen Umfeld ist ein mächtiges Analyse- und Lösungswerkzeug. Das beweisen zahlreiche Untersuchungen zur Lösung von Problemen hochkomplexer Struktur und Abläufe (siehe stellvertretend Vester 1990; Ulrich et. al. 1988 und zum Stichwort Bildungsreform: Pelz 2009). Das Wirkungsnetz beinhaltet sowohl die Erfassung der Ganzheit einer zu bearbeitenden Aufgabe in einem komplexen System beliebiger Art, als auch die Durchführung von fokussierten Lösungswegen unter Berücksichtigung der ganzheitlichen Vernetzung. Die Handhabung von Wirkungsnetzen basiert auf einfachen Elementen bzw. Funktionen. Sie sind das Grundgerüst - die so genannte **Hardware**. Die **Software** entspricht dem Denkmuster und der Erfahrungheit der Akteure im Umgang mit komplexen Problemlösungen. Zur Einführung sind in Abbildung 2 zwei überschaubare Wirkungsnetze mit drei bekannten Einflussgrößen und deren Beziehungen untereinander skizziert. Die (+)-Pfeile deuten eine gleichgerichtete Verstärkung oder Schwächung von Einflussgröße auf Einflussgröße hin. Andererseits weisen die (-)-Pfeile auf eine gegengerichtete Stärkung oder Schwächung von Einflussgröße auf Einflussgröße. Der Dreiviertelkreis mit dem „+“ in der Mitte besagt, dass sich die Beziehungen untereinander selbstverstärken oder selbstschwächen. Entweder führt dies zu einem „Teufelskreis“ mit ständig zunehmendem oder abnehmendem Wirkungsergebnis. Beide Richtungen führen an quantitative Grenzen und von dort direkt in einen Absturz oder Stillstand. Beides ist i.d.R. unerwünscht, tritt aber in der Realität vernetzter Bildungseinflussgrößen oder allgemein in unserer vernetzten Gesellschaft öfter ein. Fachlich wird hier von **positiver Rückkopplung** in einem System gesprochen.

Der Dreiviertelkreis mit dem „-“ in der Mitte besagt, dass die Beziehungen untereinander zwar schwanken aber letzt-

Abbildung 2: Wirkungsnetze mit positiven und negativen Rückkopplungen



lich ausgleichend regulierend wirken. Exakt diese Beziehungen sind es, nach denen in komplexen Wirkungsnetzen gesucht wird. Sie heißen **negative „systemstabilisierende“ Rückkopplungen**. Ob Wetterbedingungen, Verkehrsstaus durch Baustellen, Großtechnologien oder Bildungsstrategien: negative Rückkopplungen im komplexen Einflussgrößen-Verbund sorgen für eine stabile systemische Entwicklung, wohingegen reine positive Rückkopplungen in einem komplexen Wirkungsnetz nicht selten geradewegs zu unkontrollierbaren chaotischen Zuständen führen.

Zu lesen sind die beiden skizzierten Wirkungsnetze aus je drei Einflussgrößen mit den in der Realität oft anzutreffenden Beziehungen in einem studienbezogenen Beispiel folgendermaßen:

Linkes Wirkungsnetz – von oben im Uhrzeigersinn

Je verständlicher, interessanter, spannender der vorgetragene Lehrstoff ist, desto größere Aufmerksamkeit wird bei Studierenden geweckt. Das führt zu einer steigenden Teilnehmerzahl in der Veranstaltung, z.B. einer Vorlesung. Steigende Teilnehmerzahlen verstärken aber auch den „Pegel des Hintergrundrauschens“. Eine gewisse Unruhe kommt zwangsläufig auf, was zu kleinen Störungen während der Vorlesung führen kann. Dieser Umstand belastet wiederum die kommunikative Mitarbeit während der Vorlesung und schwächt die Konzentration interessierter Studierender auf den Lehrstoff.

Zwar wirken Mitarbeit und Konzentration auf den Lehrstoff wiederum positiv verstärkend auf die Qualität der Lehre und motiviert die Lehrperson, da sie sich in ihrem Konzept bestätigt fühlt. Aber der Grad einer konzentrierten kommunikativen Mitarbeit ist bei hoher Studentenzahl oft niedriger als bei geringer Studentenzahl. Bei einer großen Teilnehmerzahl von Studierenden wird sich der Mangel einer konzentrierten Mitarbeit für alle eher durch studentisches Desinteresse, Fächerwechsel oder im Extrem durch Studienabbruch bemerkbar machen. Schließlich sorgen Fächerhäufung und unkoordinierte Prüfungstermine zusätzlich für belastende Einflussgrößen auf ein Studium.

Das Regulativ liegt zwischen **„Zahl der Studierenden in Vorlesung“** und **„Mitarbeit und Konzentration auf Lehrstoff“**! Interessieren sich zu wenig Studierende für das Fach, so scheint die Lehrqualität zweifelhaft. Wird die Vorlesung zu einer „Massenveranstaltung“, dann ist eine kommunikative qualitative Lehre nicht mehr gewährleistet. Somit wird sich

für das gegebene Lehrfach eine optimale Grenzzahl Studierender einpendeln, bei denen sich Lehrperson und Studenten wechselseitig motivieren.

Drehen wir nun den Spieß um.

Ist die Qualität der Lehre minderwertig, tadelnswert, grundschlecht, dann interessieren sich weniger Studierende – wenn sie nicht müssen – für das Fach. Die „**Mitarbeit und Konzentration auf Lehrstoff**“ wird aber verstärkt, weil Störungen – wie bei größerer Teilnehmerzahl – i.d.R. ausbleiben. Höhere Konzentration und kommunikative Mitarbeit kann die Lehrperson bei überschaubarer Studentenzahl zu einer Stärkung seiner Lehrqualität verleiten ggf. zu neuen motivierenden „Lehr-Experimenten“ führen. Darauf können wiederum Studierende außerhalb der Vorlesung aufmerksam werden und im nächsten Vorlesungszyklus die Veranstaltung belegen, was folgerichtig zu einer Erhöhung der Teilnehmerzahl führt.

Ein derartiges Wirkungsnetz fördert Mitarbeit und Studienmotivation.

Rechtes Wirkungsnetz – von oben im Uhrzeigersinn

Erdrückt die „**Lehrstofffülle**“ durch ihre schiere Menge – und ggf. durch ihre mäßige didaktische Qualität – die Studierenden, dann wirkt sich dies unmittelbar oder mittelbar auf die „**Lern- und Prüfungsbelastung**“ der Studierenden aus. Dies verstärkt wiederum die „**Versagensängste**“ bei Prüfungen oder sogar bei der Bewältigung des Studiums. Oft führt diese Situation zu noch stärkerem und intensivem Lernen mit ggf. zusätzlichen Lernquellen. Hier arbeitet ein „**Teufelskreis**“ in Richtung physischer und psychischer Grenzen von Lernenden, mit u.U. belastenden Folgen.

Ist demgegenüber die „**Lehrstofffülle**“ gering und leicht zu bewältigen, wird die „**Lern- und Prüfungsbelastung**“ der Studierenden gering sein; ebenso die Versagensängste. Zusätzliche Hilfen in Form von mehr Lernstoffen sind nicht vonnöten. Das Studienfach oder sogar das ganze Studium scheint leicht absolvierbar – die Betonung liegt auf scheint! Auch hier können – wie im nachfolgenden Resümee angedeutet – externe Einflussgrößen die „gefühlte“ Leichtigkeit des Studiums schnell umkippen lassen.

Dieses Wirkungsnetz erschwert Mitarbeit und Studienmotivation.

Resümee zu den beiden Wirkungsnetz-Beispielen

Die Schlussfolgerungen aus beiden Beispielen von Wirkungsnetzen aus dem Bildungsbereich sind in sich logisch und konsequent. Aber: Es sind natürlich herausgestellte Rückkopplungen aus einem viel größeren Netzwerk. Beispielsweise scheint bei standardisierten „Pflichtvorlesungen“ mit regelmäßig mehreren Hundert Studierenden die „**Qualität der Lehre**“ und „**Zahl der Studierenden in Vorlesung**“ ohne externe Einflussgrößen eine allgemeine „**Mitarbeit und Konzentration auf Lehrstoff**“ kaum durchsetzbar. Alle sechs (drei + drei) hier gezeigten „externen“ Einflussgrößen können noch weiteren Einflussnahmen unterliegen und weitere Beeinflussungen auf andere ausüben. Unter anderem daher sind auch diese leicht überschaubaren Rückkopplungen nicht statisch zu werten und können gegebenenfalls tatsächlichen komplexen Realitäten adaptiv angepasst werden. Aber dies war nicht das Ziel der beiden Beispiele. Es kommt bei der Festlegung von Beziehungen

bzw. Beziehungsmustern zwischen so genannten Einflussgrößen eines Wirkungsnetzes wesentlich auf die Sicht des Beobachters und auf die **Wahrnehmung der realen Situation** an. Wirkungsnetze entstehen meistens aus einem subjektiven Empfinden für die realen Zusammenhänge heraus. Daher können Wirkungsnetze im Detail durchaus unterschiedlich sein wenn verschiedene Personen denselben Sachverhalt qualitativ bzw. quantitativ beschreiben. Die generelle Zielrichtung – und dies zeigt die langjährige Erfahrung im Umgang mit diesem Werkzeug – wird in der Regel dieselbe sein.

Vernetztes System Bachelor-Studium

Ziel

Ein Ziel oder mehrere Ziele **vorab** für eine zu lösende Aufgabe zu definieren, erst recht, wenn die Aufgabe in einem komplexen gesellschaftlichen dynamischen Umfeld angesiedelt ist, ist nicht nur hinreichende sondern unbedingte Voraussetzung für die Lösung des Problems.

So trivial es sich anhört, so oft ist es unklar: Je präziser, eindeutiger und verständlicher ein Ziel vorab definiert wird, umso unkomplizierter die Bearbeitung des dem Ziel zugrunde liegenden Wirkungsnetzes. Wird beispielsweise pauschal von Verantwortung, Energiebedarf, Kommunikationsfähigkeit oder Landesschutz als Ziel gesprochen, dann bleibt es verschwommen, unklar, diffus. Die Zusammenstellung und zielorientierte Bearbeitung eines Wirkungsnetzes läuft Gefahr, in seiner Aussage ebenso unklar, zudem noch kompliziert zu enden.

Übrigens: So perfekt die Natur Wirkungsnetze zwischen Organismen „gespannt“ hat, so wenig spezifische Ziele hat sie vorgegeben, wenn man von dem von Darwin erkannten Naturprinzip der Überlebensfähigkeit des Tauglichsten in seiner dynamischen Umwelt einmal absieht. Ein Widerspruch? Nein! Die Natur beherrscht Komplexität – wir nicht.

Nachfolgend skizzieren wir grob das Wirkungsnetz eines Bachelor-Studierenden an einer Universität. Sein Ziel ist die Erlangung eines Bachelor-Abschlusses innerhalb der Regelstudienzeit von 6 Semestern. Die strukturelle Umstellung von nationalen Diplom- zu internationalen Bachelor- und Master-Studiengängen ist aber noch im Fluss. Daher werden in der laufenden Übergangszeit – mangels ausreichender Vorbereitung – immer noch herkömmliche, kaum veränderte Inhalte eines Diplom-Studienganges von mehr als 6 Semestern in das enge Korsett eines Bachelor-Studiengangs „gepresst“. Vielfache Folgen sind zu volle Vorlesungen in ermüdender Atmosphäre, verdichtete Lehrinhalte und anderes mehr. Hinzu kommen persönliche Belastungen durch Semestergebühren, erzwungene Nebentätigkeiten aus Geldmangel und vieles mehr. Damit werden Studierende zu Beginn und im Verlauf des Studiums konfrontiert.

Elemente (Einflussgrößen)

Sie bilden punktuell den Komplex des zu untersuchenden Systems ab und bestimmen durch ihre Anzahl die Untersuchungsdichte des zu bearbeitenden Raums.

Beispiele hierfür sind aus dem Pädagogischen: Bildung, Erziehung, Migration; aus dem Wirtschaftlichen: Umsatz, Fixkosten, Zahl Fachpersonal, Steuern, Maschinenstundensatz; aus der Natur: Selbstreinigungskraft, Artenzahl, Energieumsatz; oder aus dem Politischen: Subventionshöhe, Steuer-

einkünfte, Lebensmittelkontrolle, Zahl Militäreinsätze, Sicherheit der Bürger, Freiheit der Bürger, Zahl der Grenzkontrollen u.v.m. Es kommt bei der Bestimmung der Einflussgrößen im komplexen System Bildung maßgeblich - nicht in jedem Fall - auf die Präzisierung der jeweiligen Eigenschaft bzw. Eigenheit an. Sicher können Definitionen bzw. Bewertungen von Einflussgrößen durch unterschiedliche Problemlöser zu einer differenzierten Dynamik des Wirkungsnetzes führen, die Erfahrung zeigt aber - trotz zwangsläufig unterschiedlicher Bewertungskriterien und Gewichtungen - immer wieder einen generellen Trend in Richtung des vorab festgesetzten Ziels bzw. der Ziele, wie bereits weiter oben erwähnt wurde. Die neun ausgewählten und nummerierten Einflussgrößen für unser konstruiertes Beispiel aus dem Bildungsbereich sind:

1. Lernmotivation,
2. Bachelor-Regel-Abschluss (in 6 Semestern),
3. Didaktische (VL-)Vorlesungs-Mängel komprimierte Lehrinhalte,
4. Prüfungsstress und Versagensängste,
5. Lernumgebung (marode Infrastruktur),
6. mangelhafte Lernmittelausstattung (und Platz in Bibliotheken),
7. Kommunikation zwischen Professor und Studierenden,
8. Semestergebühren (u. U. Zwang zur Kreditaufnahme),
9. „Jobs“ zur Leben (Lebenserhaltung, Nebentätigkeiten).

Natürlich können weitere Einflüsse die Abbildung der Realität im Wirkungsnetz noch detailreicher präzisieren. Fürs erste reicht aber diese Basis vollkommen, um in die Thematik des systemischen Denkens einzuführen.

Verbindungen (Interaktionen, Kausalitäten)

Sie sind die Verknüpfungen zwischen den Elementen und geben den Elementen einen netzwerkähnlichen Zusammenhalt. Verbindungen wirken zwischen Elementen und rückführend auf ein Element selbst. Verbindungen zwischen den Einflussgrößen lassen sich grundsätzlich nach vier Kriterien unterscheiden:

- I. nach Art der Wirkung (linear, nicht linear),
- II. nach Richtung der Wirkung,
- III. nach Stärke der Wirkung (in unserem Beispiel durch die Linienstärke bestimmt) und

IV. nach Geschwindigkeit der Wirkung.

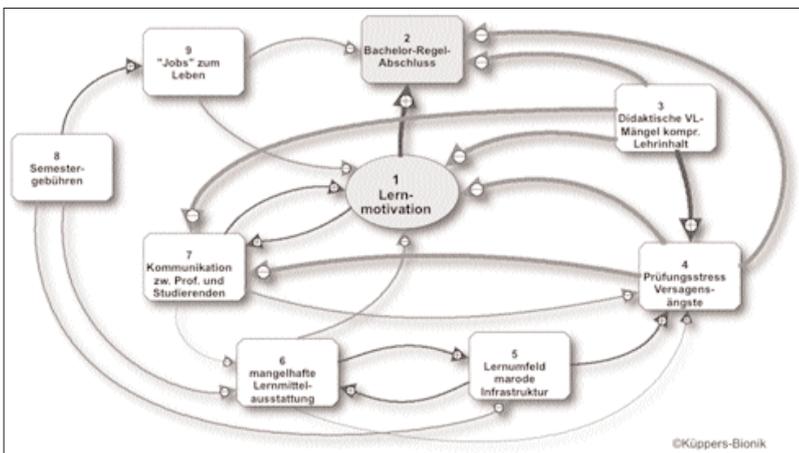
Wir betrachten hier aus der Sicht einer qualitativen Bewertung ausschließlich die Kriterien II und III. Die neun ausgewählten Einflussgrößen sind durch ein Netz aus Verbindungen miteinander verknüpft. Logische kausale Beziehungen sind durch Richtung und die Vorzeichen kenntlich gemacht. Die relativen Stärken der Beziehungen lassen sich in quantitativen empirischen Studien statistisch funktional oder in qualitativ vorgehenden Studien durch Experten-Urteile bzw. professionelle Beobachter oder Betroffenen-Urteile (aus Gruppendiskussionen oder Einzelinterviews) gewinnen. Sie sind aus drei vorgegebenen Verbindungs-Strichdicken erkennbar (zunehmende Strichdicke führt zu stärkerem Einfluss und umgekehrt).

In diesem Grundgerüst für das Wirkungsnetz sind alle 9 Einflussgrößen platziert. Ohne Frage kann sich die Netzkonstruktion durch andere Problemlöser abweichend gestalten. Aber das ist normal. Ein Wirkungsnetz folgt der Dynamik der Entwicklung und ist nie statisch.

Durch die Interaktionen nimmt das Wirkungsnetz deutliche Konturen an. Welcher Einfluss ist mit welchem verbunden und wie ist die Richtung und Wirkungsstärke. All dies und noch viel mehr tritt nun auf realistische Weise offen und sichtbar zu Tage. Im Gegensatz dazu würde eine Tabellenform mit allen 9 Einflussgrößen dem Betrachter nicht die gleiche realistische Möglichkeit bieten, Beziehungen von Einflussgrößen untereinander zu erkennen. Aber darauf kommt es bei realer Betrachtung und Lösung eines Problems im komplexen Umfeld entscheidend an!

Die Komplexität des Wirkungsnetzes in Abbildung 3 hat gegenüber der Komplexität beider Netze aus Abbildung 2 deutlich zugenommen. Und noch immer handelt es sich um eine relativ harmlose gut überschaubare Variante, was die Zahl der Einflussgrößen und Verbindungen betrifft. Für Anfänger systemischen Denkens stellt es jedoch schon eine kleine Herausforderung dar. Warum sind einige kausale Verbindungen stärker bewertet als andere, oder warum verlaufen sie in die angegebene Richtung? Welche Wirkungsunterschiede bestehen zwischen positiven und negativen Rückkopplungsschleifen? Wie ist die Dynamik ausgewählter Teile des Wirkungsnetzes oder des Gesamtsystems zu bewerten? Ein nicht enden wollender Katalog von Bewertungskriterien offenbart sich. Ohne geeignete Computerhilfe ist bereits dieses harmlos aussehende Wirkungsnetz in seiner Gesamtheit kaum in einem vernünftigen Zeitrahmen zu bewältigen.

Abbildung 3: Wirkungsnetz des Bachelor-Studiums (TATA 2004)

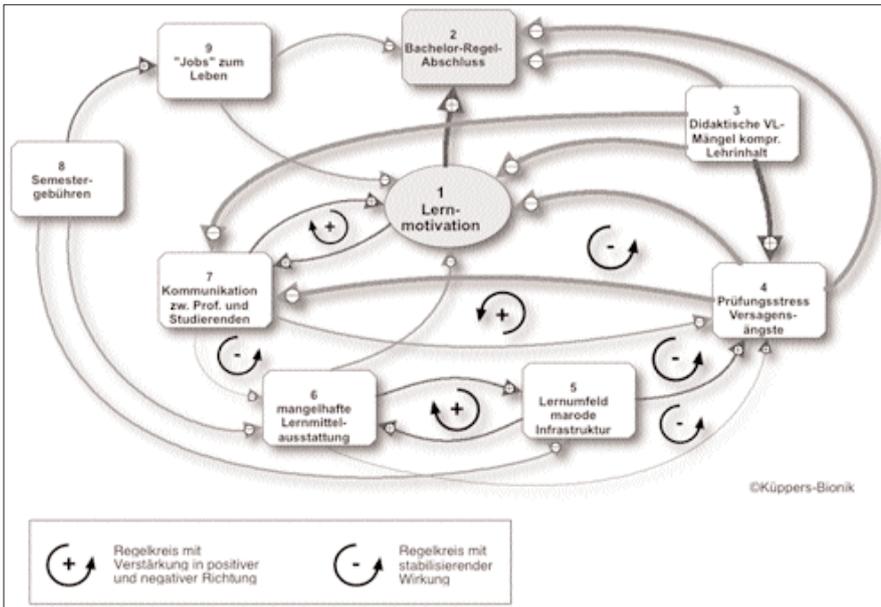


Rückkopplungen, Zirkulationen, Kreisläufe (gerichtete und gewichtete mono- und multi-kausale Verbindungen)

Sie sind spezielle Verbindungen, die zu qualitativen und quantitativen Aussagen über die Dynamik des Netzwerkes bzw. Wirkungsnetzes stark beitragen. Dabei kann der dynamische Verlauf einer initiierten Änderung linear oder nichtlineare, vorausseilende beschleunigte oder verzögerte Wirkung entfalten, die sich direkt oder indirekt auf alle vernetzten Elemente auswirkt. Aus Abbildung 2 kennen wir bereits die beiden grundlegenden Rückkopplungsschleifen und deren Wirkungen.

Im Einzelnen erkennen wir aus Abbildung 4

Abbildung 4: Herausgestellte Rückkopplungsschleifen im Wirkungsnetz des Bachelor-Studiums nach Abbildung 3



- nach Abb. 3 bzw. 4 - in Abbildung 5 näher an.

Die höchste Stufe 8 verarbeitet und bewertet **in Summe** alle **direkten und indirekten**, kaskadenartig wirksamen Verknüpfungen.

Wie lesen wir Abbildung 5?

Zum einen erkennen wir die erwarteten beiden dominierenden Einflussgrößen **Lernmotivation (1)** und **Bachelor-Regelabschluss (2)**. Sie unterliegen der stärksten Beeinflussung, wirken aber selbst gar nicht (2) oder nur sehr gering (1) auf die anderen Einflussgrößen im Wirkungsnetz. Das kritischste Element von allen ist die **Kommunikation zwischen Professoren und Studierenden (7)**, weil es sowohl auf andere stark wirkt als auch von anderen stark beeinflusst wird. Etwas schwächer im kritischen Wirkungsquadranten liegen **Prüfungstress Versagensängste (4)** und **mangelhafte Lernmittel-Ausstattung (6)**.

- insgesamt sieben Rückkopplungsschleifen.
- drei als relativ stark angenommene positive Rückkopplungen (a bis c) und
- vier superpositionierte, als relativ schwach angenommene, stabilisierende Rückkopplungen (d bis g). In Abbildung 4 sind die Einflussgrößen der Rückkopplungsschleifen durch ihre zugehörige Nummer gekennzeichnet.

Positive Rückkopplungsschleifen	Negative Rückkopplungsschleifen
a. 1 → 7 → 1	d. 1 → 7 → 4 → 1
b. 4 → 7 → 4	e. 1 → 7 → 6 → 1
c. 5 → 6 → 5	f. 1 → 7 → 6 → 4 → 1
	g. 1 → 7 → 6 → 5 → 4 → 1

hafte Lernmittel-Ausstattung (6). Stark aktiv wirkend auf andere Einflussgrößen ohne selbst beeinflusst zu werden sind **Didaktische Vorlesungsmängel und komprimierte Lehrinhalte (3)**, **Lernumfeld marode Infrastruktur (5)** und **Semestergebühren (8)**.

Die einzige Einflussgröße im konstruierten Wirkungsnetz, die zwar nicht unbedeutend ist sich aber völlig neutral verhält ist **„Jobs“ zum Leben (9)**.

Bis hierher wurde die beispielhafte systemische Analyse für den Bachelor-Regelstudiengang nach subjektiven-objektiven Kriterien ermittelt, die dem Stand der Technik entsprechen. Wir haben eine Lösung erarbeitet, die qualitativen

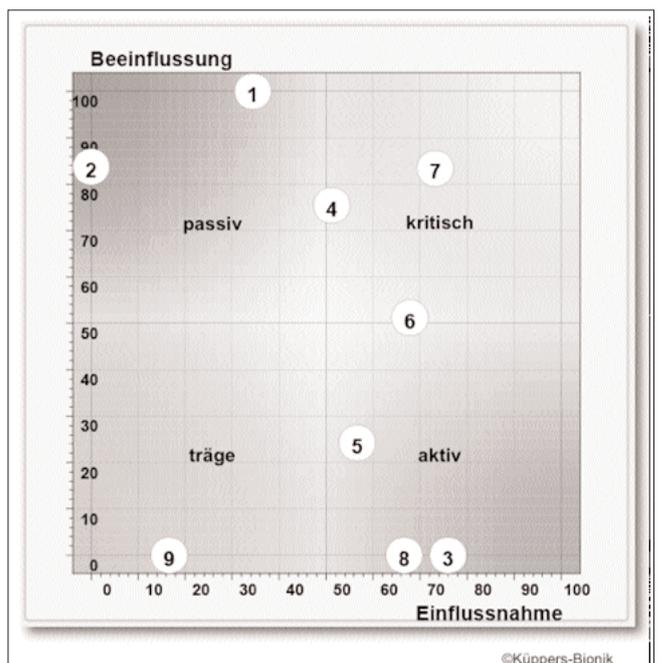
Computerunterstützte Analyse und Wertung des Wirkungsnetzes zum Bachelor-Studium

Eine programminterne Berechnung der Abhängigkeiten, Wirkungsstärken, Richtungen etc. führt schließlich zu der kumulierenden **Analyse-Matrix**, die in vier Werte-Felder (Quadranten) für die dort positionierten 9 Einflussgrößen gegliedert ist:

- Kritische Einflussgrößen im rechten oberen Quadranten der Analyse-Matrix = hohe Beeinflussung und hohe Einflussnahme
- Aktive Einflussgrößen im rechten unteren Quadranten der Analyse-Matrix = geringe Beeinflussung und hohe Einflussnahme
- Passive Einflussgrößen im linken oberen Quadranten der Analyse-Matrix = hohe Beeinflussung und geringe Einflussnahme
- Träge Einflussgrößen im linken unteren Quadranten der Analyse-Matrix = geringe Beeinflussung und geringe Einflussnahme

In diesen vier Quadranten sind die Elemente optisch so positioniert, dass der Grad der Beeinflussung (Ordinate, 0-100%) und Grad der Einflussnahme (Abszisse, 0-100%) deutlich wird (für Details siehe Lit. TATA). Schauen wir uns daraufhin die qualitative Analyse-Matrix zum Wirkungsnetz

Abbildung 5: Qualitative Analyse-Matrix zum Wirkungsnetz der höchsten kumulierten Analysestufe 8



Charakter besitzt. Allein dieser Versuch der ganzheitlichen Modellierung eines in der politischen und gesellschaftlichen Öffentlichkeit emotional und engagiert diskutierten Problems zeigt deutlich die sensiblen und weniger sensiblen „Stellschrauben“ im komplexen Bildungsumfeld. Auch wenn die hier präsentierte Lösung im Detail mit der einen oder anderen gefühlsmäßigen „gesicherten“ Ahnung oder auf kausalen isolierten Wegen ermittelten Lösung übereinstimmen mag, so legt doch nur eine systemische Wirkungsanalyse die unbedingte Grundlage für das Erkennen realistischer Zusammenhänge in komplexen Situationen. Sind auf diesem systemischen Pfad erst einmal sensible Stellschrauben im rückgekoppelten Einflussgrößen-Verbund erkannt worden, ist die Basis für die weitere Arbeit gewonnen.

Wie verändere oder beeinflusse ich die sensiblen Einflussgrößen, um das Gesamtsystem aus einer labilen oder instabilen in eine nachhaltig gefestigte und stabile Lage zu bringen? Müssen neue Einflussgrößen hinzukommen und wenn ja welche? Welche Dynamik folgt aufgrund partieller Veränderungen im Wirkungsnetz?

Diese und weitere Fragen müssen beantwortet werden. In unserem Beispiel haben sich Lehrmängel, Kommunikationsprobleme, Prüfungsstress, Versagensängste, aber auch Semestergebühren und deren Handhabung als mehr oder weniger sensible „Stellschrauben“ des Systems Bachelor-Regelstudium herauskristallisiert. Als Leitorientierung gilt: Je weniger Einflussgrößen in den beiden rechten Quadranten (kritische und aktive Elemente) der Analyse-Matrix vorhanden sind, desto größer die Wahrscheinlichkeit für eine Stabilisierung des Wirkungsnetzes. Für das Bachelor-Studium interpretiert folgt daraus: je stabiler das Wirkungsnetz desto reibungsfreier, stressärmer, motivierender und/oder weniger ängstlich können i.d.R. Studierende ihr Studienziel erreichen. Aber zudem könnten auch Kosten für Bildungsmaßnahmen zielgerichteter eingesetzt werden, neue „Einfluss-Verknüpfungen“ einbezogen werden, da man jetzt besser durch das komplexe Zusammenspiel erkennt, wo sich Schwachstellen befinden und wo nicht.

Hier beginnt die eigentliche Arbeit! Welche Varianten bieten sich nun für ein „stabiles“ Bachelor-Studiums an? Das Ergebnis von Diskussionsrunden, hauptsächlich mit Studie-

renden und Lehrenden, ergab 3 neue Einflussmerkmale mit den Nummern:

- 10. Regelmäßige Sprechstunden und Tutorien,
- 11. Kontrollinstanz zur Qualität der Lehre, insbesondere systemisches Lehren und Lernen,
- 12. Kontrolle im Umgang und Verantwortung.

Diese richtig platziert und verknüpft sollten zu einer Stabilisierung des Gesamtsystems beitragen. Soweit das realitätsnahe Modell. Die Abbildungen 6 und 7 zeigen die gefundene Lösung.

Zuerst wird die Einflussgröße **Didaktische Vorlesungsmängel und komprimierte Lehrinhalte (3)** und der von ihr beeinflusste, **stark zunehmende Prüfungsstress und Versagensängste (4)** der Studierenden einer übergeordneten „Kontrollinstanz“ zur **Qualität der Lehre, insbesondere „Systemisch Lehren und Lernen“ (11)** unterworfen. Sie soll bewirken, dass die ursprünglich starke Beeinflussung „zweifelhafter“ Lehre und Lehrinhalte auf die Studierenden abgeschwächt wird, was mit einer geringeren Stärke der Verbindung zwischen beiden Einflussgrößen dargestellt ist. Über einen zweiten Kontroll- bzw. Bewertungskanal zu **Prüfungsstress und Versagensängsten (4)** sollen sowohl das Lernverständnis – wirksam als ganzheitliches Lernen – beeinflusst als auch die Fortschritte über eine „bessere“, kommunikative und verständliche Lehre überprüft werden.

Mit dieser neuen **Kontroll-Einflussgröße (11)**, die z.B. eine paritätisch besetzte Gruppe aller am Bachelor-Studium beteiligten oder das Studium tangierenden Personen sein kann, wird auch die Erwartung geknüpft, dass die indirekten Verbindungen zu **Lernmotivation (1)**, **Bachelor-Regelabschluss (2)** und **Kommunikation zwischen Professoren und Studierenden (7)** weniger belastend wirken, was durch die geringeren Verbindungsstärken ausgedrückt wird.

Eine zweite zusätzliche Einflussgröße **Kontrolle im Umgang und Verantwortung (12)**, z.B. besetzt mit einem „Controller“-Funktion ausübenden studiennahen Gremium, prüft und handhabt vorteilhaft und nachhaltig für Studierende die Nutzung von Semestergebühren über einen stabilisierenden Regelkreis negativer Rückkopplung. Daraus könnte sich ein vorteilhafter **Selbstorganisationsprozess** entwickeln, der für

die Hochschule, Lehrende und Lernende zu einer Win-Win-Situation führen kann. Und schließlich sollen Überprüfungen der Existenz **regelmäßiger Sprechstunden und Tutorien (10)** die Kommunikation zwischen Professoren und Studierenden fördern. Alle 12 Einflussgrößen sind nun derart miteinander verknüpft, dass ein Bachelor-Studium nach diesem Modell theoretisch problemärmer absolviert werden kann, als es die gängige Praxis zeigt. Die höchste Stufe 8 verarbeitet und bewertet in Summe alle direkten und indirekten, kaskadenartig wirksamen Verknüpfungen.

Abbildung 6: Beispiel-Wirkungsnetz Bachelor-Studium Lösung

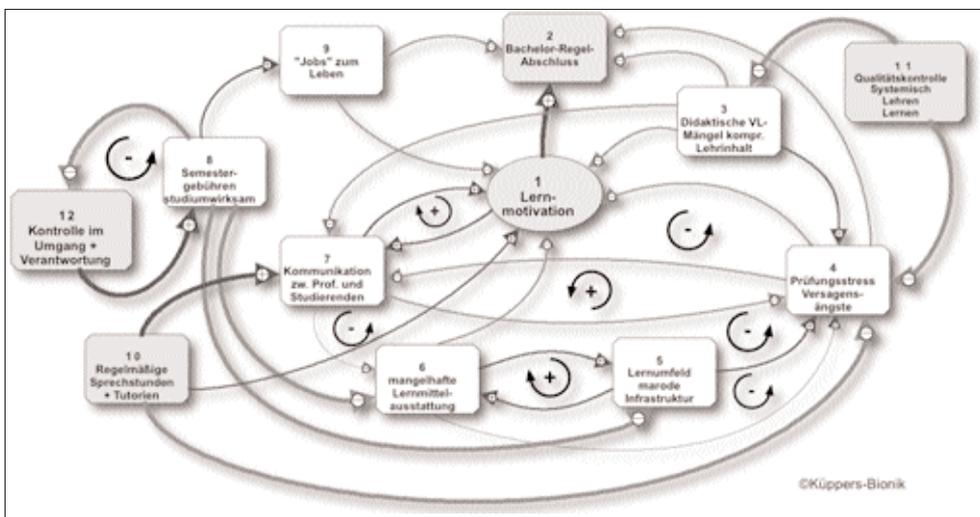
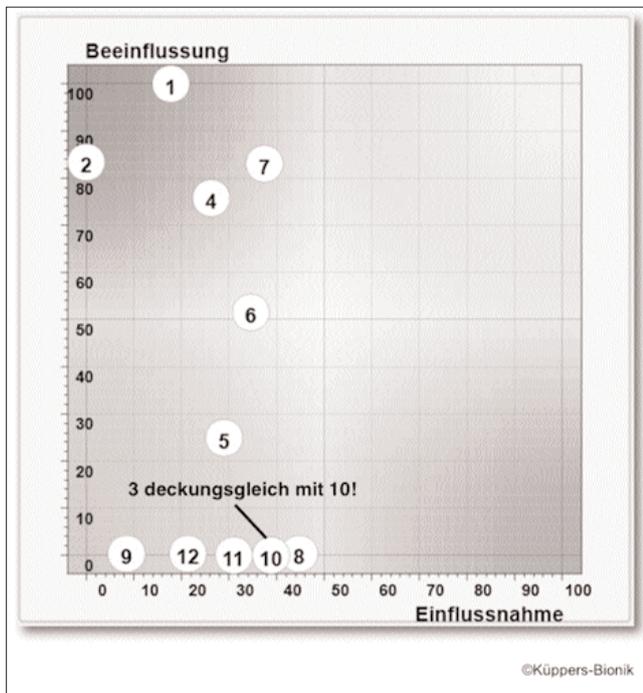


Abbildung 7: Qualitative Analyse-Matrix Lösung zum Wirkungsnetz der höchsten kumulierten Analysestufe 8



Auf den ersten Blick und im Vergleich zu Abbildung 5 ist erkennbar, dass die Einflussgrößen der Lösungsmatrix ausschließlich die beiden Quadranten für reaktive passive (oben links in Abbildung 7) und für träge puffernde Einflussgrößen (unten links in der Abbildung 7) besetzt halten. Exakt dies ist gewollt. Durch das Hinzufügen der drei „Kontroll-Instrumente“ (10, 11 und 12) konnte der Umgang mit dem Unerwarteten problemvorbeugender gestaltet werden.

Prüfungsstress und Versagensängste (4) sowie die **Kommunikation zwischen Professoren und Studierenden (7)** sind auf weniger sensible kritische Positionen gerückt, unterliegen aber nach wie vor demselben Grad der Beeinflussung. Außer den zwangsläufig dominierenden Beeinflussungsgrößen **Lernmotivation (1)** und **Bachelor-Regel-Abschluss (2)** spielen alle anderen Einflussgrößen eher neutrale träge Rollen im komplexen System.

Fazit

Wenn ein kurzes knappes Fazit aus dem detailreichen Ablauf des vorab beschriebenen Bachelor-Regel-Abschluss-Komplexes gezogen werden soll, so dieses:

Die folgende These würde unseren gegenwärtigen Trend in der Bildung auf den Punkt bringen:

Wir kennen vor allem zunehmend die Kosten und abnehmend den Wert der Bildung.

Und daher kann nach meiner Auffassung nur eine ganzheitliche Sicht der Dinge nachhaltig anpassend auf eine wirksame Lehr- und Lern-Änderung innerhalb der vorhandenen Bildungsstrukturen einwirken – und diese schließt nun einmal Kontrollen mit regelungstechnischem Charakter ein. Jedenfalls ist mit den herkömmlichen „Bastelansätzen“ eine neue erforderliche Bildung Humboldtschen Geistes in weite Ferne gerückt. Eine Einschränkung bleibt in unserem Praxisbeispiel jedoch ausgespart und kann auch trotz bestem Vorsatz nicht berücksichtigt werden: Der politische Gestaltungswille. Diejenigen Entscheidungsträger unserer Gesellschaft, die die Notwendigkeit einer grundlegenden Änderung der Sicht für nachhaltige Lösungen im komplexen vernetzten Gemeinwesen erkannt haben und es zudem in der Hand haben, diese Änderungen auch durchzusetzen, müssen es zumindest durchsetzen wollen.

Literaturverzeichnis

Begon, M.E./Harper, J.L. Townsend, C.R. (1998): Ökologie, Spektrum. Heidelberg, Berlin.
 Briggs, J./Peat, F. D. (1993): Die Entdeckung des Chaos. München.
 Dörner, D. (1990): Logik des Misslingens. Hamburg.
 Dieckmann, J. (2004): Luhmann-Lehrbuch. München.
 Lewin, R. (1996): Die Komplexitätstheorie. München.
 Mainzer, K. (2008): Komplexität. Paderborn.
 Pelz, G. R. (2009): Vernetztes Denken als Grundlage einer jeden Bildungsreform. In:Schulkritik.de, H. 1, S. 7-18.
 Perrow, Ch. (1987): Normale Katastrophen..Frankfurt, New York.
 Priddat, B. (2006): zitiert in Lotter, W.: Einfach mehr. H. 1, S. 48.
 Richter, K./Rost, J.-M. (2002): Komplexe Systeme. Frankfurt a.M.
 TATA Interactive Systems GmbH (2004): Gamma-Tutor. Tübingen.
 Ulrich, H./Probst, G. J. B. (1988): Anleitung zum ganzheitlichen Denken und Handeln. Bern, Stuttgart.
 Vester, F. (1990): Ausfahrt Zukunft. München.
 Vester, F. (2007): The Art Of Interconnected Thinking. München.
 Witzmann, E. (2010): Bachelorstudium: „Die Schwachstellen von Bologna“, Interview mit dem deutschen Bildungswissenschaftler Rolf Schulmeister „Die Presse“, 25.01.2010.

■ **Dr.-Ing. E. W. Udo Küppers**, Ingenieurwissenschaftler, Bioniker und Systemiker,
 E-Mail: kueppers@uni-bremen.de

Reihe Hochschulwesen:
 Wissenschaft und Praxis

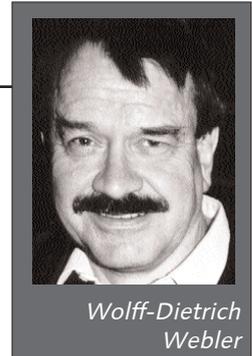
im Verlagsprogramm:

Peer Pasternack: Politik als Besuch

ISBN 3-937026-40-1, Bielefeld 2005, 253 Seiten, 29.70 Euro

Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Wolff-Dietrich Webler



Wolff-Dietrich Webler

Internationale Vergleichbarkeit von Noten im Hochschulbereich?

Problematik der Notenvergabe, Referenzgrößen und der Verwendung der Gaußschen Normalverteilung

1. Wunsch nach internationaler Vergleichbarkeit der Noten

Die Ziele der Bologna-Reformen beziehen sich nicht nur darauf, einen gemeinsamen Hochschulraum in Europa zu schaffen, sondern auch für einen transparenten, durchlässigen europäischen Arbeitsmarkt zu qualifizieren. Für Leistungstransparenz bei Bewerbungen müssten Leistungen vergleichbar festgestellt und dokumentiert werden. Das gelingt schon nicht auf nationaler Ebene und führt immer wieder zu spektakulären Konflikten (s.u.). Trotzdem wird es auch auf europäischer Ebene versucht. Die Umsetzung dieser politischen Ziele stößt jedoch auch dort auf Schwierigkeiten. Ein Mittel, dieses Ziel zu erreichen, besteht in der Verwendung des European Credit Transfer Systems (ECTS) in Form der Leistungspunkte und der ECTS-Noten. Mit beiden gibt es - unterschiedlich gravierende - Schwierigkeiten: Leistungspunkte auf der Basis eines zeitlichen Maßes an Arbeit sind noch verhältnismäßig vergleichbar. Auf die Differenzprobleme, die sich im Laufe eines Studiums akkumulieren und daraus entstehen, dass einige Länder einem Leistungspunkt 25 Stunden, andere 30 Stunden Arbeitspensum der Studierenden zu Grunde legen, soll hier nicht eingegangen werden.

Aber auf die Probleme mit der Vergleichbarkeit von ECTS-Noten muss hier eingegangen werden. Sowohl die EU-Richtlinien zur Vergabe von ECTS-Noten, als auch Empfehlungen der HRK dazu gehen bei der Festlegung von Noten im Hochschulbereich von der **Gaußschen Normalverteilung** aus (Glockenform der Ergebnisse).

Weiter wird erkennbar, dass die Notenvergabe sich am **Leistungsbild der lokalen Kohorte** orientiert, in der die Prüfungsleistung erbracht wird. Aus Sicht der Prüfungsforschung erscheint beides nicht vertretbar.

2. Orientierung an der lokalen Kohorte

Hochschulen qualifizieren für einen überregionalen Arbeitsmarkt. Vor allem Examensnoten sollen nicht nur eine Bilanz der im Studium gezeigten Leistungen gegenüber den Prüflingen ziehen, sondern in Bewerbungssituationen den potentiellen Arbeitgebern eine Information zum Leistungsstand der Bewerber geben. Diese Information muss vergleichbar sein, wenn sie ihre Funktion erfüllen soll. Die Vergleichbarkeit ist aber prinzipiell nicht gegeben, wenn sie lediglich einen von lokalen Zufälligkeiten der Lerngruppe (und Unzulänglichkeiten der Prüfungskompetenz) abhängi-

gen Leistungsspiegel enthält. Ein Arbeitgeber irgendwo in Europa (oder auch nur Deutschland) kann nicht klären, ob die Jahrgangskohorte in den Examensprüfungen in Tübingen zufällig besser zusammen gesetzt war als die in Hamburg; schon gar nicht kann er dies ausgleichend bewerten. Auf solche Weise entstandene Zeugnisnoten sind nicht vergleichbar, der Leistungsstand der konkurrierenden Bewerber also ebenfalls nicht.

Statt des lokalen Leistungsbildes der Kohorte muss ein überregionaler Maßstab zu Grunde gelegt werden, an dem sich die lokalen Leistungen (und die ihnen vorausgehenden Veranstaltungen!) zu orientieren haben. Dann spielt die jahrgangszufällige Zusammensetzung der Kohorte keine Rolle mehr. In der Berufsbildung wird von solchen Maßstäben ausgegangen. Solche Überlegungen haben einst auch zu Formen wie dem bundesweit einheitlich geregelten Juristischen Staatsexamen geführt (auf dessen Probleme hier nicht eingegangen werden kann).

Ein solcher überregionaler Maßstab kann heute das Kompetenzprofil der Absolventen sein, das ein Studiengang erwerben lassen will. Diese Profile sollen sich an wissenschaftlichen Merkmalen, aber ebenso an (relativ abstrakt gefassten und daher vielfältig einzulösenden) beruflichen Anforderungen orientieren.

Dementsprechend reichen Prüfungen auch nicht mehr aus, wenn sie Wissensmengen feststellen; moderne Prüfungen müssen das Kompetenzprofil mindestens auch als Handlungskompetenz in seiner erreichten Intensität erfassen. Prüfungsforschung und Prüfungsentwicklung haben hier Lösungen anzubieten. Die Alltagspraxis der Prüfungen in den Fachbereichen steht da allerdings noch ganz in den Anfängen.

3. Orientierung an der Gaußschen Normalverteilung?

Ein eher statistischer Einwand vorweg: Die Normalverteilung wird in den Empfehlungen der EU-Kommission und der HRK an eine Gruppengröße von mindestens 30 Studierenden gebunden. Sicherer wären mindestens 60 Studierende als Basis, denn Gruppengrößen darunter gelten nur unter bestimmten Voraussetzungen als zulässig, die keineswegs als Normalfall vorausgesetzt werden können. Das Leistungsbild einer kleinen Gruppe von 30 Studierenden auf die „Glockenform“ bringen zu wollen, erscheint auch anhand der im Alltag erfahrbaren realen Verteilung der Leistungsbilder als nicht zulässig.

Ein ebenso wichtiger, grundsätzlicher Einwand besteht darin, dass die Grundannahmen der Normalverteilung im Hochschulbereich *zumindest bei Abschlussexamina* nicht gegeben sind. Zu diesen Grundannahmen zählen, dass die Mitglieder der Lerngruppe quasi im Gleichschritt auf einen gleichen Prüfungszeitpunkt zu laufen. Das ist bei Semester-/Modulprüfungen zwar der Fall. Tatsache aber ist, dass kein Lehrender an Hochschulen einem Studierenden raten würde, sich einem Examen zu stellen, wenn dieser noch nicht erkennen lässt, dass er wohlpräpariert ein gutes Examensergebnis erwarten lässt. Also sind Prüfungsjahrgänge im Hochschulbereich, die ihren Examenszeitpunkt zumindest teilweise mit bestimmen können, viel leistungshomogener und überdurchschnittlich hoch qualifiziert. Typischerweise ist daher für Hochschulen die Annahme normalverteilter Prüfungsvorbereitung *zumindest bei Abschlussexamina* nicht zutreffend (und um die geht es bei den EU-Regelungen). Da kann es ohne weiteres vorkommen, dass eine Verteilung von 10% sehr guten und 25% guten Leistungen (also als guten Studienerfolg nur rund ein Drittel der Prüfungsgruppe umfassend!) den real gezeigten Leistungen in keiner Weise entsprechen würde. Und dann kommt es zu einem Vorgang, der jeden Testtheoretiker zu heftiger Kritik herausfordert: Weil das Examen „zu gut“ ausgefallen ist (in seiner mangelnden Professionalität also „zu leicht“ gestellt war?) wird das ganze Notenniveau für diesen Prüfungstermin als Einzelfall „im Fahrstuhl“ nach unten transponiert! Und umgekehrt. Alles das ist unvermeidbar. Außerdem besteht kein Zusammenhang zwischen der statistischen Größe „Normalverteilung“ und den inhaltlichen Prüfungsanforderungen (s.u.).

4. Hauptursachen der Unvergleichbarkeit der Noten

4.1 Nichteinhaltbarkeit testtheoretischer Forderungen bzw. Verstöße gegen diese Forderungen

Auch wenn für Prüfungen uneingeschränkt die Forderungen der Objektivität, Validität und Reliabilität gelten, so sind sie in der Alltagspraxis nur eingeschränkt oder gar nicht einzuhalten. Denn was heißt das im Prüfungsbereich?

Objektivität: Das Prüfungsergebnis ist allein durch die (vorab durch Kriterien zu definierende) Leistung des Prüflings bestimmt.

Randbedingungen:

1. Allen Prüflingen stehen die gleichen Informationsquellen zur Verfügung (z.B. Sprechstundenzeit, Beratungsgespräch).
2. Der Prüfer gibt allen Prüflingen die gleichen (formalen) Informationen (z.B. Leistungserwartungen, Standards, Kriterien).
3. Die situativen Bedingungen sind für alle Prüflinge gleich.
4. Verschiedene voneinander unabhängige Prüfer kommen bei der Bewertung der Leistung eines Prüflings zu dem gleichen oder annähernd gleichen Ergebnis.
5. Bei der Bewertung ist der Ermessens-/Interpretationsspielraum für den Prüfer eingegrenzt, d.h. die Bewertung erfolgt anhand von vorab definierten Kriterien in (teils) standardisierter Form (Schablonen, Auswertungsbögen, -vordrucke u.ä.)

Reliabilität: Die Prüfung diagnostiziert das, was sie diagnostizieren will, zuverlässig.

1. Die Beurteilung bei wiederholter Prüfung stimmt interpersonell überein.
2. Bei Wiederholung ist die Prüfung „über die Zeit stabil“.

Validität: Die Prüfung misst das, was sie messen soll. Fragen der Validität sind Fragen nach der Lehrzielorientierung und den Bewertungskriterien.

Prüfungen sind valide, wenn sie

1. nur das bewerten, was auch geprüft worden ist,
2. nur das geprüft wird, was durch die Vorgabe von Lehrzielen festgelegt worden ist,
3. die Prüfungsthemen das Prüfungs-Gebiet angemessen repräsentieren.

Aus dem vorstehenden Text geht schon bei einem kollegialen Vergleich mit der je eigenen Alltagserfahrung aus dem durchschnittlichen Prüfungsbetrieb hervor, dass solche Forderungen im Alltag nur selten eingehalten werden. Das bedeutet, dass Noten im Alltag von Variablen beeinflusst werden, die in ihrer Wirkung nicht kontrolliert sind.

4.2 Weitere Forderungen

An eine professionelle (d.h. dem Stand der Prüfungsforschung und Testentwicklung entsprechende) Gestaltung von Prüfungen und an ihre Bewertung sind weitere Forderungen zu richten:

1. Zu fordern ist die Orientierung an einem überregional gültigen Leistungsmaßstab, an dem die lokalen Leistungen jahrgangsunabhängig eingeschätzt (vielleicht im Einzelfall sogar gemessen) werden (vgl. oben Abschnitt 1). Dieser Maßstab entspricht der Tatsache, dass für einen überregionalen, mehr und mehr europaweiten Arbeitsmarkt ausgebildet wird. Dieser Leistungsmaßstab zeichnet sich über Kompetenzprofile der Hochschulabsolventen bereits ab. Hochschulen und Beschäftigungssystem stehen hierzu schon lange im Dialog.
2. Der Leistungsmaßstab muss so konkret und detailliert wie möglich formuliert sein, sodass die Interpretationsspielräume der Prüfer für die Bewertung der Leistung so weit als möglich eingeschränkt sind. Nur so sind (z.B. prüfungsfremd verursachte) Abweichungen vom Bewertungsmaßstab vermeidbar.
3. Die Leistungsmaßstäbe (sprich: die Lehrziele = Merkmale für Qualität der Ausbildung) sind hochgradig von jeweils national bzw. fachkulturell geprägten Werten abhängig. Von daher bewerten Noten sehr unterschiedliche Sachverhalte. Schon Gutachter des gleichen, lokalen Fachbereichs können aufgrund unterschiedlicher Prioritäten und Gewichtungen (z.B. Inhalt vs. Form) bei der gleichen Bachelor- oder Master-Arbeit zu sehr unterschiedlichen Noten kommen. Nur vergleichbare Maßstäbe führen auch zu vergleichbaren Noten. Eine solche Vergleichbarkeit ist auch innerhalb gleicher Fächer zwischen Fachbereichen noch sehr unterschiedlich erreicht, besonders weit auseinander liegend bei interpretativen Fächern.
4. Die Noten spiegeln außerdem Leistungen auf unterschiedlichen Ebenen kognitiver Denkopoperationen wider. Mangels Ausbildung sind den meisten Prüfern diese Unterschiede nicht bekannt, damit aber auch das Problem

nicht bewusst. Eine Klausur, die weitgehend auf der Ebene der Wissensabfrage (Ebene 1) stattgefunden und eine Note festgesetzt hat, ist im Ausmaß und der Qualität festgestellter Fähigkeiten nicht vergleichbar mit einer Klausur, die sich weitgehend auf analytischen, bewertenden/beurteilenden oder auf kreativen, entwickelnden, problemlösenden Ebenen bewegt hat (also den Ebenen 4,5 und 6 auf der Sechser-Skala). Die so ermittelten Noten beziehen sich daher auf völlig unterschiedliche Leistungsbilder. Gegenüber einem potentiellen Arbeitgeber haben sie keine Aussagekraft über die nachgewiesenen Fähigkeiten.

5. Die nächste Quelle der Unvergleichbarkeit von Noten ist deren fehlerhaftes Zustandekommen. Bei schriftlich und mündlich gestellten Fragen können 21 Standardfehler der Frageformulierung auftreten (s. Anhang), bei mündlichen Prüfungen noch einmal 24 Standardfehler bei Bewertungen. So hochgradig fehleranfällig zu Stande gekommene Noten sind nicht geeignet, als Basis eines Leistungsvergleichs zu dienen.
6. Eine weitere Dimension der Unterscheidung von festgestellten Fähigkeiten bilden die Arten des Wissens, die dabei geprüft wurden: Faktenwissen, konzeptionelles Wissen, prozedurales Wissen und metakognitives Wissen (Anderson/Krathwohl 2001). Darunter ist zu verstehen:

“Knowledge Dimensions Defined

Factual Knowledge is knowledge that is basic to specific disciplines. This dimension refers to essential facts, terminology, details or elements students must know or be familiar with in order to understand a discipline or solve a problem in it. Conceptual Knowledge is knowledge of classifications, principles, generalizations, theories, models, or structures pertinent to a particular disciplinary area.

Procedural Knowledge refers to information or knowledge that helps students to do something specific to a discipline, subject, area of study. It also refers to methods of inquiry, very specific or finite skills, algorithms, techniques, and particular methodologies.

Metacognitive Knowledge is the awareness of one’s own cognition and particular cognitive processes. It is strategic or reflective knowledge about how to go about solving problems, cognitive tasks, to include contextual and conditional knowledge and knowledge of self” (Wilson 2006).

Noten allein geben einem potentiellen Arbeitgeber auch hier keine ausreichende Information über die tatsächlich vorhandenen Fähigkeiten eines Bewerbers. Sie sind daher isoliert betrachtet - ungeeignet.

5. Resümee und Lösungsversuch

Ein Teil dieser Prüfungspraxis kann professionalisiert und damit weniger fehleranfällig gestaltet werden. Aber nur ein Teil. Angesichts der schweren Mängel, die zu konstatieren sind, scheinen faire Vergleiche noch längere Zeit nicht möglich. Dies sind aber nicht die einzigen Probleme, die die Notenfindung zu bewältigen hat. Wie schon erwähnt, sind die Kriterien der Notenvergabe hochgradig abhängig von den Wertesystemen - Studienzielen - in den

jeweiligen Wissenschaftskulturen. Und die werden sich nicht leicht vereinheitlichen lassen - was auch nicht erstrebenswert wäre.

Der Traum von europaweit vergleichbaren, objektiven Noten als Ergebnis eines Leistungstests bleibt schon deshalb unerfüllt, weil die verschiedenen fachlichen und nationalen/regionalen Wissenschaftskulturen unterschiedliche Wertekataloge hervorbringen, nach denen gewertet wird. Daraus resultieren deutlich unterschiedene Profile für Studienziele (oder billigend in Kauf genommene Folgen dieser Kulturen), die in einer ersten Annäherung größere Gemeinsamkeiten in einer mittel-/osteuropäischen Wissenschaftskultur, einer romanisch-mediterranen und einer eher west- und nordeuropäischen Wissenschaftskultur mit je eigenen Regeln erkennen lässt. In Mittel-, West- und Nordeuropa wird besonderer Wert auf Selbständigkeit der Studierenden bzw. Absolventen, Methodensicherheit, Handlungsfähigkeit usw. gelegt. Hierarchien und soziale Distanzen werden vergleichsweise geringer gelebt. In den romanischen Ländern und teilweise den Ländern des ehemaligen Ostblocks wird diese Distanz stärker betont - verbunden mit der dort nach wie vor bevorzugten Form der Vorlesung, der frontalen Instruktion, wird wesentlich mehr Wert auf Wissen und (bestenfalls) Verstehen gelegt und deutlich weniger auf den Erwerb eigener Urteilsfähigkeit, Eigenständigkeit von Sichtweisen und Standpunkten sowie die Entwicklung eigener Lösungen.

Solche Unterschiede sind nicht dramatisch - sie sollten nur durch Reflexion und Explikation und systematische internationale Vergleiche transparent gemacht werden. Schon innerhalb Deutschlands ist einem Arbeitgeber, der öfter Hochschulabsolventen bestimmter Fachrichtungen einstellt, bekannt welche Kompetenzprofile bestimmte Fachbereiche bei ihren Absolventen ausprägen. Dass mit Noten nicht wie mit einer Einheitswährung „Euro“ umgegangen werden kann, ist dann kein Nachteil mehr. Individuelle Differenzierungen der Bewerber unterhalb allgemeiner Differenzen der Wissenschaftskulturen lassen sich dann über andere Bewerbungsunterlagen erkennen.

Literaturverzeichnis

- Anderson, L.W./Krathwohl, D.R. (eds.)(2001): A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives. New York: Longman
 Wilson, L.O. (2006): <http://www.uwsp.edu/education/lwilson/curric/newtaxonomy.htm> (28.03.2010)

Anhang

Fehlerquellen der Fragetechnik in Prüfungen (insbesondere mündlichen)

1. Keine Doppelfragen stellen! Sie verwirren den Prüfling und lenken ab (Welche der beiden Fragen ist tatsächlich gemeint? Sind sie voneinander abhängig, bauen sie aufeinander auf? In welcher Reihenfolge soll ich die Fragen beantworten?)
2. Keine „Romanaufträge“ erteilen! („Erzählen Sie mir etwas über ...“ Beliebte Eröffnungsfrage, um dem Prüfling freie Hand zu lassen, womit er/sie beginnen will; stark verwirrend, da Prüflinge die Wahlfreiheit nicht be-

- greifen, sich statt dessen fragen, was der Prüfer hören will)
3. Keine unpräzisen Fragen stellen! („Was macht der Apotheker am Monatsende?“)
 4. Entscheidungsfragen nach Möglichkeit vermeiden! („Kann das So-und-So-Präparat auch als so und so bezeichnet werden?“) Fehlerwahrscheinlichkeit nur 50% - nicht zulässig
 5. Negationsfragen sollten vermieden werden! (Viel zu vage; der Prüfling weiß nicht, worauf der Prüfer hinaus will: „Was darf der Apotheker nicht tun?“ Übertriebenes Beispiel, um die Vagheit ganz deutlich zu machen: „An welchem Fluss liegt Hamburg nicht?“)
 6. Inverse Frageformen vermeiden! („Die Einsatzmöglichkeiten von Aspirin bestehen worin?“) Schlechtes Deutsch; beginnt als Mitteilung - die eigentliche Prüfungsfrage wird erst anschließend erwartet; erst mit dem letzten Wort dann „Überfall“ als Frage
 7. Keine Ratefragen! (Nachgeschobene Prüferfrage: „Worauf will ich wohl hinaus?“ oder: „Was meine ich wohl damit? Können sie sich unter meiner Aussage etwas vorstellen?“)
 8. Als Prüfer keine „Romane“ erzählen! (Stiehlt den Kandidat/innen die Prüfungszeit zur Darstellung ihres Könnens)
 9. Denk- oder Wissensfragen stellen? (Mehr als das; würde nur die ersten beiden Ebenen der Denkopoperationen einbeziehen; bis zu den höchsten Denkkategorien vordringen! Ebenen 4,5 & 6)
 10. Keine „Gesinnungsfragen“ stellen, aber z.B. Pro- und Contra-Argumente eines Problems erfragen. (Nicht: „Was halten Sie davon, dass der Betriebsrat mehr Rechte bekommen sollte?“ Dagegen: „Welche Vor- und Nachteile hätte es, wenn der Betriebsrat mehr bzw. weniger Rechte bekäme?“)
 11. Den Schwierigkeitsgrad der Fragen beachten! Ebenen nicht verwechseln. (Nennen Sie mir,... (Ebene 1); Erklären Sie mir ... (Ebene 2)
 12. Die Fragen deutlich und laut aussprechen! (Der Prüfling wagt nur 1-2 mal zurück zu fragen, wenn er etwas tatsächlich akustisch nicht verstehen konnte, weil die Frage zu leise oder undeutlich war)
 13. Soll man Hilfsfragen stellen? (Ein weiterer Hinweis ist sinnvoll, wenn er nicht zu Punktabzügen führt!)
 14. Wie soll man sich bei ausweichenden Antworten verhalten? (Auf der gestellte Fragen bestehen!)
 15. Keine Scheinprüfungsfragen stellen! („An welchem Fluss liegt Frankfurt a.M.?“)
 16. Nicht auf Einzelbegriffen herumreiten! Insbesondere nicht, wenn der Kandidat den Sachverhalt sichtlich verstanden hat und erklären kann - nur auf den Kernbegriff jetzt nicht kommt. („Wat is een Dampfmashin'? - Vortrag aus der „Feuerzangenbowle“. Oder: „Was ist das Wesen der Arzneimittel?“)
 17. Keine ironischen oder gar sarkastischen Fragen stellen! („Lassen wir mal jegliche Logik beiseite und betrachten nur Ihre Antwort“... - „Tun wir mal so, als hätten Sie Ahnung von ...“)
 18. Keinen Druck erhöhen von der Art: „Na ja, eine weitere Chance haben Sie ja noch mit der letzten Frage!“
 19. Soll der Prüfer richtige Antworten bestätigen? (Mit zurückhaltender Bekräftigung - ja!)
 20. Soll der Prüfer falsche Antworten korrigieren? (Zurückhaltend, ohne Verniedlichung - ja! „Ihre Antwort ist so noch nicht korrekt“ Nicht: „Da liegen sie mit ihrer Antwort jetzt nur ein klitzekleines bisschen daneben!“)
 21. Nach welchen Gesichtspunkten soll der Schwierigkeitsgrad der Fragen gewählt werden? (Ansetzend am mutmaßlichen Leistungsvermögen der Kandidat/innen, immer vom Leichterem zum Schwierigeren, damit der Kandidat Zuversicht und Sicherheit gewinnen kann. Niemals auf eine bestimmte Note hin prüfen, auch nicht auf die nächstbessere als das evtl. vorangegangene schriftliche Prüfungsergebnis. Der Prüfling muss die Chance haben, auch auf höchste Noten zu kommen. Vielleicht hat er/sie Lücken in den reinen Fakten-/Wissensbeständen, aber seine Stärken auf den höchsten Bewertungsebenen der Analyse, Urteilsfähigkeit und kreativen Entwicklungsfähigkeit bei Problemlösungen. Also müssen die Fragen Antworten auf dem gesamten Spektrum der Anforderungen ermöglichen - auch in ausreichendem Umfang für eine hohe Note).

■ **Dr. Wolff-Dietrich Webler**, Professor of Higher Education, University of Bergen (Norway); Ehrenprofessor der Staatlichen Pädagogischen Universität Jaroslawl/Wolga; Leiter des Instituts für Wissenschafts- und Bildungsforschung Bielefeld (IWBB), E-Mail: webler@iwbb.de

Jonas Kunze & Andreas Geyer-Schulz

ECTS-Notenzuweisung: Pragmatische Kohortenbildung



Jonas Kunze



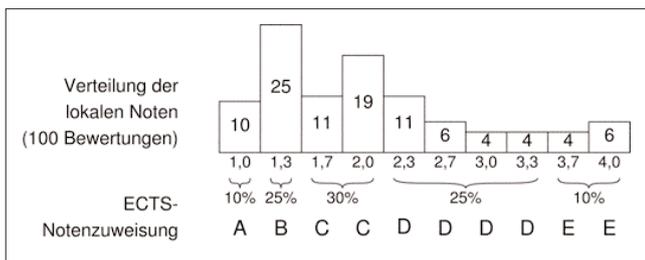
Andreas Geyer-Schulz

Die Berechnung und Ausweisung von ECTS-Noten ist an vielen Institutionen ein nicht gelöstes Problem, das auch im aktuellen ECTS Guide (EU-Kommission 2009) kommentiert wird. Auf europäischer und nationaler Ebene finden sich Empfehlungen und Vorgaben für die ECTS-Notenberechnung. Dieser Beitrag erläutert anhand von Beispielen, welche Probleme bei der Verwendung dieser Vorgaben zu erwarten sind und weist auf die gravierenden Folgen für die verantwortliche Institution hin. Um eine praktikable Lösung zu erreichen, wird eine grundlegende Forderung (Mindestgröße der Referenzgruppe) hinterfragt. Abschließend wird ein Verfahren erläutert, welches die genannten Probleme vermeidet.

ECTS-Noten stellen eine Ergänzung zu den lokalen Noten dar, die für Leistungsbewertungen auf allen Ebenen, d.h. Abschluss-, Fach-, Modul- und ggf. Teilmodulebene ausgewiesen werden sollten. Die ECTS-Noten treffen eine grobe Aussage darüber, wie die Leistung eines Studierenden im Bezug auf die üblicherweise vergebenen Bewertungen einzuordnen ist.¹ Nach den EU-Vorgaben werden Bewertungen, die von den 10% besten Studierenden erreicht werden, die ECTS-Note A zugewiesen. Die darauffolgenden 25% erhalten die ECTS-Note B. Für die folgenden 30%, 25% und die letzten 10% werden respektive die ECTS-Noten C, D und E vergeben. Abbildung 1 veranschaulicht beispielhaft die Zuordnung.

Das Beispiel legt nahe, dass die ECTS-Noten direkt aus der Notenverteilung einer Prüfung bestimmt und den Prüfungsteilnehmern neben ihren lokalen Noten (z.B. 1,3) gleichzeitig ECTS-Noten zugewiesen werden können. Jedoch stellen Nachklausuren, mündliche Nachprüfungen und Anerkennungen die Frage, welche Notenverteilung in diesem Falle herangezogen werden soll. Gleiches gilt für die Be-

Abbildung 1: ECTS Notenzuweisung (Beispiel)



stimmung von ECTS-Noten für ein Fach, das sich aus verschiedenen Modulen zusammensetzt.

Ein intuitiver, naheliegender Ansatz ist die Verwendung einer Referenzgruppe (EU-Kommission 2009, S. 41). Im operativen Bereich wird häufig der Begriff Kohorte verwendet. Zur Bestimmung der Referenzgruppe (bzw. Kohortenbildung) finden sich Vorgaben bzw. Empfehlungen auf EU-Ebene:

- Zusammenfassen der Benotungen vorheriger Jahrgänge (EU-Kommission 2009, S. 41; HRK 2005)
- Zusammenfassen der Benotungen mehrerer verwandter Studienprogramme (EU-Kommission 2009, S.41)
- Als Mindestgröße werden 30 Benotungen empfohlen (EU-Kommission 2005, S.14)

Auf nationaler Ebene existieren analoge Empfehlungen der HRK (HRK 2005). Die Umsetzung einer aussagekräftigen ECTS-Notenausweisung kommt jedoch offensichtlich nur zögerlich voran, wenn im aktuellen ECTS-Guide (EU-Kommission 2009, S. 41) bemerkt wird: "[...], the use

¹ Die ECTS-Noten ergänzen somit die in Deutschland verwendete absolute Leistungsbewertung (Rating-Verfahren) um eine relative Bewertung (Ranking-Verfahren). Diese beiden Arten der Leistungsbewertung miteinander zu verbinden, wirft grundlegende Fragen auf. In Webler 2010 wird beispielsweise erläutert, dass aufgrund der Verwendung einer lokalen Kohorte und der fehlerhaften Annahme von normalverteilten Bewertungen eine internationale Vergleichbarkeit nicht erreicht werden kann.

Dieser grundlegenden Kritik ist jedoch eine pragmatische Sichtweise zur Seite zu stellen. Neben dem Ziel der internationalen Vergleichbarkeit von Noten kann die zusätzliche Information, die mit aussagekräftigen ECTS-Noten verbunden ist, als wertvolle Verbesserung zum Status Quo und damit als Ziel an sich angesehen werden. Unter der Verwendung einer lokalen Kohorte liefern die ECTS-Noten Information über die relative Leistungsfähigkeit eines Studierenden. Somit wären ECTS-Noten (wie auch die lokalen Noten) nicht direkt, sondern immer nur institutionsbezogen vergleichbar. Ein Vergleich der Institutionen wird hierdurch mittelfristig verbessert, indem beispielsweise im Rahmen von Austauschprogrammen die ECTS-Noten von Studierenden an ihrer lokalen Universität und ihrer Gastuniversität verglichen werden und daraus eine grobe Einschätzung des Leistungsniveaus der beiden Institutionen abgeleitet werden kann. Aufgrund von Sprachbarrieren und fehlerhaften Einstufungen von Gaststudierenden bleibt eine exakte Einschätzung sehr problematisch. Werden ECTS-Noten unter dieser Zielsetzung gesehen, stellt die Verwendung einer lokalen Kohorte kein Problem an sich mehr dar.

Bezüglich der Normalverteilungsannahme ist zu bemerken, dass ein Ranking-Verfahren grundsätzlich eine eindeutige Rangordnung der Leistungen voraussetzt. Ob die (absoluten) Leistungen normalverteilt sind, ist hierbei unerheblich. Problematisch wird es hingegen, wenn das Bilden einer eindeutigen Rangordnung nicht möglich ist, beispielsweise wenn die bei einer Erfolgskontrolle die Teilnehmer nur als bestanden bzw. als nicht bestanden bewertet werden. Auf diese Problematik soll in diesem Beitrag jedoch nicht eingegangen werden.

Abbildung 2: Entwicklung der Referenzgruppen (Wandernde Kohorte)



of the ECTS scale by European institutions has been rather limited."

Im Folgenden betrachten wir das Verfahren der wandernden Kohorte, welches eine direkte Umsetzung der genannten Vorgaben darstellt. Anhand von Beispielen werden Probleme auf operativer Ebene beschrieben. Abschließend wird die Ursache auf konzeptioneller Ebene identifiziert und eine Lösung entwickelt.

1. Probleme der wandernden Kohorte

Eine Voraussetzung zur Bildung der Referenzgruppe ist die Verfügbarkeit von Prüfungsergebnissen aus der Vergangenheit. Im Folgenden wird davon ausgegangen, dass diese Prüfungsergebnisse vorliegen.

Um das Verfahren der wandernden Kohorte zu beschreiben, betrachten wir nun ein Modul, welches in der Vergangenheit von jeweils 20 Studierenden pro Jahr absolviert wurde. Das Verfahren der wandernden Kohorte bestimmt die Referenzgruppe als die Absolventen des Moduls der vorigen drei Jahre. Die Größe der Referenzgruppe hat somit eine ungefähre Größe von 60 Absolventen. Am Anfang eines jeden Jahres wird die Referenzgruppe aktualisiert: Der älteste Jahrgang wird entfernt und ein neuer Jahrgang wird in die Referenzgruppe aufgenommen. Das Verfahren ist in Abbildung 2 dargestellt.

Werden nun, bei gleichbleibender Absolventenzahl, die Prüfungsergebnisse mit der Zeit besser respektive schlechter, passt sich die Kohorte in den nächsten Jahren auf diesen neuen Zustand an.

1.1 Beispiel 1: Ausreißer

Als Erstes wird diskutiert, wie ein einmaliger Ausreißerjahrgang auf die Entwicklung der ECTS-Notenvergabe wirkt. Ausgehend von den Absolventenzahlen entsprechend Abbildung 2 nehmen wir an, dass im Jahre 2006 markant bessere Prüfungsergebnisse erreicht wurden. Die Referenzgruppe für die Jahre 2007 bis 2009 beinhaltet die Absolventen aus 2006 mit den sehr guten Prüfungsergebnissen und weist somit tendenziell schlechtere ECTS-Noten aus. Dies ist soweit die direkte und logische Folge der guten Prüfungsergebnisse des Jahres 2006. Im Jahre 2010 entfällt der Jahrgang 2006 aus der Referenzgruppe und es werden

wieder bessere ECTS-Noten zugewiesen. Diese Änderung in der ECTS-Notenzuweisung beim Übergang von 2009 zu 2010 wird von den Absolventen als nicht intuitiv wahrgenommen, da die Prüfungsergebnisse von 2007 bis 2009 durchgehend miteinander vergleichbar sind. Diese nicht intuitive Vergabe von ECTS-Noten kann zu rechtlichen Einsprüchen gegen die Notenzuweisung führen, welche wiederum mit einem hohen Arbeitsaufwand für die Institution verbunden ist.

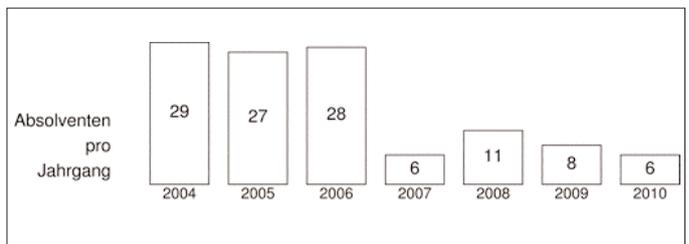
Das Problem der abrupten Änderung ohne erkennbaren Grund in der ECTS-Notenzuweisung wird weiter verschärft, wenn sich die Studierendenzahlen ändern. Nehmen wir ergänzend zum vorigen Beispiel an, dass im Jahre 2006 nicht nur sehr gute, sondern auch noch sehr viele Prüfungsergebnisse, beispielsweise 83, erreicht wurden. Die Absolventen dieses Jahrgangs stellen nun nicht mehr, wie bisher, ungefähr ein Drittel der Referenzgruppe, sondern machen über zwei Drittel der Referenzgruppe aus. Somit ist die Veränderung bei der ECTS-Notenzuweisung sowohl beim Übergang von 2006 auf 2007, als auch beim problematischen Übergang von 2009 auf 2010, deutlich abrupter als im vorigen Beispiel.

Zusammenfassend führen Ausreißerjahrgänge mit extrem guten bzw. schlechten Bewertungen zu unerwartetem Verhalten bei der ECTS-Notenzuweisung in der Zukunft.

1.2 Beispiel 2: Zu kleine Referenzgruppe

Als Zweites wird der Fall von zu kleinen Referenzgruppen illustriert. Betrachten wir die in Abbildung 3 dargestellte Entwicklung der Absolventenzahlen. Während es bis 2006 vergleichsweise viele Absolventen gab, reduzierte sich ab 2007 die Absolventenzahl auf ungefähr ein Drittel des ursprünglichen Wertes. Als Grund für eine solche Entwicklung mag ein Dozentenwechsel in einem Wahlbereich eines Studiengangs dienen. Auch sind ähnliche Effekte beim Übergang der deutschen Schulen vom G9- auf das G8-System in den nächsten Jahren zu erwarten.

Abbildung 3: Sinkende Absolventenzahlen



Im genannten Beispiel entsteht das Problem einer zu kleinen Referenzgruppe: Die Referenzgruppe für 2010 besteht nur aus 25 Absolventen, was die empfohlene Mindestgröße von 30 Absolventen unterschreitet. Die empfohlenen Lösungsansätze lauten:

- Die Noten mehrerer Module zusammenfassen (EU-Kommission 2005, S. 15).
- Mehrere vorherige Jahrgänge in die Referenzgruppe einbeziehen (EU-Kommission 2005, S. 15).
- Pragmatische Lösung wählen (HRK 2005).

Die ersten beiden Lösungsansätze führen zu einer strukturellen Änderung der ECTS-Notenzuweisung. Im gegebenen Beispiel bedeutete dies, dass die Bestimmung der Referenzgruppe bis 2009 nach einem anderen Schema erfolgte als ab 2010. Diese Änderung bedarf der detaillierten Dokumentation, um bei juristischen Auseinandersetzungen die Bestimmung der Referenzgruppe nachvollziehbar zu halten. Ein weiteres und folgenschweres Problem ist jedoch das Fehlen von objektiven Kriterien für die gewählten Änderungen. Wie viele Vorgängerjahre werden zusammengefasst? Welche Module werden zusammengefasst? Wann wird die Zusammenfassung aufgehoben? Das Fehlen von objektiven Kriterien, im Besonderen bei der von der HRK angedachten pragmatischen Lösung, führt zur Ungleichbehandlung von Absolventen, sodass der Einspruch gegen die ausgewiesenen ECTS-Noten mit hoher Wahrscheinlichkeit vor Gericht Erfolg haben wird. Die Aussagefähigkeit der ECTS-Note wäre damit nicht mehr gegeben. Abschließend ist von einer strukturellen Änderung des Verfahrens noch aus einem dritten Grunde abzuraten: Erfolgt die Bildung der Referenzgruppe feingranular auf Fach-, Modul- und Teilmodulebene ist bei einem Studiengang mit einem großen Wahlangebot leicht von hunderten von Prüfungen auszugehen, bei der für jede eine Referenzgruppe gebildet werden muss. Eine strukturelle Änderung des Verfahrens würde pro Prüfung bzw. Referenzgruppe manuell erfolgen, was mit einem sehr hohen Personalaufwand verbunden wäre.

1.3 Beispiel 3: Kaltstart

In den letzten Jahren wurden fast alle Studiengänge in Deutschland auf das Bachelor- und Master-System umgestellt. Während in den vorigen Beispielen von einem laufenden Prüfungsbetrieb ausgegangen werden konnte, werden in den neu eingerichteten Studiengängen die Prüfungen erstmalig absolviert. Die offene Frage ist, wie die Referenzgruppe in diesem Fall gebildet werden soll. Diese Problematik wird im Folgenden auch als Kaltstartproblem bezeichnet.

Ein grundsätzlich praktikabler Ansatz ist es, Absolventen eines verwandten Studiengangs heranzuziehen, z.B. des vormals angebotenen Diplomstudiengangs. Es können sich jedoch Probleme wie in Beispiel 1 und 2 ergeben, wenn die Absolventenzahl im neuen Studiengang nicht der des verwandten Studiengangs ähnelt. Auch wenn die Absolventenzahlen vergleichbar sind, stellen Wahlmöglichkeiten ein weiteres Problem dar, welches im Folgenden illustriert wird.

Ein neuer Bachelorstudiengang verfügt über ein Fach „Vertiefung“, in dem eines von drei angebotenen Modulen absolviert werden muss. Wie groß soll die Referenzgruppe für jedes der drei Module sein, wenn ungefähr 100 Absolventen pro Jahr erwartet werden? Eine gleichmäßige Verteilung der erwarteten Absolventen auf die drei Module ist zwar pragmatisch, entspricht jedoch meist nicht dem tatsächlichen Wahlverhalten der Studierenden, sodass Effekte wie in Beispiel 1 und 2 mit hoher Wahrscheinlichkeit zu erwarten sind. Noch komplizierter wird die Referenzgruppenbestimmung, wenn in dem Fach „Vertiefung“ weitere Module angeboten werden oder ein Modul wegfällt. Die erwartete Anzahl an Absolventen wird sicherlich sinken respektive steigen, sodass eine nachträgliche Anpassung der Referenz-

gruppe angeraten ist. Dies wiederum führte jedoch zu einer strukturellen Änderung der ECTS-Notenzuweisung mit den bereits erläuterten negativen Konsequenzen.

Zusammenfassend lauten die zentralen Probleme:

- Das Ausscheiden von Absolventengruppen eines Jahrgangs aus der Referenzgruppe kann zu unerwarteten Änderungen bei der ECTS-Notenzuweisung führen.
- Änderung von Größen der Absolventengruppen erfordern bzw. erzwingen strukturelle Änderungen bei der Bestimmung der Referenzgruppe. Dies ist jedoch mit äußerst negativen organisatorischen, juristischen und schlußendlich politischen Konsequenzen verbunden.
- Das Kaltstartproblem kann nicht zufriedenstellend gelöst werden.

Positiv formuliert lassen sich verschiedene Anforderungen an ein gewünschtes Verfahren zur ECTS-Notenzuweisung identifizieren. Nur wenn diese Eigenschaften erfüllt sind, kann die Ausweisung von aussagekräftigen ECTS-Noten unter vertretbarer Belastung der organisatorischen und juristischen Abteilungen erfolgen.

- Unerwartete Änderungen der ECTS-Notenzuweisungen sollen vermieden werden. Eine Anpassung der ECTS-Notenzuweisung soll kontinuierlich verlaufen.
- Das Verfahren zur ECTS-Notenzuweisung soll (a) nach objektiven Kriterien erfolgen, (b) keine manuellen Anpassungen benötigen und (c) das Problem zu kleiner Referenzgruppen lösen.
- Das Kaltstartproblem soll auf einfache und transparente Weise gelöst werden können.

2. Konzeptionelle Ursache und Lösung

Fasst man die Beobachtungen aus dem vorigen Abschnitt zusammen, lässt sich ein grundlegendes, konzeptionelles Problem identifizieren. **Das konzeptionelle Problem liegt in der Tatsache, dass die Referenzgruppe aus Absolventen besteht.** Obwohl diese Aussage offensichtlich und intuitiv ist, ergeben sich hieraus die genannten folgenschweren Probleme.

Erstens ist ein Absolvent entweder in der Referenzgruppe oder nicht, d.h. neben dem Eintritt in die Referenzgruppe kann er auch austreten, was zu den in Beispiel 1 geschilderten Problemen führt.

Zweitens geht jeder Absolvent mit einem Gewicht von Eins in die ECTS-Notenzuweisung ein. Große Jahrgänge erhalten somit ein stärkeres Gewicht.

Drittens hat die Referenzgruppe eine exakte Größe, identisch zu der Anzahl der Absolventen, die in ihr enthalten sind. Dadurch besteht die Möglichkeit, dass die Referenzgruppe zu klein wird.

Als Hinführung auf eine Lösung werden nun das Ausscheiden aus der Referenzgruppe, sowie die Anforderung nach der Mindestgröße der Referenzgruppe diskutiert.

Das Ausscheiden von Jahrgängen aus der Referenzgruppe würde zu weniger Problemen führen, wenn es gleichmäßig verlaufen würde. Im Falle des Verfahrens der wandernden Kohorte fällt ein Jahrgang als Ganzes aus der Referenzgruppe. Würde stattdessen ein Jahrgang mit jeder jährlichen Aktualisierung an Einfluss auf die Referenzgruppe verlieren, könnten unerwartete Änderungen in der ECTS-Notenzuweisung verhindert werden.

Im Hinblick auf die Größe der Referenzgruppe ist die Notwendigkeit einer Mindestgröße zu hinterfragen. Die Forderung nach einer Mindestgröße ist statistisch begründet. Würde man eine ECTS-Notenzuweisung auf einer sehr kleinen Referenzgruppe aufbauen, ist die Wahrscheinlichkeit sehr hoch, dass die enthaltenen Absolventen nicht repräsentativ sind und dass mehrere ECTS-Notenklassen auf eine lokale Note zusammenfallen. Weiterhin würde bereits eine geringe Anzahl an neu hinzugefügten Absolventen reichen, um die Referenzgruppe und damit die ECTS-Notenzuweisung grundlegend zu verändern. Dies bedeutete wiederum unerwartete Änderungen, die den geforderten Eigenschaften widersprechen. Um das Problem zu lösen, wird folgende Aussage getroffen: Wenn sichergestellt werden kann, dass die Referenzgruppe über eine hinreichende Größe bzw. statistische Qualität verfügt, so ist nur noch die statistische Qualität des jeweils hinzuzufügenden Absolventenjahrgangs zu kontrollieren. Ist diese nicht ausreichend, müssen hinreichende Maßnahmen bezüglich der Aktualisierung der ECTS-Notenzuweisung getroffen werden. Der Lösungsansatz ist also, die Anforderung nach einer Mindestgröße zu verlagern. Anstatt eine Mindestgröße der Referenzgruppe zu fordern, wird eine Mindestgröße für die Absolventenjahrgänge bei der Aktualisierung vorausgesetzt.

2.1 Lösung

Die Lösung lautet wie folgt:

1. Statt einer Referenzgruppe wird eine relative Referenzverteilung in den Mittelpunkt gestellt.
2. Statt dem vollständigen Entfernen von Absolventen nach einer gewissen Zeit aus der Referenzgruppe werden alle Bewertungen im Laufe der Zeit entwertet.
3. Ein neuer Absolventenjahrgang geht mit einem vorgegebenen Gewicht in die Referenzverteilung ein, d.h. unabhängig von der Größe des Jahrgangs. Ausgenommen davon sind Absolventenjahrgänge, die zu klein sind. In diesem Falle wird das Gewicht reduziert, um die statistische Qualität der Referenzverteilung zu gewährleisten.

Die drei Punkte werden nun im Detail erläutert. Bezüglich des ersten Punktes stellt die Verwendung einer Referenzgruppe zur Bestimmung der ECTS-Notenzuweisung zu viele Informationen bereit. Die Referenzgruppe trifft Aussagen darüber, wie viele Absolventen die Note 1; 1,3; 1,7; usw. erreicht haben. Wie in der Einführung erläutert, ist für die ECTS-Notenzuweisung jedoch nur relevant, wie viel Prozent der Absolventen die jeweilige Note erreichen. Dementsprechend kann statt einer Referenzgruppe (absolute Häufigkeitsverteilung) eine relative Häufigkeitsverteilung verwendet werden. Im Folgenden wird für diese relative Häufigkeitsverteilung der Begriff *Referenzverteilung* benutzt. An dieser Stelle ist anzumerken, dass eine Referenzverteilung größenlos ist. Sie trifft also keine Aussage über die Anzahl an Absolventen, die in die Referenzverteilung eingegangen sind. Doch gerade da sie keine Aussage hierzu trifft, erleichtert die Verwendung der Referenzverteilung die Lösung des Kaltstartproblems. Bei der Verwendung einer Referenzgruppe muss zur Lösung des Kaltstartproblems sowohl (a) die Verteilung der Bewertungen als auch (b) die Anzahl der Bewertungen bekannt sein. Wie bereits

erläutert, stellt der Teil (b) im Hinblick auf Wahlbereiche und sich ändernde Studiengangsinhalte ein nicht zufriedenstellend lösbares Problem dar. Bei der Verwendung einer Referenzverteilung entfällt der Teil (b). Es ist ausschließlich die Angabe einer relativen Verteilung vonnöten.

Der zweite Punkt verfolgt die angedachte kontinuierliche Entwertung von Jahrgängen. Das anteilige Entfernen von alten Jahrgängen ist durch die Verwendung der Referenzverteilung einfach zu realisieren. Dementsprechend erlaubt die Referenzverteilung die Verwendung von glättenden Verfahren zur Aktualisierung der ECTS-Notenzuweisung. Ein unendlicher Verbleib von Absolventen in der Referenzverteilung, auch wenn es sich um Ausreißer handelt, ist denkbar und stellt kein Problem dar. Wenn bei der jährlichen Aktualisierung ältere Jahrgänge nur noch mit immer geringerem Gewicht eingehen, üben diese Jahrgänge auf die ECTS-Notenzuweisung einen ebenfalls sehr geringen, vernachlässigbaren Effekt aus.

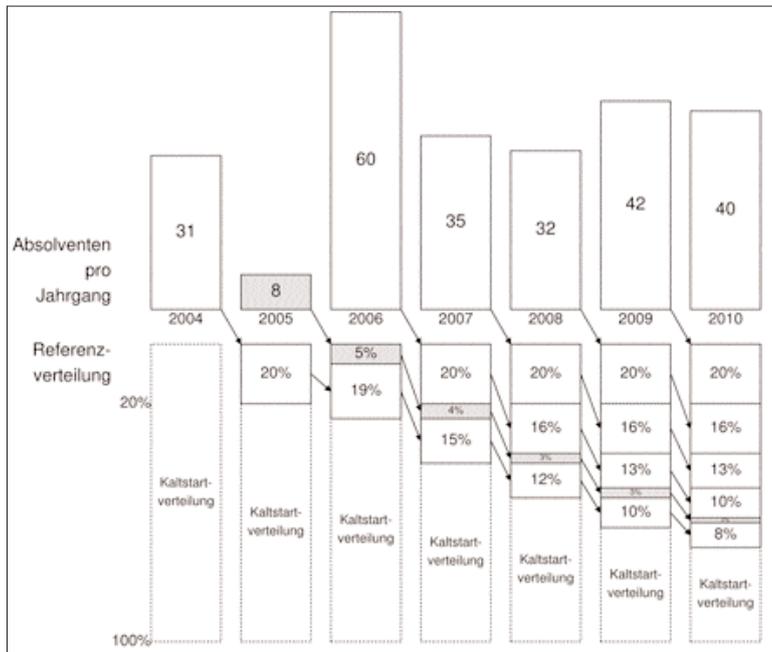
Zum dritten Punkt existieren zwei Leitideen: Die erste ist, dass ab einer gewissen Größe eines Absolventenjahrgangs die statistische Qualität ausreichend ist, d.h. eine repräsentative Verteilung mit hinreichend hoher Wahrscheinlichkeit angenommen werden kann. Wird diese gewisse Größe überschritten, ist die exakte Zahl von Absolventen irrelevant. Folglich wird in diesem Falle der Einfluss auf die Anpassung der ECTS-Notenzuweisung konstant gehalten. Die zweite Leitidee behandelt den Fall, dass der Absolventenjahrgang klein ist. In diesem Falle sind Extremverteilungen wahrscheinlicher. Wenn beispielsweise der Jahrgang aus nur zwei Absolventen besteht, werden maximal nur zwei unterschiedliche Noten vergeben. In diesem Falle soll der Einfluss auf die aktualisierte ECTS-Notenzuweisung entsprechend der Größe des Absolventenjahrgangs angepasst werden.

3. Verfahren

Um die vorgestellte Lösung operativ einsetzen zu können, sind weitere Festlegungen zu treffen. Als Erstes wird festgelegt, dass bei einer Aktualisierung alle alten Absolventenjahrgänge um den gleichen prozentualen Anteil entwertet werden. Somit lassen sich alle alten Absolventenjahrgänge zu einer Verteilung zusammenfassen und müssen nicht mehr getrennt gespeichert werden. Außerdem werden alte Absolventenjahrgänge nie aus der Referenzverteilung entfernt. Als Zweites ist der Anteil festzulegen, mit dem ein neuer Absolventenjahrgang üblicherweise in die Referenzverteilung eingeht. Angelehnt an den vorgeschlagenen Fünfjahreszeitraum (EU-Kommission 2005, S. 15) wird das Gewicht für den neuen Absolventenjahrgang auf $1/5$ gesetzt. Als Mindestgröße wird ebenfalls auf die empfohlenen 30 Absolventen zurückgegriffen. Besteht der neue Jahrgang aus weniger als 30 Absolventen, wird das Gewicht linear zwischen $1/5$ (30 Absolventen) und 0 (keine Absolventen) interpoliert. Zusammenfassend entspricht diese Vorgehensweise dem Verfahren der exponentiellen Glättung (Brown 1963), wobei das Gewicht als Funktion der Größe des Absolventenjahrgangs definiert ist.

Die operative Anwendung wird nun exemplarisch an einem neu eingerichteten Modul illustriert. Das Kaltstartproblem ist durch die Angabe einer relativen Verteilung (Kaltstart-

Abbildung 4: Entwicklung der Referenzverteilung (prozentual gerundet, Höhe des Balkens entspricht Anteil)



Drei- statt an einen Fünfjahreszeitraum bei der Bestimmung des maximalen Gewichts für neue Absolventenjahrgänge bei der Aktualisierung.

4. Zusammenfassung

Die ECTS-Notenzuweisung basiert auf einer Datenbasis, die Aussagen über die Verteilung der Bewertungen unter den Absolventen eines Moduls, Faches oder Studiengangs trifft. Die Datenbasis wird nach den existierenden Empfehlungen der EU (EU-Kommission 2005; 2009) oft als Zusammenfassung der Absolventen der vorigen Jahre gebildet. Diese Referenzgruppe oder Kohorte enthält Informationen zur Verteilung der Bewertungen und zur Anzahl der vergebenen Bewertungen im gewählten Zeitraum. Letztere Information ist für die Berechnung der ECTS-Notenzuweisung nicht erforderlich, führt aber in verschiedenen, plausiblen Situation zu Problemen. Diese Probleme erfordern einen entscheidenden organisatorischen und juristischen Mehraufwand und es besteht die Gefahr, dass das politische Ziel der ECTS-Noten ad absurdum geführt wird. Wird als Datenbasis jedoch eine (relative) Referenzverteilung verwendet und das vorgestellte

Verfahren zur Aktualisierung der Datenbasis verwendet, werden die benannten Probleme vermieden. Das Verfahren zeichnet sich durch Stabilität, Einfachheit und Transparenz aus und ist auch für die Automatisierung der ECTS-Notenzuordnung auf jeder Ebene (Studiengang, Fach, Modul, ...) geeignet.

Literaturverzeichnis

Brown, R.G. (1963): Smoothing, Forecasting and Prediction of Discrete Time Series. Prentice Hall.
 EU-Kommission (2005): ECTS Users' Guide. Directorate-General for Education and Culture.
 EU-Kommission (2009): ECTS-Users' Guide. Directorate-General for Education and Culture.
 HRK (2005): ECTS als System zur Anrechnung, Übertragung und Akkumulierung von Studienleistungen. In: HRK Arbeitsbericht 2004/2005, S. 89ff.
 Kunze, J./Geyer-Schulz, A. (2010): ECTS-Noten: Verfahren zur Kohortenberechnung, Technischer Bericht em2010-1, KIT, http://www.em.uni-karlsruhe.de/literatur/lit_show.php?id=15613
 Webler, W.-D. (2010): Internationale Vergleichbarkeit von Noten im Hochschulbereich? Problematik der Notenvergabe, Referenzgrößen und der Verwendung der Gaußschen Normalverteilung, In: Qualität in der Wissenschaft, Jg. 4/H. 1, S. 20-23.

■ **Jonas Kunze**, Dipl.-Inform.Wirt, Studienprogrammkoordination, Karlsruher Institut für Technologie (KIT), E-Mail: jonas.kunze@kit.edu
 ■ **Dr. Andreas Geyer-Schulz**, Lehrstuhl für Informationsdienste und elektronische Märkte, Institut für Informationswirtschaft und -management, Karlsruher Institut für Technologie (KIT), E-Mail: andreas.geyer-schulz@kit.edu

verteilung) einfach lösbar. Die Entwicklung der Referenzverteilung über die Jahre hinweg ist in Abbildung 4 dargestellt. An dieser Stelle ist zu erwähnen, dass das Gewicht der einzelnen Jahrgänge nur aus Verständnisgründen in der Abbildung dargestellt ist. Da alle alten Absolventenjahrgänge mit dem gleichen Faktor entwertet werden, ist es algorithmisch problemlos möglich, diese Jahrgänge jeweils zusammenzufassen.

Es ist zu sehen, dass schwankende Absolventenzahlen ohne verzerrenden Einfluss auf die Verteilung bleiben. Bei sehr geringen Absolventenzahlen wie im Jahre 2005 wird die Referenzverteilung nur minimal angepasst, während ein großer Jahrgang wie im Jahre 2006 kein herausragendes Gewicht erhält. Für einen detaillierten Vergleich dieses Aktualisierungsverfahren mit dem Verfahren der wandernden Kohorte und einer weiteren Variante der exponentiellen Glättungen siehe (Kunze/Geyer-Schulz, 2010).

Das vorgestellte Verfahren ist nicht das einzig denkbare Verfahren. Es ist jedoch eines der einfachsten Verfahren, welches alle genannten Probleme deutlich abschwächt bzw. verhindert. In den in (Kunze/Geyer-Schulz, 2010) durchgeführten Untersuchungen erweist sich das Verfahren als stabil und gleichzeitig ausreichend flexibel, sowie als sehr einfach umzusetzen. Die konkrete Wahl der Parameter ist durch die Vorgaben der EU-Kommission motiviert und stellt eine vergleichsweise konservative Wahl dar, welche die Stabilität in den Vordergrund stellt. Es ist durchaus denkbar, dass zukünftige empirische Untersuchungen eine Anpassung der Parameter nahelegen, z.B. die Anlehnung an einen

Wie kann Qualitätsmanagement zur Entwicklung der Hochschullehre beitragen? Tagung der Internationalen Bodensee-Hochschule IBH Dienstag, 29. Juni 2010

Qualitätsmanagement und Qualitätssicherung haben an Hochschulen eine große Bedeutung. Ziel der Tagung ist es, aufgrund aktueller Forschungsergebnisse die Diskussion zur Qualität in der Lehre an Hochschulen anzuregen, verschiedene Maßnahmen zur Qualitätsoptimierung im Sinne von Best-Practice-Beiträgen kennen zu lernen und deren Wirkung zu diskutieren.

Das Hochschulmanagement, die Bildungswissenschaft und die Hochschulangehörigen haben unterschiedliche Erwartungen an Qualitätsmanagement. Referentinnen und Referenten aus der Schweiz, Deutschland und Österreich beleuchten Qualitätsmanagement an Hochschulen aus verschiedenen Perspektiven. Sie stellen konkrete Massnahmen zur Qualitätsoptimierung der Lehre aus ihren Hochschulen vor. Im Weiteren werden erste Ergebnisse des laufenden Forschungsprojektes „Von der Qualitätsevaluation zur Qualitätsoptimierung in der Hochschullehre“ der Pädagogischen Hochschulen St.Gallen, Vorarlberg und Weingarten präsentiert und diskutiert.

Die Tagung richtet sich an Hochschulleitungen, Qualitätsverantwortliche, Dozent-innen und Dozenten, Personen der Hochschulpolitik sowie weitere am Thema Interessierte.

Workshops

WS 1 Der Beitrag von Qualitätszirkeln zur Weiterentwicklung der Lehre und anderer Leistungsbereiche
Prof. Dr. Traugott Zech, PH Vorarlberg

WS 2 Qualitätszirkel praktisch erproben
Prof. Dr. habil. Ursula Pfeiffer & Luzia Sievi M.A., PH Weingarten

WS 3 Von der Lehrevaluation zum Qualitätszirkel zur Optimierung der Lehre
Lic.phil. Verena Messerli & Lic.phil. Benita Affolter, PH St.Gallen

WS 4 Im Praxistandem die eigene Lehre optimieren
Prof. Dr. habil. Diethelm Wahl, PH Weingarten

WS 5 Auf dem Weg zum Prozessmanagement an einer Hochschule
Dr. Heidrun Schöch, Fachhochschule Vorarlberg

WS 6 Gute Lehrende – gute Lernende – gute Lehre: Die Bedeutung von Input-variablen für die Qualität von Lehr-/Lern-Prozessen
Prof. Dr. José Gomez, dipl. Hdl. HSG, Zentrum für Hochschulbildung der FHS St.Gallen

Hinweise

Kosten
CHF 100.- bzw. EUR 70.-
Studierende: CHF 20.- bzw. EUR 15.-
(inklusive Mittags- und Pausenverpflegung)

Anmeldung
Bis Montag, 17. Mai 2010 mit beiliegendem Anmeldeformular oder online unter www.phsg.ch/qm-tagung.

Kontakt
Für die Tagungsleitung:
Lic.phil. Verena Messerli, E-Mail: verena.messerli@phsg.ch
Tagungssekretariat:
Debora Egli-Keller, E-Mail: debora.egli@phsg.ch

Tagungsort
PHSG, Hochschulgebäude Stella Maris
Müller-Friedbergstrasse 34
CH-9400 Rorschach

Lageplan
Einen Lageplan zum Tagungsort finden Sie unter www.phsg.ch > Campus > Hochschulgebäude > Stella Maris

Liebe Leserinnen und Leser,

nicht nur in dieser lesenden Eigenschaft (und natürlich für künftige Abonnements) sind Sie uns willkommen.

Wir begrüßen Sie im Spektrum von Forschungs- bis Erfahrungsberichten auch gerne als Autor/in.

Wenn das Konzept der „Qualität in der Wissenschaft“ Sie anspricht - wovon wir natürlich überzeugt sind - dann freuen wir uns über Beiträge von Ihnen in den ständigen Sparten

- Qualitätsforschung,
- Qualitätsentwicklung/-politik,
- Anregungen für die Praxis/Erfahrungsberichte, aber ebenso
- Rezensionen,
- Tagungsberichte,
- Interviews.

Die Hinweise für Autorinnen und Autoren finden Sie unter: www.universitaetsverlagwebler.de.

Hauptbeiträge der aktuellen Hefte Fo, HSW, HM, P-OE und ZBS

Auf unserer Homepage www.universitaetsverlagwebler.de erhalten Sie Einblick in das Editorial und Inhaltsverzeichnis aller bisher erschienenen Ausgaben.

Fo

Forschung

Politik - Strategie - Management

Fo 3+4/2009

Deutsch als Wissenschaftssprache -
Wissenschaft ist vielsprachig

Forschung über Forschung

Arie Rip

Towards Post-Modern Universities

Forschungsentwicklung/-politik

Tim Flink

Außenwissenschaftspolitik: ein neues
Handlungsfeld?

Wolff-Dietrich Webler

Ausbau der Promotions- und Postdoc-
Phase für vielfältige Aufgaben über For-
schung hinaus - Teil IV

Konrad Ehlich

Deutsch als Wissenschaftssprache
für das 21. Jahrhundert

Svetlana Nikolova

On the use of "Lingua Franca" and local
languages in the publications on the hu-
manities

*Ralph Mocikat, Wolfgang Haße &
Hermann H. Dieter*

Sieben Thesen zur deutschen Sprache
in der Wissenschaft

Falk Reckling & Christoph Kratky

Die Wissenschaft spricht englisch - aber
nicht nur

Christoph Kratky

Die Positionierung des Fonds zur Förde-
rung der wissenschaftlichen Forschung
(FWF) zum Thema Antragssprache

Forschungsgespräche

Gespräch mit Thomas Brunotte

Gespräch mit Prof. Dr. Ralph Mocikat

Gespräch mit Dr. Peter Gauweiler

HSW

Das Hochschulwesen

Forum für Hochschulforschung, -praxis und -politik

HSW 1/2010

Wissenschaftliches Schreiben
im Studium

Hochschulentwicklung/-politik

Tobias Sander

„Den Menschen da abholen wo er
steht“ – Kompetenzkonzept und
Hochschulbildung

*Roman Banzer, Anja Zwingenberger &
Christina vom Brocke*

Schreibintensive Seminare
Ein Seminar-konzept zum wissen-
schaftlichen Schreiben im Bachelor-
Studium

Swantje Lahm

Lehrend in die Wissenschaft.
Die Qualifizierung von
Doktorand/innen für schreibintensive
Lehre am John S. Knight Institute
for Writing in the Disciplines, Cornell
University, USA

Anregungen für die Praxis/
Erfahrungsberichte

Philipp Mayer

Wissenschaftlich schreiben heißt vor
allem denken – zwölf Techniken für
mehr Effizienz

Jörg Erik Kinner

Förderung der Sprachkompetenz
durch vorlesungsbegleitende
Schreibübungen

HM

Hochschulmanagement

Zeitschrift für die Leitung, Entwicklung und Selbstverwaltung von
Hochschulen und Wissenschaftseinrichtungen

HM 1/2010

Entwicklung, Gestaltung und
Verwaltung von Hochschulen
und Wissenschaftseinrichtungen

Christian Marettek & Ákos Barna

Aktuelle Probleme des Hochschul-
managements im Rahmen der
„deregulierten Hochschule“

Otto Hüther

Konflikte zwischen Hochschulräten
und akademischen Selbstverwaltungs-
gremien

*Shiho Futagami, Uschi Backes-Gellner
& Kerstin Pull*

Stand und aktuelle Herausforderungen
des japanischen Hochschulsystems

Dorothea Jansen & Insa Pruiskien

Transaktionskosten vermeiden -
Performanzprofile beachten -
Chancen für unkonventionelle
Forschung bereit stellen.

P-OE

Personal- und Organisationsentwicklung
in Einrichtungen der Lehre und Forschung

Ein Forum für Führungskräfte, Moderatoren, Trainer, Programm-Organisatoren

POE 1/2010

Varianten der Betreuung studentischen Lernens: Tutorien und Mentoring

Personal- und Organisationsentwicklung/-politik

Christian K. Karl

Ausgezeichnet – Hochschuldidaktische Innovationen in den Bauwissenschaften

Ingeborg Stahr & Franz Bosbach

Gut beraten: Das Mentoring-System der Universität Duisburg-Essen

Nicole Auferkorte-Michaelis & Annette Ladwig

Kompetenzentwicklung ECTS-kreditiert: Lernen im Tutorienprogramm an der Universität Duisburg-Essen

Tobina Brinker, Anne Barkey & Eva-Maria Schumacher

Studienbegleitende Tutorien zur Beratung und Unterstützung der Studierenden im Fachbereich Sozialwesen der Fachhochschule Bielefeld

Wolff-Dietrich Webler

Sekretärinnen mit besonderen Aufgaben - eine vergessene Gruppe der PE? Konzept zur Aus- und Weiterbildung von exponierten Sekretärinnen bzw. Assistentinnen der Hochschulleitung

Anregungen für die Praxis/ Erfahrungsberichte

Anja von Richthofen, Ina Voigt & Michael Lent

Weiterentwicklung der Berufungsverfahren an Hochschulen Teil 2: Erfahrungsberichte

ZBS

Zeitschrift für Beratung und Studium

Handlungsfelder, Praxisbeispiele und Lösungskonzepte

ZBS 1/2010

Beratung zum MINT-Studium

Beratungsentwicklung/-politik

Annemarie Cordes

tasteMINT: ein Potenzial-Assessment-Verfahren für Abiturientinnen im Übergang Schule – Hochschule

Wolfgang Loggen

Die Auswirkungen der MINT-Aktivitäten auf die Beratung in der Zentralen Studienberatung – eine Einschätzung

Gunvald Herdin

Fächerpräferenzen von Studienanfänger/innen 2006/2007 - Regionale und geschlechtsspezifische Muster. Eine Sekundäranalyse der amtlichen Statistik unter besonderer Berücksichtigung der MINT-Fächer - Zusammenfassung der Ergebnisse der Studie

Anregungen für die Praxis/ Erfahrungsberichte

Joachim Klaus

Das Studienzentrum für Sehgeschädigte (SZS) des Karlsruher Instituts für Technologie: Förderung von Schlüsselqualifikationen für Studium und Beruf

Jessica Kleinhelftwes

Das Learning Skills Center der University of Texas in Austin. Aufbau, Organisation und Arbeitsweise

Petra Gerlach

Die Aufbauphase der Informationskampagne „Gscheit studiert“



Für weitere Informationen

- zu unserem Zeitschriftenangebot,
- zum Abonnement einer Zeitschrift,
- zum Erwerb eines Einzelheftes,
- zum Erwerb eines anderen Verlagsproduktes,
- zur Einreichung eines Artikels,
- zu den Autorenhinweisen

oder sonstigen Fragen, besuchen Sie unsere Verlags-Homepage:

www.universitaetsverlagwebler.de

oder wenden Sie sich direkt an uns:

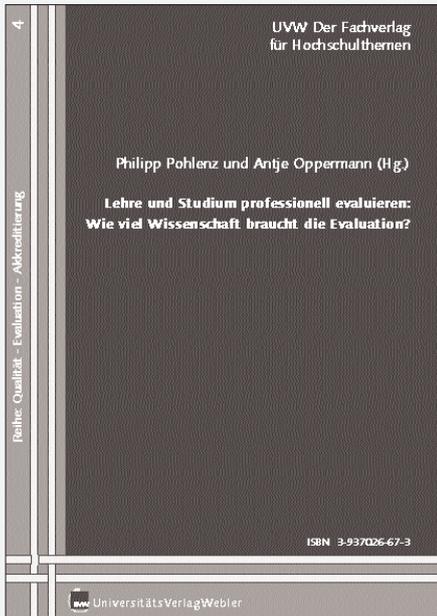
E-Mail: info@universitaetsverlagwebler.de

Telefon: 0521/ 923 610-12

Fax: 0521/ 923 610-22

Postanschrift:
UniversitätsVerlagWebler
Bünder Straße 1-3
33613 Bielefeld

**Philipp Pohlenz und Antje Oppermann (Hg.):
Lehre und Studium professionell evaluieren: Wie viel Wissenschaft braucht die Evaluation?**



In den letzten Jahren ist das Aufgabenspektrum der Evaluation von Lehre und Studium deutlich breiter geworden: Sie soll Leitungsentscheidungen auf eine informierte Grundlage stellen, soll Wirkungen der Lehre im Sinne eines Wissens- und Kompetenzgewinns der Studierenden messen, soll dadurch zur Qualitätssicherung und –entwicklung und zu Innovationen in Lehre und Studium beitragen. Analog steigen die Anforderungen, die an die „Evaluatorinnen und Evaluatoren“ zu richten sind. Diese beziehen sich bspw. auf methodische Kompetenzen oder Feldkenntnisse im Bereich des Managements von Hochschulen.

Gleichzeitig ist Evaluation zumeist nicht unmittelbar im Wissenschaftsbetrieb in Forschung und Lehre angesiedelt. Vielmehr ist sie dabei, sich zu einer Leitungsaufgabe zu entwickeln, die sich ihrerseits auf die Unterstützung des Wissenschaftsbetriebes bei der akademischen Selbstreflexion bezieht.

Die 10. Jahrestagung des Arbeitskreises Qualitätsmanagement und Evaluation der Berliner und Brandenburger Hochschulen (im März 2009 an der Universität Potsdam) machte diese Entwicklungen zum Thema. Sie fragte danach, wie viel Wissenschaft die Evaluation braucht, um ein wissenschaftsadäquates Qualitätssicherungsinstrument zu sein.

ISBN 3-937026-67-3, Bielefeld 2010, 257 Seiten, 29.50 Euro

Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

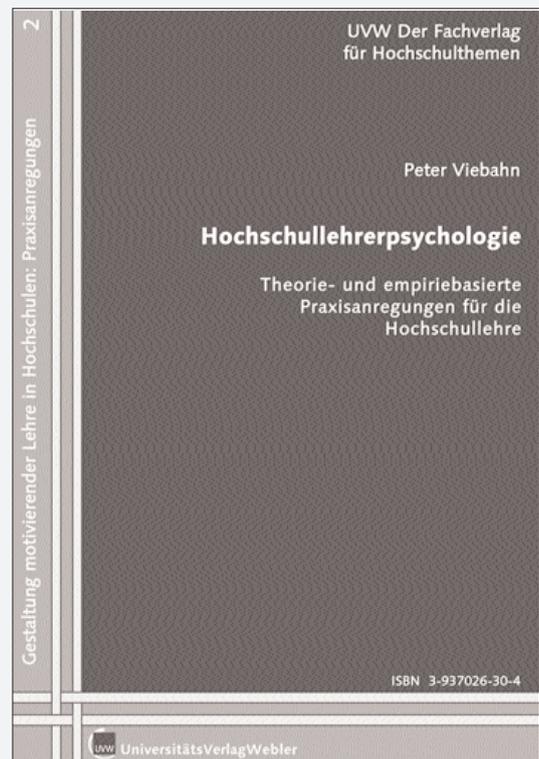
**Peter Viebahn:
Lernerverschiedenheit und soziale Vielfalt im Studium
Differentielle Hochschuldidaktik aus psychologischer Sicht**

Mit der Einführung der gestuften Studiengänge und der Internationalisierung der Ausbildung hat sich das Bildungsangebot von Hochschulen in hohem Maße ausdifferenziert und es werden zunehmend unterschiedliche Studierendengruppen angesprochen. Diese Entwicklung konfrontiert die Hochschuldidaktik in verschärfter Weise mit der grundsätzlichen Problematik: Wie kann die Lernumwelt Hochschule so gestaltet werden, dass dort ganz unterschiedliche Studierende ihr Lernpotential entfalten können? Eine Antwort auf diese Frage gibt diese Arbeit. Sie führt in das Konzept der Differentiellen Hochschuldidaktik ein. Im allgemeinen Teil werden hochschuldidaktisch relevante Modelle zur Individualität des Lernens (z.B. konstruktivistischer Ansatz) und die bedeutsamen psychischen und sozialen Dimensionen studentischer Unterschiedlichkeit in ihrer Bedeutung für das Lernen erläutert. Im angewandten Teil wird eine Vielzahl von konkreten Anregungen zur Optimierung des Lernens für die verschiedenen Lernerguppen geboten.

Ein Autoren- und ein Sachwortverzeichnis ermöglichen eine gezielte Orientierung.

Dieses Buch richtet sich an Hochschuldidaktiker, Studienplaner und Lehrende, die einen produktiven Zugang zur Problematik und Chance von Lernerheterogenität finden wollen.

ISBN 3-937026-57-6, Bielefeld 2008, 225 Seiten, 29.80 Euro



Bestellung - Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22