

**Andreas MÜLLER & Bernhard SCHMIDT<sup>1</sup> (München)**

## **Prüfungen als Lernchance: Sinn, Ziele und Formen von Hochschulprüfungen**

### **Zusammenfassung**

Der Artikel greift die Frage nach den Zielen und der Gestaltung von Prüfungen vor dem Hintergrund der veränderten Anforderungen in modularisierten Studiengängen auf. Dabei wird für die Herstellung einer Passung zwischen Lernzielen, Lehrmethoden und Prüfungsformen plädiert und auf die Gestaltung von Prüfungen in lehrer- wie in lernerzentrierten Lernarrangements eingegangen. Aus hochschuldidaktischer Perspektive wird verdeutlicht, welche – bisher noch wenig oder nur rudimentär genutzten – Chancen von Konzepten des peer- und self-assessment in lernerzentrierten Settings ausgehen. Daran anschließend werden die unterschiedlichen Funktionen und Ziele von Prüfungen vertieft analysiert und diskutiert, um schließlich an exemplarisch ausgewählten Lehr-Lern-Situationen die Möglichkeiten einer Stärkung der Lernmöglichkeiten durch Prüfungen über Feedback-Verfahren aufzuzeigen.

### **Schlüsselwörter**

Entwicklungsfunktion, Modularisierung, Prüfung, Rückmeldung, Lernen

## **Assessment as Opportunity to Learn: Significance, Aims and Forms of Higher Education Assessment**

### **Abstract**

This article concentrates on aims as well as the arrangement of assessments based on the requirements created by modularised study programmes. Therefore the article argues that learning targets, learning methods and assessment methods need to be matched. Moreover, the adequate design of tests in learner-centred as well as teacher-centred learning arrangements is addressed. From the perspective of teaching in higher education we discuss the prospects given by concepts of peer-assessment and self-assessment in learner-centred settings, which are rarely or only rudimentarily used so far. Afterwards functions and aims of assessments are analysed in depth. Finally by means of exemplary learning arrangements the options assessments and feedback methods provide for strengthening opportunities to learn are highlighted.

### **Keywords**

Development, modularisation, assessment, feedback, learning

---

<sup>1</sup> e-Mail: [b.schmidt@lmu.de](mailto:b.schmidt@lmu.de)

# 1 Rahmenbedingungen

Die Europäisierung der Hochschullehre im Rahmen des Bologna-Prozesses brachte tiefgreifende Reformen und eine völlige Umstrukturierung der Hochschullehre in Deutschland mit sich. Die Folgen der umfassenden Einführung gestufter Studiengänge und eines Credit-Point-Systems werden erst allmählich absehbar und die damit verbundenen Probleme, die zu erheblichen Teilen durch eine nicht intendierte Überregulierung von Seiten der Hochschulen (BÜLOW-SCHRAMM, 2008) verursacht werden, sind noch längst nicht gelöst. Eine Verbesserung der Qualität von Hochschullehre wird beispielsweise im Kommuniqué der Bologna-Folgekonferenz in Berlin 2003 als zentrales Ziel formuliert wird.

„The quality of higher education has proven to be at the heart of the setting up of a European Higher Education Area” (BERLIN-KOMMUNIQUE, 2003, S. 3).

Eine Annäherung an dieses Ziel ist bislang jedoch nicht erkennbar, geschweige denn dessen Realisierung. Durch die starke Expansion der Hochschulen in den 1960er bis 1980er Jahren, die nicht mit einem entsprechenden Ausbau der Universitäten selbst einher ging, sind die Voraussetzungen für nachhaltige Verbesserungen im Bereich der Hochschullehre keineswegs optimal (vgl. EBEL-GABRIEL, 2004). Gleichzeitig sind die Hochschulen durch die Europäisierung einem verschärften Wettbewerb ausgesetzt und müssen – um sich im europäischen Bildungsmarkt zu positionieren – den Anforderungen von Studierenden, Arbeitsmarkt und Bildungspolitik in gleichem Maße gerecht werden. Die daraus resultierenden Interessenskonflikte spiegeln sich auch in der Lehr- und Prüfungskultur an den Hochschulen wieder.

Herausforderungen ergeben sich insbesondere aus der Modularisierung. Durch die Modularisierung erhalten Prüfungen einerseits schon durch ihre höhere Dichte zusätzliches Gewicht, andererseits sind sie enger als in traditionellen Studiengängen mit dem Lernprozess selbst verknüpft sind (BÜLOW-SCHRAMM, 2008). Letzteres ergibt sich aus der Verknüpfung von Modulen mit einer klaren Definition der zu erwerbenden Kompetenzen. Zwar ist die Definition von Lernzielen Grundvoraussetzung für erfolgreiche Lehre sowie für sinnvolle Prüfungen und daher von grundlegender Bedeutung, die Formulierung von Kompetenzen bringt aber in der Praxis – durch die mangelnde Operationalisierbarkeit – erhebliche Schwierigkeiten mit sich (BRAUN, SOELLNER & HANNOVER, 2006). Spürbar werden hier die Diskrepanz zwischen bildungspolitischen Vorgaben sowie wissenschaftlich fundierten Anforderungen einerseits und durch Unterausstattung sowie institutionelle Anforderungen eingeschränkte Handlungsspielräume andererseits.

## 2 Wozu prüfen? Eine systematische Annäherung an Funktionen von Prüfungen

Prüfungen spielen seit jeher im Kontext von (institutionellen) Lehr- und Lernprozessen eine bedeutsame Rolle, zudem gewinnen sie aktuell – wie bereits geschildert – im universitären Kontext durch die Umstellung auf Bachelor- und

Masterstudiengänge an Bedeutung<sup>2</sup>. Die zunehmende Popularität des Themenbereichs drückt sich in einer steigenden Anzahl von Veröffentlichungen aus, die sich vertieft mit der Notwendigkeit der Kompetenzorientierung von Prüfungen beschäftigen, Konzepte aus der universitären Praxis vorstellen (vgl. zusammenfassend DANY et al., 2008) oder sogar die Etablierung einer eigenen (Teil-) Wissenschaft 'Prüfungsdidaktik' anregen (TINNEFELD, 2002). Diesem praktischen Gewicht des Gegenstands und Handlungsfeldes 'Prüfen' zum Trotz, fällt bei einem ersten Blick auf die einschlägige Literatur zusehends auf, dass eine theoretische, inhaltliche sowie systemische Fundierung bislang noch größtenteils aussteht. So bleibt etwa die grundlegende Frage in welchem Verhältnis z.B. Prüfungen und Evaluationsmaßnahmen zueinander stehen, welche Funktionen Prüfungen für verschiedene beteiligte Personen oder Institutionen erfüllen oder auch wie bestimmte Funktionen realisiert werden können bislang weitgehend offen. In den folgenden Abschnitten soll versucht werden, anhand einiger praxisnaher Überlegungen einen Beitrag zur diesbezüglichen Diskussion zu leisten und somit die Grundlage für weiterführende Überlegungen zur zeitgemäßen Nutzung und Gestaltung von Prüfungen zu legen.

Die wohl differenzierteste Systematisierung von Prüfungsfunktionen stammt von FLECHSIG (1976). Er unterscheidet zwischen insgesamt elf Funktionen, die sich den drei Kategorien Herrschafts- und Sozialisierungsfunktion, Rekrutierungsfunktion sowie Didaktische Funktion zuordnen lassen (vgl. BÜLOW-SCHRAMM & GIPSER, 1994; MÜLLER & BAYER, 2007; REIS & RUSCHIN, 2008) (vgl. Tabelle 1).

Tabelle 1: Funktionen von Prüfungen (nach FLECHSIG, 1976)

Herrschafts- und Sozialisierungsfunktion (Systemreproduktion)
Initiation
Statusverteilung
Legitimation
Rekrutierungsfunktion (Selektion, Auswahl, Zuordnung, Linearisierung)
Platzierung in Kohorte
Auslese aus Kohorte
Qualifikations- und Kompetenznachweis
Didaktische Funktion (Systemoperationen)
Zeitliche und inhaltliche Gliederung des Studiengangs
Orientierung der Lehrenden und Lernenden über die Studienziele
Extrinsische Lernmotivation
Diagnoseinstrument
Rückmeldung des Lehr-Lernerfolgs an Lehrende und Lernende

<sup>2</sup> Dies wird eindrücklich verdeutlicht durch den Titel „Prüfungen auf die Agenda!“ (DANY, SZCZYRBA & WILDT, 2008) einer der letzten Ausgaben der Reihe Blickpunkt Hochschuldidaktik.

Im Folgenden werden die angeführten Funktionskategorien getrennt voneinander kurz vorgestellt und anhand von Beispielen für die drei beteiligten Gruppierungen Studierende, Hochschule und Gesellschaft veranschaulicht (vgl. Tabelle 2).

## 2.1 Herrschafts- und Sozialisierungsfunktion

Die erste Kategorie beschreibt nach FLECHSIG (1976) Funktionen von Prüfungen unter dem Fokus von Herrschaftsausübung und Sozialisierung. Prüfungen dienen hierbei aus systemischer Perspektive vor allem dem Erhalt und der Reproduktion vorhandener Zustände und Machtkonstellationen (REIS & RUSCHIN, 2008). So sind Prüfungen etwa vergleichbar einem Initiationsritual, das den Zugang zu bestimmten Gruppierungen und Ressourcen erlaubt oder sie legitimieren einen sozialen Status, der unabhängig von der tatsächlich ausgeübten Tätigkeit mit einem gewissen Prestige verbunden ist.

## 2.2 Rekrutierungsfunktion

Die zweite Kategorie von Funktionen wird von FLECHSIG (1976) mit Rekrutierung betitelt – im Kontext von Evaluation ist hierfür derzeit der Begriff „Selektion“ gebräuchlich. Deutlich weniger belastet erscheint der von REIS & RUSCHIN (2008) eingebrachte Terminus ‚Linearisierung‘ oder auch das von INGENKAMP & URBAN (2005, S. 34) im Rahmen der pädagogischen Diagnostik vorgeschlagene ‚zuordnen‘ von Personen zu Angeboten oder Bedingungsklassen. Inhaltlich gemein ist allen Bezeichnungen, dass über Prüfungen eine Auswahl getroffen wird, sei es aus sozialer Perspektive durch Rangreihen (Platzierung) oder mittels einem vorab festgelegten Kriterium bzw. Schwellenwert, der überschritten werden muss (Auslese). In Anlehnung an die funktionale Differenzierung von Evaluationen lässt sich dieser Funktionsbereich auch summative Bewertung bezeichnet werden. Summativ bezieht sich hier jedoch weniger auf die zeitliche Anordnung einer Prüfung am Ende einer Veranstaltung oder eines Studiums (denn Moduleinzelnoten könnten durchaus als zeitlich-formative Prüfungen im Rahmen eines gesamten Studiums angesehen werden), sondern beschreibt primär deren Charakter als Instrument der verbindlichen Notengebung und Selektion (vgl. IRONS, 2008).

Aus studentischer Sicht können über selektierende Mechanismen zum einen künftige Laufbahnoptionen erschlossen bzw. ausgeschlossen werden (z.B. über den Numerus Clausus) oder aber im Rahmen von Bewerbungen Zugang zu Institutionen erlangt werden. Auch die Hochschulen selbst treten als rekrutierende Einrichtungen auf, sei es eher intern (Anstellung von wissenschaftlichen Hilfskräften) oder auch nach außen gerichtet (Neubesetzung von Lehrstühlen). Im Verlauf eines Studiengangs sind Prüfungen per se notwendig für einen erfolgreichen Studienfortschritt, dies konkretisiert sich bei bestimmten Veranstaltungen deren Bestehen die Zugangsvoraussetzung für nachfolgende Module darstellt<sup>3</sup>. Auch hinsichtlich

---

<sup>3</sup> Hiermit sind insbesondere so genannten Grundlagen- und Orientierungsprüfungen gemeint, die von Studierenden im ersten Fachsemester abgelegt werden müssen (vgl. [http://www.uni-muenchen.de/aktuelles/amtl\\_bekannt/292-11pae-ba-08-ps00.pdf](http://www.uni-muenchen.de/aktuelles/amtl_bekannt/292-11pae-ba-08-ps00.pdf)).

der Zulassung zu weiterführenden Masterstudiengängen treten Hochschulen (hoch) selektiv auf. So wird aus derzeitiger Perspektive wohl nur für 20-30% der Bachelorstudierenden die Möglichkeit eines Anschlussstudiums bestehen. Aus gesellschaftlicher Perspektive manifestiert sich die Selektionsfunktion von Hochschulprüfungen vor allem über den Beleg einer formalen Qualifikation sowie in der Steuerung von Übergängen innerhalb des Bildungssystems und hin zum Beschäftigungssystem (DANY et al., 2008, S. 5).

### 2.3 Didaktische Funktion

Die dritte Kategorie umfasst nach FLECHSIG (1976) didaktische Funktionen. Prüfungen dienen demnach zum einen als zeitliche und inhaltliche Gliederungs- und Orientierungspunkte (vgl. REIS & RUSCHIN, 2008) sowie als Instrument der extrinsischen Lernmotivation (BRUNSTEIN & HECKHAUSEN, 2007; HECKHAUSEN, 1989).

Zum anderen erlauben Prüfungen aus diagnostischer Perspektive Rückschlüsse auf bis dato realisierte Lern- und Bildungsprozesse sowie mögliche künftige Verbesserungsmaßnahmen. Wie von verschiedensten Seiten betont wird, ist die hiermit verbundene Lernchance im Kontext von Prüfungen nicht nur auf die Prüflinge selbst zu beziehen, sondern vielmehr auch auf die prüfende Person sowie das prüfende System als gesamtes (DANY et al., 2008, S. 5; FLECHSIG, 1976, S. 305; TREMP & EUGSTER, 2006, S. 165).

Analog zur Funktionsdifferenzierung im Evaluationskontext lässt sich die skizzierte Entwicklungsperspektive als formative Bewertung kategorisieren. Idealerweise gehen mit einer solchen formativen Prüfung keine Zensuren einher. IRONS (2008, S. 7) betont jedoch unter Bezug auf die universitären Praxis in modularisierten Studiengängen, dass der Unterschied zu summativen Maßnahmen vor allem darin zu sehen ist, ob mit einer Prüfungssituation eine Entwicklungsmöglichkeit und Lernchance für die Studierenden verbunden ist. Realisiert werden kann dieser Förderimpuls über formative feedback welches verstanden wird als

„any information, process or activity which affords or accelerates student learning based on comments relating to either formative assessment or summative assessment activities“

Aus studentischer Sicht wird durch Prüfungen der aktuelle Lern- bzw. Kompetenzstand erfasst, woraus in einem nachgelagerten Schritt potentielle Entwicklungsfelder abgeleitet werden können. Welche Lernfelder im Detail aufgezeigt werden, hängt stark von Form und Kontext der Prüfung ab. Für lernerzentrierte Veranstaltungen wären offene Formen der Rückmeldung zu bearbeiteten Aufgaben, angewandten Vorgehensweisen und der Kooperation in der Lerngruppe nahe liegend.

Im Rahmen der Forschung zur Rückmeldung von Testresultaten gibt es unterschiedlichste Ansätze, die eine motivationale Wirkung von Feedback beschreiben: Einerseits lassen sich günstige Effekte einer positiven Rückmeldung in Anlehnung an THORNDIKES (1913; 1922) 'law of effect' oder auch kognitivistisch fundierte Ansätze wie DECI & RYANS (1985) 'cognitive evaluation theory' oder BANDURAS (1986; 1991) 'self efficacy theory' postulieren. Aus behavioristischer Perspektive wirken positive Rückmeldungen hierbei als Verstärker, die kognitive

Sichtweise betont dagegen die Erhöhung von intrinsischer Motivation und Selbstwirksamkeitserwartungen.

Andererseits gehen die Kontrolltheorie (CARVER & SCHEIER, 1981, 1982; PODSAKOFF & FARH, 1989) oder auch die Zielsetzungstheorie (LOCKE & LATHAM, 1990) davon aus, dass eher durch das Anzeigen einer negativen Abweichung von einer Bezugsnorm oder einem Standard Motivation generiert wird, aktiv zu werden.

Diese unterschiedlichen theoretischen Ansätze zur Wirkungsweise von Feedback spiegeln sich in indifferenten empirischen Befunden wieder. Wie HATTIE & TIMPERLEY (2007) in einem aktuellen Review betonen, können sowohl positive als auch negative Rückmeldungen leistungsförderlich und handlungsinitiierend wirken. Demgegenüber kommen KLUGER & DeNISI (1996, S. 275) am Ende ihrer viel zitierten Metaanalyse zu dem Ergebnis, dass weniger positives Feedback (Lob), sondern vor allem Rückmeldungen, die eine Zielverfehlung anzeigen, günstige Effekte nach sich ziehen.

Ergänzend zur didaktischen Funktion für Studierende, findet sich auf Seiten der Hochschule Entwicklungsaspekte auf individueller und systemischer Ebene: Der Dozent selbst kann im Falle einer schriftlichen Multiple-Choice-Prüfung über eine statistische Analyse der vorliegenden Antworten eine Modifikation der künftigen Prüfungsfragen vornehmen<sup>4</sup>. Unabhängig von der konkreten Prüfungsform ist es auch möglich, die Erreichung der selbst gesetzten Lernziele zu überprüfen (Kompetenz- oder Wissenserwerb). Bei einer Abweichung sind Anpassungen der Veranstaltung z.B. hinsichtlich des Inhalts, der Struktur (z.B. Verhältnis von Beiträgen der Studierenden und Dozent), des Workloads (z.B. Literaturvorgaben) oder der Betreuungsangebote denkbar.

Auch auf systemischer Ebene könnte im Anschluss an Prüfungen ein Ausbau von Betreuungsangeboten für Studierende ins Auge gefasst werden. Zusätzlich wären auch inhaltliche oder strukturelle Modifikationen bestehender Veranstaltungsreihen oder des Modulaufbaus denkbar, wenn etwa festgestellt wird, dass in einem späteren Semester Wissens- oder Kompetenzlücken in bestimmten Domänen bestehen.

Hinsichtlich der zuletzt skizzierten Entwicklungsfelder auf Hochschulebene bleibt jedoch kritisch zu beachten, dass die Rückmeldungen an die Prüfer sowie die Institution in diesem Fall eher indirekter Natur sind. D.h. Anregungen für Dozenten und das System Hochschule ergeben sich nur über den Umweg der studentischen Prüfungsleistung. Wenn man so will erfolgt hier also keine direkte Evaluation der Lehrleistung eines Dozenten, sondern eher eine versteckte Selbstevaluation seitens der prüfenden Partei.

Aus Entwicklungsperspektive ist fraglich, inwieweit die für die Ableitung, Initiierung und Realisierung von Veränderungsmaßnahmen relevante interne Attribuierung von Leistungsrückmeldungen durch die prüfenden Instanzen bei solch einem Setting überhaupt vorliegen kann. Realistisch erscheint hier, dass vor allem

---

<sup>4</sup> z.B. Analyse der Trennschärfe einzelner Items oder der Itemschwierigkeit.

bei negativen Prüfungsleistungen von Dozenten- und Hochschuleseite eine primäre Ursachenzuschreibung auf studentische Faktoren erfolgt (vgl. z.B. MUSCH & BRÖDER, 1999).

Dieser kritische Aspekt tritt auf gesellschaftlicher Ebene nochmals verstärkt zu Tage: per se können aus Prüfungen (v.a. Studienabschlussprüfungen), unter Beachtung einer längeren Zeitspanne, durchaus auch Rückschlüsse für politische Entscheidungsfindungsprozesse gezogen werden. So deuten überdurchschnittlich hohe Ausfallquoten auf eine mangelnde Passung von Prüfungsanforderungen und Lehr-Lernangeboten hin, die auf Ebene zentraler Staatsexamine (z.B. Lehramt oder Rechtswissenschaften) auch politisch bedingt sein können. Ob jedoch solche Zusammenhänge in der politischen Realität wirklich hergestellt werden bleibt fraglich.

Tabelle 2: Funktionen von Prüfungen differenziert nach Beteiligten

	Reproduktions-/ Selektions-/ Allokationsfunktion	Didaktische Funktion/ Entwicklungsfunktion
Individuum (Student)	Offenlegung von Laufbahnoptionen Bewerbung z.B. für Stipendien	Erfassung des aktuellen Lern- und Kompetenzstands Ableitung von Entwicklungsfelder Generierung von Motivation
Institution (Hochschule)	Auswahl bei Stellenbesetzung Zulassungsvoraussetzung für nachfolgende Module (insb. GOP) Bestehen notwendig für erfolgreichen Studienverlauf Zulassung zu weiterführendem Studium (Master)	Individuelle Ebene (Dozent): Anpassung künftiger Prüfungen (Itemkonstruktion) Erreichung von Lernzielen → Anpassung der Veranstaltung systemische Ebene (Institution): Betreuungsangebot Inhaltliche und strukturelle Modifikation von Veranstaltungsreihen/ Modulaufbau
Gesellschaft	Steuerung von Übergängen innerhalb des Bildungssystems und zum Beschäftigungssystem Beleg einer formalem Qualifikation	Rückschlüsse für politische Entscheidungsfindungsprozesse

### 3 Prüfungen als Teil der Lehre

BOUD (2007) stellt in einer vergleichenden Analyse von Prüfungsordnungen im angloamerikanischen Kulturraum fest, dass sich in diesen offiziellen Dokumenten primäre wie sekundäre Funktionen von Prüfungen erkennen lassen. Im Vordergrund stehen Begriffe wie Messung und Lernergebnis (outcome) sowie Integrität als primärer Fokus von Prüfungen, also Aspekte der Herrschafts-, Sozialisierungs- und Rekrutierungsfunktion von Prüfungen (vgl. FLECHSIG 1976). Erst in zweiter Linie werden die Rückmeldung an die Lernenden, deren Weiterentwicklung und die Funk-

tion von Prüfungen als Teil des Lernprozesses thematisiert, die in der Terminologie von Flechsig den didaktischen Funktionen von Prüfungen zuzuordnen sind. Diese Analyse dürfte auch für deutsche Prüfungsordnungen kaum anders ausfallen. Im Mittelpunkt stehen Lernstandserhebungen, die Messung von angeeignetem Wissen und erst sekundär – wenn überhaupt – wird auf die Bedeutung von Prüfungen als Rückmeldung für die Lernenden und deren Lernprozess verwiesen.

Der prägende Einfluss von Prüfungen auf den Lernprozess selbst wird dabei in der hochschulpolitischen Diskussion in Deutschland ebenso ausgeblendet, wie BOUD (2007) das für den angloamerikanischen Hochschulraum kritisiert. Solange die Messung von Lernergebnissen im Vordergrund steht, überrascht es nicht, dass auch die Hochschulforschung sich vorrangig der Sicherung von Validität und der Entwicklung reliabler Testverfahren verschreibt oder die Objektivität von Hochschulprüfungen diskutiert, deren Ergebnisse sich keineswegs immer nur an den Leistungen der Geprüften festmacht (DAXNER, 1999; MÜLLER-BENEDICT, 2005).

Dagegen liegen kaum Untersuchungen zur Wirkung von Prüfungsformen auf den Lernprozess selbst vor, weder auf das der jeweiligen Prüfung zeitlich vorangehende Lernen noch auf die daran anknüpfenden Lernprozesse. Letztendlich beeinflusst aber kaum ein anderes hochschuldidaktisches Element den Lernprozess so nachhaltig (vgl. BOUD 2007), wie das Format der daran anschließenden Prüfungen, da Studierende ihren Lernprozess eben auf das erfolgreiche Absolvieren dieser Prüfungen hin ausrichten (vgl. NORTON 2004, p. 688). Die Erfahrungen mit den dort gestellten Anforderungen, mit Erfolgen und Misserfolgen wirken wiederum auf anschließende Lernprozesse. Weitgehend ungenutzt bleibt dabei das Potenzial von Prüfungen, den Lernprozess zu unterstützen und selbst konstruktiver Teil des Lernprozesses zu werden.

Natürlich können auch tradierte Prüfungsformen an Hochschulen – wie Klausuren, Hausarbeiten oder mündliche Prüfungen – für sich einen Lerneffekt erzielen, indem sie eine Explikation und Strukturierung von Wissen sowie dessen präzise Darstellung erfordern. Darüber hinaus unterstützen sie den Lernprozess als Quelle extrinsischer Motivation bzw. bewegen Studierende über deren Wunsch nach möglichst guten Studienleistungen („achievement motivation“ vgl. NEWSTEAD & HOSKINS, 2001) zu Lerneranstrengungen. Allerdings läuft der eigentlich gewünschte Lernprozess, die intensive und kritische Auseinandersetzung mit den Inhalten, dabei Gefahr in den Hintergrund zu rücken und auf Seiten der Lernenden an dessen Stelle das Ziel einer möglichst effizienten Vorbereitung auf die jeweilige Prüfungsform zu treten (NORTON, 2004).

Um ein Auseinanderdriften von Lernzielen und Prüfungsanforderungen zu verhindern, ist eine erweiterte Sichtweise auf Prüfungen notwendig, die diese als wesentlichen Teil des Lernprozesses selbst begreift. Die logische Konsequenz aus dieser Perspektive ist, dass Prüfungen – ebenso wie die eingesetzten Lehr- und Lernmethoden – an den jeweiligen Lernzielen ausgerichtet werden müssen. Dabei

können – wenn die Vermittlung deklarativen Wissens<sup>5</sup> im Vordergrund steht – eine lehrerzentrierte Unterrichtsgestaltung und abschließende Wissenstests legitime und erfolgreiche Strategien sein. Die Vermittlung anwendungsorientierten prozeduralen Wissens dagegen bedarf nicht nur anderer Vermittlungs- sondern insbesondere auch anderer Prüfungsformen.

Dies gilt umso mehr, wenn in einem konstruktivistischen Lehr-Lern-Verständnis die Selbststeuerung der Lernenden gefördert werden soll, eine Zielsetzung die gerade mit den bislang dominanten Prüfungsformen kaum korrespondiert. Über die Passung von Lernzielen und Prüfungsformaten in einzelnen Veranstaltungen hinaus verweisen MAYRBERGER & MERKT (2008) darauf, dass eine Weiterentwicklung der Lernkultur an Hochschulen auch auf eine Veränderung der Prüfungskultur angewiesen ist und umgekehrt.

Es ist von daher auch kein Zufall, dass gerade in den Hochzeiten hochschuldidaktischer Neuerungen in den 1960er und 1970er Jahren auch das Prüfungswesen auf die Agenda gesetzt wurde. Die Veränderungen und Innovationen in diesem Bereich blieben allerdings deutlich hinter anderen hochschuldidaktischen Handlungsfeldern zurück (BÜLOW-SCHRAMM, 2008).

### 3.1 Prüfen in lehrerzentrierten Lernarrangements

Auch in den so genannten „weichen Fächern“, zu welchen die Sozial- und Geisteswissenschaften zählen, ist die Vermittlung von schwerpunktmäßig deklarativem Wissen notwendiger Bestandteil des Studiums und nimmt insbesondere zu Beginn der Studienzeit Raum ein. Zur Vermittlung grundlegender Begriffe, Konzepte und Ideen sind – nicht nur aus Gründen der Effizienz und der begrenzten Ressourcen – lehrerzentrierte Lehr-Lern-Arrangements weithin etabliert und auch berechtigt. Selbst Vertreter eines gemäßigten Konstruktivismus bestreiten nicht die didaktische Notwendigkeit von Vorlesungen oder anderen Formen lehrerzentrierter Veranstaltungen. Das in diesen Lehrformen von den Studierenden erworbene Wissen ist zunächst allerdings ein rein theoretisches und losgelöst von konkreten Anwendungskontexten. Dies entspricht dem formulierten Ziel der Vermittlung deklarativen Grundwissens, das dann auch im Zentrum einer Überprüfung der Lernzielerreichung stehen muss.

Vor diesem Hintergrund sind Klausuren und die, mit zunehmender Prüfungsdichte in den gestuften Studiengängen (BÜLOW-SCHRAMM, 2008) stärker in den Blick rückenden, Multiple-Choice-Prüfungen als ein adäquates Mittel der Überprüfung von Lernerfolgen anzusehen. Voraussetzung ist, dass gerade diese hochstandardisierte Form der Wissensüberprüfung strikten Qualitätskriterien genügt. Auf Ebene der einzelnen Items können klassische testtheoretische Kriterien wie Schwierigkeit, Trennschärfe, Homogenität formuliert werden, auf Testebene sind Objektivität, Reliabilität und Validität sicherzustellen (BRAUNS & SCHUBERT, 2008).

---

<sup>5</sup> Zu Differenzierung verschiedener Wissensformen siehe REINMANN-ROTHMEIER & MANDL (2001).

Multiple-Choice-Prüfungen können somit zwar den Korrekturaufwand insbesondere bei Einsatz von Datenscannern erheblich verkürzen, sind aber mit einem erheblichen Aufwand bei ihrer Erstellung und Erprobung verbunden, der nur durch ein hohes Maß an Professionalität in der Testentwicklung sichergestellt werden kann (HEES, HERMANN & HUSON, 2008). Diese Prüfungsform lässt sich auch in digitale Medien übertragen und online umsetzen (HEIDBRINK, 2003; SCHAFFERT, 2005), wenngleich hier hinsichtlich der Selektionsfunktion von Prüfungen Probleme der Kontrollierbarkeit der Bearbeitungssituation noch einen relevanten Einwand darstellen.

Steht dagegen statt der Selektions- und Legitimationsfunktion von Prüfungen deren Beitrag für den Lernprozess im Mittelpunkt, so ergeben sich bislang kaum genützte Möglichkeiten aus dem Einsatz standardisierter Leistungstests. So geben Selbsttests, die sich sehr gut online abbilden lassen (vgl. MAYRBERGER & MERKT 2008, S. 148), den Lernenden die Möglichkeit, den eigenen Lernfortschritt eigenständig zu überprüfen und eine Rückmeldung zum eigenen Wissensstand zu erhalten, die nicht an Sanktionen gekoppelt ist.

Dieser Grundgedanke liegt auch dem Modell des Progress-Testing zugrunde, wie es von NOUNS & BRAUNS (2008) vorgestellt wird. Hier können Studierende der Medizin in regelmäßigen Abständen einen Test absolvieren, der Fragen zu den Inhalten des gesamten Studiums erfasst und anhand des Anteils der richtig beantworteten Items ihren eigenen Wissenszuwachs von Semester zu Semester überprüfen. Die Autoren schreiben dieser formativen Form der Lernstandserhebung u.a. ein hohes motivationales Potenzial zu. Insbesondere aber lässt das Testverfahren aufgrund der Breite des abgefragten Stoffes keine gezielte Vorbereitung zu und misst so „spontan abrufbares“ (ebd., S. 115) anstatt kurzfristig angeeignetem Wissen.

Fragt man allerdings nach dem Beitrag der Prüfungen zum Lernprozess selbst, also zum Lerngewinn durch die jeweiligen Prüfungen, so ist dieser – zumindest bei stark standardisierten Tests – jenseits motivationaler Aspekte eher gering einzuschätzen. Die Prüfung selbst ist in den beschriebenen Settings nicht als Lernprozess, sondern lediglich als Lernstandserhebung konzipiert. Bei Prüfungsformen, die insbesondere zur Erfassung anwendungsorientierter und situationsbezogener Kompetenzen eingesetzt werden, können dagegen auch durch in der Prüfungsform selbst Lernprozesse angelegt sein.

### **3.2 Prüfen in lernerzentrierten Lernarrangements**

Der Aufbau anwendbaren und auf verschiedene Problemstellungen transferierbaren Wissens ist auf entsprechende didaktische Szenarien verwiesen, wie sie seit den hochschuldidaktischen Diskussionen der 1960er und 1970er Jahren diskutiert wurden. Die damals vordergründigen Konzepte des forschenden und entdeckenden Lernens (HUBER, 2001; JENKINS, BREEN, LINDSAY & BREW, 2003), die Ideen des problemorientierten Lernens (vgl. KOPP, BALK & MANDL, 2002) und die insbesondere durch die lernpsychologischen Forschungen der 1980er und 1990er gespeisten Konzepte des situierten Lernens (vgl. MANDL, GRUBER & RENKL, 1997) rücken die Lernenden in den Mittelpunkt von Lehr-Lern-Settings

und werden daher auch als lernerzentrierte Zugänge charakterisiert (vgl. auch SCHMIDT & TIPPELT, 2005).

In diesem Kontext wird häufig auf einen Perspektivenwechsel von einer Orientierung am Lehren hin zur Orientierung am Lernen (KNOLL, 1998, S. 27) gesprochen. Der Einbezug der Teilnehmenden in die Gestaltung der Lehre, eine klare Strukturierung von Lehr-Lern-Situationen, die Berücksichtigung von Gruppenprozessen, Akzeptanz gegenüber Widerständen und Ermutigung zu deren Artikulation sowie die angemessene Gestaltung der Rahmenbedingungen durch die Seminarleitung sind Kernelemente dieses neuen Paradigmas (vgl. KNOLL, 1998; SCHMIDT & TIPPELT, 2007).

Das in diesen Lernarrangements aufgebaute Wissen, das v.a. auf die kompetente Bearbeitung authentischer Problemstellungen ausgerichtet ist, entzieht sich den vorangehend skizzierten standardisierten Wissenstests. Wenn die Vermittlung anwendbaren prozeduralen Wissens im Vordergrund steht, so müssen neben den Lernformen auch die Prüfungsformen auf dieses Ziel hin ausgerichtet sein. Denkbar sind z.B. die Bearbeitung von authentischen Fällen (vgl. SCHULMEISTER, 2002), praxisnahen offenen Aufgabenstellungen (vgl. HUBER, 2001), Simulationen (FREIMANN & SCHWADERLAPP, 1994) sowie die Erstellung von Portfolios (siehe überblicksartig BRINKER, 2007).

Die Bewertung dieser offenen Prüfungsformen wurde wiederholt kritisch diskutiert, da die Interrater-Reliabilität der Bewertungen von Lösungsansätzen, Essays oder Portfolios insbesondere dann unbefriedigend ausfällt, wenn die Bewertungskriterien nicht ausreichend oder nur unscharf definiert sind (BAUME, YORKE & COFFEY, 2004). Klar definierte Kriterien für die Bewertung der jeweiligen Leistungen und Produkte sind ein wichtiger Schritt zur Erhöhung der Reliabilität offener Prüfungsverfahren.

Die Kommunikation dieser Kriterien gegenüber den Lernenden ist für den Lernprozess selbst aber noch wesentlich wichtiger. Durch klar formulierte Anforderungen, die mit der jeweiligen Aufgabe verbunden sind, erhalten Studierende im Sinne eines Feedforward zusätzliche Anhaltspunkte für relevante Qualitätskriterien im Kontext der Aufgabenstellung und Rückmeldungen werden von einem schwer nachvollziehbaren Urteil zu einem hilfreichen Feedback. Die Motivation Studierender hängt von regelmäßigen, fairen und validen Rückmeldungen ebenso ab, wie von den persönlichen Lerninteressen bzw. Lernbarrieren, der wahrgenommenen Relevanz der Inhalte und der Lernerkontrolle (JENKINS et al., 2003).

Das Prinzip der Lernerkontrolle – also die Verlagerung der Steuerung des Lernprozesses weg von den Lehrenden hin zu den Lernenden – bezieht sich nicht nur auf die Wahl der Lernwege und -methoden sowie des Lerntempos, sondern kann ebenso auf den Prüfungsprozess bezogen werden. Nicht nur in die Wahl der Prüfungsinhalte und -formen können die Lernenden im Sinne eines Empowerment-Ansatzes (LEACH, NEUTZE & ZEPKE, 2001) einbezogen werden, sondern auch in die Bewertung der erbrachten Leistungen und die sich daraus ableitenden Rückmeldungen (MacLELLAN, 2001). Vor diesem Hintergrund werden unter dem Stichwort des „peer-assessment“ Verfahren diskutiert und erprobt, die Studierende in die Bewertung von Prüfungsleistungen und vor allem in die Formulierung von

Rückmeldungen einbeziehen und damit auch die Verantwortungsübernahme der Lernenden für den eigenen Lernprozess fördern (vgl. ebd.).

Die Bedeutung und die positiven Effekte kooperativen Lernens sind weithin bekannt und verschiedene Formen des kooperativen Lernens in der Hochschullehre etabliert. FALCHIKOV (2007) sieht den Einbezug der Mitlernenden in die Bewertung von Lernergebnissen und deren Rückmeldung als eine logische Konsequenz des kooperativen Lernens und sogar als herausragend hinsichtlich ihrer Lernwirksamkeit an. Die Delegation der Bewertung von Ergebnissen des Lernprozesses an die Lernenden selbst und die an diese gestellte Aufgabe, Rückmeldungen an ihre Kommilitonen zu formulieren, trägt einerseits zu höheren Lernerfolgen bei (vgl. BANGERT, 1995; GREER, 2001) und fördert andererseits Reflexionsfähigkeit, Autonomie, Selbstwirksamkeit und Verantwortungsübernahme (FALCHIKOV, 2007). Damit avanciert peer-assessment selbst zum wesentlichen Bestandteil des Lernprozesses und erschöpft sich längst nicht in den motivationalen Effekten, die aus der Partizipation Lernender im Bewertungsprozess resultieren können.

Probleme, die aus peer-assessment entstehen können, haben im Wesentlichen zwei Ursachen. Zum einen führen unklare oder gar nicht festgelegte Bewertungskriterien zu fehlender Akzeptanz und zu wenig validen Bewertungen (PAPINCZAK, YOUNG & GROVES, 2007), zum anderen stößt die ungewohnte Form der Leistungsbewertung anfangs sowohl bei Studierenden als auch bei Lehrenden auf eine gewisse Skepsis (ebd.), die es ernst zu nehmen aber auch zu überwinden gilt (WALKER, 2001).

FALCHIKOV & GOLDFISH (2000) konnten in einer Metaanalyse von 56 Studien zeigen, dass die Bewertung der Peers vor allem dann mit der Bewertung der Lehrenden korrespondiert, wenn klare Bewertungskriterien und die Studierenden bereits Erfahrung mit peer-assessment haben. Zudem bedarf es einer Vorbereitung der Studierenden auf die mit der Übernahme von Verantwortung für die Bewertung der Leistungen ihrer Mitlernenden und weiterer Unterstützungsangebote, wie z.B. Lernverträge (vgl. WALKER 2001). Darüber hinaus wird vielfach auch eine Kombination von peer-assessment und self-assessment vorgeschlagen (z.B. FALCHIKOV 2007; PAPINCZAK, YOUNG & GROVES 2007; WALKER 2001), auch um in manchen Studien berichteten Schwächen von peer-assessments hinsichtlich Genauigkeit und Validität (vgl. LI, 2001) zu kompensieren.

Einer der stabilsten Befunde aus der Prüfungsforschung ist die Bedeutung von Feedback als wichtigster Bestandteil von Prüfungen für den weiteren Lernprozess. Voraussetzung für ein effektives Feedback sind nach TARAS (2002, S. 505), dass die Lernenden sich der Anforderungen bewusst sind, diese mit ihren eigenen Leistungen abgleichen und schließlich die Weiterentwicklung der eigenen Leistungen in Richtung dieser Anforderungen vorantreiben. Diese Ziele lassen sich dann besonders gut realisieren, wenn auch der Prüfungs- und Bewertungsprozess in der Verantwortung der Lernenden liegt (ebd.).

Gleichzeitig eignet sich self-assessment – verstanden als „the involvement of students in making judgements of their learning“ (TAN, 2008, S. 16) – auch zur Sichtbarmachung von Lernfortschritten für die Lernenden selbst sowie für die Lehrenden und die Institution (MOORE & HUNTER, 1993). Eines der wichtigsten

Argumente für self-assessments ist aber die Förderung der Fähigkeit zur Selbstevaluation und zur realistischen Einschätzung eigener Lernfortschritte, wie sie für das lebenslange Lernen als elementar angesehen werden (TAN, 2007; 2008). Dieses Ziel von self-assessment wird dann in besonderem Maße erreicht, wenn nicht nur die individuellen Lernleistungen, sondern auch die gegebenen Rahmenbedingungen, die Lerninhalte des Studienprogramm und die eingesetzten Lehr-Lernmethoden im Sinne einer umfassenden formativen Evaluation Gegenstand der Bewertung durch die Lernenden sind. In diesem Kontext würde – so TAN (2007) – dann auch das Kriterium der Reliabilität von Bewertungen zugunsten der Reflexivität und der daraus hervorgehenden Anregungen für die weitere Gestaltung von Lehre und Lernen in den Hintergrund treten.

#### **4 Praktische Realisierung der Entwicklungsfunktion: Prüfungen und Feedback**

Auch wenn im Sinne der vorgestellten Kategorisierung von FLECHSIG (1976) neben der Herrschafts- und Sozialisierungsfunktion, der Rekrutierungs- bzw. Selektionsfunktion auch die Didaktische Funktion von Prüfungen allgemein anerkannt ist (vgl. BÜLOW-SCHRAMM & GIPSER, 1994; MÜLLER & BAYER, 2007; REIS & RUSCHIN, 2008), finden sich bislang nur wenige Ansätze, die sich explizit den Entwicklungsmöglichkeiten und Lernchancen durch Prüfungen widmen. Der überwiegende Schwerpunkt aktueller Publikationen liegt eindeutig bei der Frage, wie eine Kompetenzorientierung bei universitären Prüfungen realisiert werden kann. Auch die Steigerung von Gütekriterien (Validität, Reliabilität, Objektivität), Prüfungsmethodik oder die Standardisierung von Testsituationen werden viel thematisiert (vgl. zusammenfassend DANY et al., 2008; TINNEFELD, 2002).

Wenngleich diese Aspekte der Prüfungsvorbereitung wesentlich erscheinen, um universitäre Lernstandserhebungen (noch) gerechter zu gestalten, fällt doch auf, dass die Nachbereitung von Prüfungen im Sinne einer Orientierung an studentischen Entwicklungsfeldern bislang weitgehend ausgeblieben ist. Exemplarisch illustriert werden kann dies an einer Veröffentlichung von MÜLLER & BAYER (2007) mit dem Titel 'Prüfungen: Vorbereitung – Durchführung – Bewertung'. Auch anhand des Vorschlags TINNEFELDS (2002, S. 6) zur Etablierung einer eigenen wissenschaftlichen Disziplin 'Prüfungsdidaktik' wird dies deutlich: die Prüfungsdidaktik umfasst demnach „die Prüfungen selbst und deren Vorbereitung“, behandelt jedoch nicht die Rückmeldung der Resultate an die Studierenden.

Erkennt man die Entwicklungsfunktion als eine wesentliche von Prüfungen an, stellt sich die Frage, wie diese realisiert werden kann. Eine Möglichkeit wäre die Integration von Aspekten der Prüfungskonzeption in Lehrveranstaltungen, um Studierende an der Konstruktion von Multiple-Choice-Items zu beteiligen (FELLENZ, 2004). Curriculumsunabhängig stellt die Rückmeldung der Prüfungsergebnisse eine, wenn nicht sogar die (vgl. HATTIE, 1987) wesentliche Möglichkeit dar, Entwicklungs- und Lernprozesse von Studierenden zu initiieren oder zu begleiten. Im Folgenden soll ein Modell vorgestellt werden, das die Entwicklungsfunktion durch einen Feedbackprozess im Rahmen einer lernerorientierten Veranstaltung exemplarisch anhand einer studentischen Stundengestaltung verdeutlicht (vgl. Abbildung 1).

Als Folie zur Notengebung sollte zu Beginn ein Bewertungsbogen erstellt werden, um die Beobachtungen des Prüfers zu systematisieren und „ad-hoc-Verhalten“ (TINNEFELD, 2002, S. 27) zu reduzieren. Wesentlich ist hierbei, zunächst relevante Beobachtungskriterien zu identifizieren (Schritt 0). Diese können einerseits vom Dozenten aus den vorab festgelegten Zielen der Lehrveranstaltung abgeleitet und in Form von Kompetenzdimensionen über konkrete Handlungen (z.B. Moderationsfähigkeit als Form der sozialen Kompetenz) operationalisiert werden (HUBER, 2008).

Andererseits können jedoch auch die Studierenden in den Entstehungsprozess des Beobachtungsrasters einbezogen werden, was z.B. im Kontext einer Diskussion über verschiedene Kompetenzdimensionen (personal, sozial, methodisch, aktional, vgl. z.B. ERPENBECK & ROSENSTIEL, 2007) erfolgen kann. In einem Folgeschritt gilt es Bewertungsstandards für die einzelnen Auswertungs- bzw. Beobachtungskategorien aufzustellen (Schritt 0.1) und somit inhaltliche Zielvorgaben, die als Benotungsgrundlage dienen, zu formulieren (z.B. aktives Eingehen auf Fragen aus dem Plenum als Standard für Moderationsfähigkeit). Auch hier können die Vorgaben wiederum von dem Dozenten selbst kommen oder aber im Austausch mit den Studierenden erstellt werden.

Die begriffliche Trennung zwischen Beobachtung (Schritt 0) und Bewertung (Schritt 0.1) erfolgt an dieser Stelle nicht zufällig, sondern soll verdeutlichen, dass eine Notengebung nur anhand einer Orientierung an gesetzten Standards oder Zielen erfolgen kann<sup>6</sup>. Ein Beispiel für einen möglichen Bewertungsbogen findet sich am Ende dieses Beitrags (S. 45) – es sei darauf hingewiesen, dass an dieser Stelle nur exemplarische Standards angegeben sind, die in der konkreten Seminarpraxis weiterer Erläuterungen bedürfen.

Es folgt ein wesentlicher, in der Praxis jedoch oft nur unzureichend umgesetzter Schritt, nämlich die Weitergabe der vorab festgelegten Beobachtungskriterien und Bewertungsstandards an die Studierenden (Schritt 1). Aus Sicht der Rückmeldeforschung kann die Präsentation von Standards als Feedforward verstanden werden (BOGART, 1980; HUDSPETH, 1993; MÜLLER, 2008, S. 29). Das heißt, die vorgegebenen Standards stellen hier im Sinne der Zielsetzungstheorie (LOCKE & LATHAM, 1990) verhaltenslenkende Orientierungspunkte dar, an denen die Studierenden ihre Handlungen im Seminarkontext ausrichten können.

Es folgt mit einem Referat oder einer Stundengestaltung die Prüfungssituation selbst (Schritt 2) anhand derer das bislang vernachlässigte Spannungsfeld zwischen Beobachtung und Bewertung thematisiert werden soll. Vor allem in der kommunikationswissenschaftlich fundierten und handlungspraktisch orientierten Feedbackforschung wird immer wieder auf die Notwendigkeit verwiesen, bei Rückmeldungen zwischen Beobachtung und Bewertung zu trennen (SCHULZ VON THUN, 2004, S. 72 ff.; STERZENBACH, 2004).

---

<sup>6</sup> Da diese beiden Prozessschritte durchaus schon vor dem Beginn einer Seminarreihe erfolgen können, z.B. in dem Studierende in einem früheren Semester an der Erstellung des Bewertungsbogens beteiligt waren oder dieser aber vom Dozenten eigenhändig konzipiert wird, wurden sie mit 0 bzw. 0.1 nummeriert.

Ersteres kann als wertneutraler Hinweis auf potentielle Entwicklungsfelder verstanden werden, wohingegen letzteres den Zensurcharakter einer Situation betont wobei subjektive Interpretationen des Feedbackgebers an Gewicht gewinnen. Dies spiegelt sich auch in der Diskussion um die Positionierung von Feedback im Spannungsfeld zwischen Entwicklungs- und Kontrollfunktion (vgl. BAHNERS, 2005; HORNSTEIN, 2000) wider. In unterschiedlichsten Forschungsfeldern besteht inzwischen weitgehend Einigkeit dahingehend, dass eine „Realisierung beider Funktionen mit nur einer Methode nicht möglich“ ist (MÜLLER, 2008, S. 6; KNIGHT, 2001; PREISER, 1992; SCHNEEWIND, 2007; SCHULER, 2004).

Da im Kontext der neuen BA-Studiengänge jede studentische Leistungserbringung relevant für die Abschlussnote ist (im Gegensatz zu Referaten oder Stunden-gestaltungen im Rahmen vieler Magister- und Diplomstudiengänge), finden sich Lehrende hier in einem Rollenkonflikt wieder: einerseits müssen sie studentische Leistung per Zensur beurteilen, andererseits sollen und wollen sie fördernd die Entwicklung der Studierenden begleiten (vgl. IRONS, 2008, S. 6 ff.).

Auf Seiten der Studierenden erscheint diese ambigue Situation problematisch, da sie (verständlicherweise) vor allem unter dem Primat der Notenvergabe wahrgenommen wird. Somit erlangt die Rückmeldung eine nicht zu unterschätzende selbstwertbezogene Bedeutung, was sich in emotionalen Reaktionen sowie Mechanismen der Selbst-Verteidigung, äußern kann, wie z.B. externalen Attributions-mustern oder Ablehnung des Feedbackgebers (BRETT & ATWATER, 2001; HIGGINS, HARTLEY & SKELTON, 2001; ILGEN, FISHER & TAYLOR, 1979; KLUGER & DeNISI, 1996; SMITHER, LARSEN & RICHMOND, 2005).

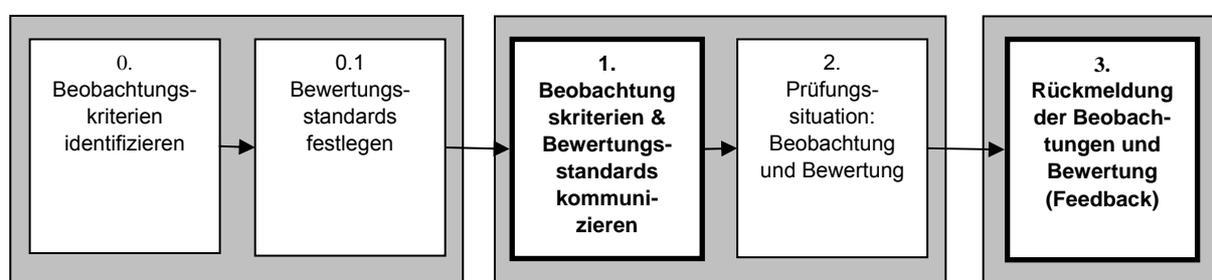


Abbildung 1: Erweiterter Feedback-Zyklus

Auflösen lässt sich dieses problematische Setting zwischen summativer und formativer Prüfungsfunktion nur durch strukturelle Veränderungen, zum Beispiel durch die Einführung nicht-notenrelevanter Lernstandserhebungen (NOUNS & BRAUNS, 2008) oder aber eher situativ durch eine personale Trennung von „Bewerter“ und „Förderer“. So vielversprechend der erste Vorschlag erscheint, so idealistisch klingt der Zweite angesichts der angespannten Personalsituation an vielen deutschen Hochschulen. Zumindest eine Reduzierung dieser Spannungssituation, die sich aus dem skizzierten Funktions-, Ziel- und Rollenkonflikt ergibt, kann jedoch durch Berücksichtigung einiger Gesichtspunkte im Rahmen der Prüfungs- und Rückmeldesituation erreicht werden.

Vor der eigentlichen Stundengestaltung bietet es sich für den Dozenten an hervorzuheben, dass die Seminarsitzung für die Studierenden auch als Experimentierfeld für didaktisch reflektierten Methodeneinsatz fungiert (vgl. RACE,1994). Dies bedeutet, dass die Bewertung nicht explizit an eine perfekt angewendete bzw. ablaufende Methode (z.B. Formen der Gruppendiskussion wie Fishbowl etc.) gekoppelt ist, sondern vielmehr an die daran geknüpften didaktischen Überlegungen und Lernziele bzgl. der Seminarteilnehmer.

Für die Rückmeldesituation sollte von Seiten des Prüfers grundlegend eine möglichst entspannte Atmosphäre ohne Zeitdruck geschaffen werden. Hierzu gehört auch, vorab einen Überblick über den Ablauf des Feedbackgesprächs zu geben – dies ist umso bedeutsamer bei Formen der peer-Rückmeldung mit mehreren Beteiligten. Zur Förderung selbstdiagnostischer Fähigkeiten empfiehlt es sich, die Studierenden zu Beginn kurz nach ihrem eigenen Eindruck zu befragen. Hier bietet es sich an, nach Fragen zur Vorbereitung der Prüfungssituation (Zusammenarbeiten im Team, Workload) der Struktur des Beobachtungsbogens zu folgen.

Die Rückmeldung selbst sollte sowohl Beobachtungen und Bewertungen beinhalten, wobei es günstig erscheint in Bezug auf Letzteres den subjektiven Wahrnehmungsgehalt auszudrücken (z.B. über ICH-Botschaften). Zum Abschluss der Rückmeldung sollte den Studierenden Raum für Rückfragen, Anmerkungen sowie einem Feedback an den Dozenten eingeräumt werden.

## 5 Fazit

Legt man Flechsigs Analyse von Prüfungsfunktionen zugrunde, so liegt der Schwerpunkt der hier dargestellten Ansätze klar auf den didaktischen Funktionen von Prüfungen, geht aber gleichzeitig darüber hinaus. Prüfungen sind auch oder sogar in erster Linie als Teile des Lernprozesses zu sehen und stärker als das bislang in Hochschulen etabliert ist, auf ihren Beitrag zum Lernprozess selbst hin zu untersuchen und entsprechend zu gestalten. Zentral ist dabei einerseits die Ausrichtung von Lehr-Lernmethoden und Prüfungen an den jeweiligen Lernzielen, andererseits den Fokus auf die Entwicklungsfunktion von Prüfungen zu richten.

Durch den Einbezug der Lernenden in die Gestaltung von Prüfungen, in deren Auswertung und insbesondere durch Rückmeldungen, die weit über die bloße Leistungsbewertung hinausgehen, können Prüfungen mehr als ein vom eigentlichen Lernprozess abgekoppelter ritueller Akt sein, sondern selbst als ein wesentlicher Teil des Lernens eine zusätzliche Qualität erlangen. Entsprechende Instrumente liegen aus der hochschuldidaktischen Forschung vor und sind vielfältig erprobt.

Für eine breite Etablierung innovativer Prüfungsverfahren und eine stärkere Betonung von Prüfungen als Teil des Lernens, wird die Übernahme dieser Methoden alleine aber nicht ausreichen, sondern eine veränderte Sicht auf das Prüfen von Seiten der Institutionen wie auch der Lehrenden und Lernenden notwendig sein. In diesem Sinne kann MAYRBERGER & MERKT (2008) zugestimmt werden, wenn sie neben einer veränderten Lernkultur auch eine neue Prüfungskultur an Hochschulen fordern.

## 6 Literaturverzeichnis

- Bahners, C.** (2005). Vorgesetztenbeurteilung mittels 360° Feedback. München: Rainer Hampp Verlag.
- Bandura, A.** (1986). Social foundations of thought and action. Englewood Cliffs: Prentice Hall.
- Bandura, A.** (1991). Social cognitive theory of self-regulation *Organizational Behavior and Human Decision Processes*, 50, 248-287.
- Bangert, A. W.** (1995). Peer assessment: an instructional strategy for effectively implementing performance-based assessments. South Dakota: University of South Dakota.
- Baume, D., Yorke, M. & Coffey, M.** (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment & Evaluation in Higher Education*, 29(4), 451 - 478.
- Berlin-Kommuniqué.** (2003). Realising the European Higher Education Area. Abgerufen von [http://www.bologna-bergen2005.no/Docs/00-Main\\_doc/030919Berlin\\_Communique.PDF](http://www.bologna-bergen2005.no/Docs/00-Main_doc/030919Berlin_Communique.PDF) am 16.02.2009.
- Bogart, D.** (1980). Feedback, feedforward, and feedwithin: strategic information in systems. *Behavioral Science*, 25(4), 237-249.
- Boud, D.** (2007). Reframing assessment as if learning were important. In D. Boud & N. Falchikov (Hrsg.), *Rethinking Assessment in Higher Education. Learning for the longer term* (S. 15 - 25). London, New York: Routledge.
- Braun, E., Soellner, R. & Hannover, B.** (2006). Ergebnisorientierte Lehrveranstaltungsevaluation. In Hochschulrektorenkonferenz (Hrsg.), *Qualitätsentwicklung an Hochschulen. Beiträge zur Hochschulpolitik 8/2006* (S. 60-67). Bonn.
- Brauns, K. & Schubert, S.** (2008). Qualitätssicherung von Multiple-Choice-Prüfungen. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 92 - 102). Bielefeld: Bertelsmann Verlag.
- Brett, J. & Atwater, L.** (2001). 360° Feedback: accuracy, reactions and perception of usefulness. *Journal of Applied Psychology*, 86/5, 930-942.
- Brinker, T.** (2007). Verschiedene Prüfungsformen zur Beurteilung studentischer Kompetenzen. Abgerufen von Fachhochschule Bielefeld [http://www.lehridee.de/data/doc/id\\_264/Pruefungsformen.pdf](http://www.lehridee.de/data/doc/id_264/Pruefungsformen.pdf) am 28.01.2009.
- Brunstein, J. & Heckhausen, H.** (2007). Leistungsmotivation. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (S. 143-192). Heidelberg: Springer.
- Bülow-Schramm, M.** (2008). Hochschuldidaktische Prüfungskritik revisited under Bologna-Bedingungen. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 27 - 44). Bielefeld: Bertelsmann Verlag.
- Bülow-Schramm, M. & Gipser, D.** (1994). "Wer Lehre sagt, muss auch Prüfung sagen..." *Handbuch Hochschullehre, Band F(1.1)*.
- Carver, C. & Scheier, M.** (1981). *Attention and self-regulation: A control theory to human behavior*. New York: Springer Verlag.

- Carver, C. & Scheier, M.** (1982). Control Theory: an useful conceptual framework for personality – social, clinical and health psychologie. *Psychological Bulletin*, 92, 111-135.
- Dany, S., Szczyrba, B. & Wildt, J.** (Hrsg.). (2008). Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen. Düsseldorf: Ulrich Welbers.
- Daxner, M.** (1999). Die blockierte Universität. Warum die Wissenschaft eine andere Hochschule braucht. Frankfurt/Main, New York: Campus Verlag.
- Deci, E. & Ryan, R.** (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.
- Ebel-Gabriel, C.** (2004). Qualitätssicherung an Hochschulen - Konzepte, Prozesse, Akteure. In W. Benz (Hrsg.), *Handbuch Qualität in Studium und Lehre*. B 1.1 (S. 1-24). Stuttgart: Raabe.
- Erpenbeck, J. & Rosenstiel, L. v.** (Hrsg.). (2007). *Handbuch Kompetenzmessung*. Stuttgart: Schäffer-Poeschel.
- Falchikov, N.** (2007). The place of peers in learning and assessment. In David Boud & N. Falchikov (Hrsg.), *Rethinking Assessment in Higher Education. Learning for the longer term* (S. 128 - 143). London, New York: Routledge.
- Falchikov, N. & Goldfinch, J.** (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3).
- Fellenz, M.** (2004). Using assessment to support higher level learning: the multiple choice item development assignment. *Assessment & Evaluation in Higher Education*, 29(6), 703-719.
- Flehsig, K.-H.** (1976). Prüfungen und Evaluation. In *Interdisziplinäres Zentrum der Universität Hamburg bei der Arbeitsgemeinschaft für Hochschuldidaktik* (Hrsg.), *Blickpunkt Hochschuldidaktik* (S. 303-336). Hamburg.
- Freimann, J. & Schwaderlapp, R.** (1994). Fast wie im richtigen Leben. Erfahrungen mit dem Verhaltensplanspiel „Pappenheim“. In H. Albach & P. Mertens (Hrsg.), *Hochschuldidaktik und Hochschulökonomie. Neue Konzepte und Erfahrungen* (S. 53-64). Wiesbaden: Gabler Verlag.
- Greer, L.** (2001). Does changing the method of assessment of a module improve the performance of a student? *Assessment and Evaluation in Higher Education*, 26(2127-138).
- Hattie, J. & Timperley, H.** (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Heckhausen, H.** (1989). *Motivation und Handeln*. Berlin: Springer Verlag.
- Hees, F., Hermanns, A. & Huson, A.** (2008). Prüfungserstellung mit Total Quality Management (TQM). In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 129 - 141). Bielefeld: Bertelsmann Verlag.
- Heidbrink, H.** (2003). Prüfungen online. *REPORT. Literatur- und Forschungsreport Weiterbildung*, 26(4), 44-54.

- Higgins, R., Hartley, P. & Skelton, A.** (2001). Getting the message across. the problem of communicating assessment feedback. *Teaching in higher education*, 6(2), 269-274.
- Hornstein, E., von, Rosenstiel, L., von** (2000). Ziele vereinbaren – Leistung bewerten. München: Wirtschaftsverlag Langen-Müller/Herbig.
- Huber, L.** (2001). Forschendes Lernen. In Technische Universität Darmstadt (Hrsg.), *Aktivierende Lehr- und Lernformen* (S. 11-30). Darmstadt: Technische Universität.
- Huber, L.** (2008). 'Kompetenzen' prüfen? In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf der Agenda!* (S. 12-26). Düsseldorf: Ulrich Welbers.
- Hudspeth, D.** (1993). Feedforward. In J. Dempsey, Sales, G. (Hrsg.), *Interactive Instruction and Feedback* (S. 287-301).
- Ilgen, D., Fisher, C. & Taylor, S.** (1979). Consequences of individual Feedback on Behavior in Organizations. *Journal of Applied Psychology*, 64, 349-371.
- Ingenkamp, K. & Urban, L.** (2005). *Lehrbuch der pädagogischen Diagnostik*. Weinheim: Beltz Verlag.
- Irons, A.** (2008). *Enhancing Learning through formative assessment and feedback*. London: Routledge.
- Jenkins, A., Breen, R., Lindsay, R. & Brew, A.** (2003). *Reshaping Teaching in Higher Education. Linking Teaching with Research*. London: Kogan Page Limited.
- Kluger, A. & DeNisi, A.** (1996). The effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), 254-284.
- Knight, P.** (2001). *Formative and summative, criterion & norm-referenced assessment*. Abgerufen von Queen's University Belfast <http://www.qub.ac.uk/directorates/AcademicStudentAffairs/CentreforEducationalDevelopment/Resources/Assessment/ResourcesRelevanttoUniversityAssessmentPolicy/MethodsofAssessmentFormativeandSummativeAssessment/> am 29.04.2009.
- Knoll, J.** (1998). Hochschuldidaktik der Erwachsenenbildung. In J. Knoll (Hrsg.), *Hochschuldidaktik der Erwachsenenbildung* (S. 13-47). Bad Heilbrunn: Julius Klinkhardt.
- Kopp, V., Balk, M. & Mandl, H.** (2002). Evaluation problemorientierten Lernens in der Münchner Medizinerbildung – Bewertung durch die Studierenden (1997 bis 2001). *Forschungsbericht Nr. 148*. München: Ludwig-Maximilians-Universität.
- Leach, L., Neutze, G. & Zepke, N.** (2001). Assessment and Empowerment: some critical questions. *Assessment & Evaluation in Higher Education*, 26(4), 293 - 306.
- Li, L. K. Y.** (2001). Some Refinements on Peer Assessment of Group Projects. *Assessment & Evaluation in Higher Education*, 26(1), 5 - 19.
- Locke, E. A. & Latham, G. P.** (1990). *A theory of goal setting and task performance*. Englewood Cliffs: Prentice Hall.
- Maclellan, E.** (2001). Assessment for Learning: the differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307 - 318.

- Mandl, H., Gruber, H. & Renkl, A.** (1997). Situiertes Lernen in multimedialen Lernumgebungen. In L. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 167-178). Weinheim: Beltz.
- Mayrberger, K. & Merkt, M.** (2008). eAssesment und Lernkulturen – ein Spagat zwischen Studienreformprozessen und Didaktik? In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektive auf Reformen im Prüfungswesen* (S. 142 - 158). Bielefeld: Bertelsmann Verlag.
- Moore, W. S. & Hunter, S.** (1993). Beyond "Mildly Interesting Facts": Student Self-Evaluations and Outcomes Assessment. *New Directions for Teaching and Learning*, 56, 65 - 82.
- Müller-Benedict, V.** (2005). Sind Universitätsprüfungen objektiv? Der langfristige historische Zusammenhang zwischen Erfolg in akademischen Prüfungen und Überfüllung der akademischen Berufe. *Soziologie*, 34(2), 191 - 208.
- Müller, A.** (2008). *Feedback: Wirkmechanismen und Einsatz in der Personalentwicklung*. Saarbrücken: VDM Verlag.
- Müller, F. & Bayer, C.** (2007). Prüfungen: Vorbereitung – Durchführung – Bewertung. In B. Hawelka, M. Hammerl & H. Gruber (Hrsg.), *Förderung von Kompetenzen in der Hochschullehre* (S. 223-238). Kröning: Asanger Verlag.
- Musch, J. & Bröder, A.** (1999). Ergebnisabhängig asymmetrisches Attributionsverhalten: motivationale Verzerrungen oder rationale Informationsverarbeitung. *Zeitschrift für Sozialpsychologie*, 30, 246-254.
- Newstead, S. E. & Hoskins, S.** (2001). Encouraging Student Motivation. In H. Fry, S. Ketteridge & S. Marshall (Hrsg.), *A Handbook for Teaching & Learning in Higher Education* (S. 70-82). London: Kogan Page Limited.
- Norton, L.** (2004). Using assessment criteria as learning criteria: a case study in psychology. *Assessment & Evaluation in Higher Education*, 29(6), 687 - 702.
- Nouns, Z. & Brauns, K.** (2008). Das Prinzip des Progress-Testing - Detaillierte Leistungsdarstellung und Lehrevaluation auf Basis der kontinuierlichen Wissensentwicklung der Studierenden. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 114-128). Bielefeld: Bertelsmann Verlag.
- Papinczak, T., Young, L. & Groves, M.** (2007). Peer Assessment in Problem-Based Learning: A Qualitative Study. *Advances in Health Sciences Education*, 12, 169-186.
- Podsakoff, P. & Farh, L.** (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes*, 44, 45-67.
- Preiser, S.** (1992). Beurteilung in sozialen Interaktionen. In R. Selbach, Pullig, K. (Hrsg.), *Handbuch Mitarbeiterbeurteilung* (S. 3-38). Wiesbaden: Gabler Verlag.
- Race, P.** (1994). *The open learning handbook: promoting quality in designing and delivering flexible learning*. London. Kogan Page
- Reinmann-Rothmeier G. & Mandl, H.** (2001). Wissen. In *Lexikon der Neurowissenschaften*, Band 3, S. 466. Heidelberg: Spektrum Akademischer Verlag.
- Reis, O. & Ruschin, S.** (2008). Kompetenzorientiert prüfen – Baustein eines gelungenen Paradigmenwechsels. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf der Agenda!* (S. 45-57). Düsseldorf: Ulrich Webers.

- Schaffert, S.** (2005). Online-Prüfungen zur Zertifizierung informell erworbener beruflicher Kompetenzen. from [www.bwpat.de](http://www.bwpat.de)
- Schmidt, B. & Tippelt, R.** (2005). Besser lehren – Neues von der Hochschuldidaktik. In U. Teichler & R. Tippelt (Hrsg.), Hochschullandschaft im Wandel (S. 103-114): Beiheft der Zeitschrift für Pädagogik.
- Schmidt, B. & Tippelt, R.** (2007). Aktivierendes Lehren – Problemorientiertes Lernen: Herausforderungen für Hochschule und Beratung. Zeitschrift für Beratung und Studium, 1, 13-18.
- Schneewind, J.** (2007). Wie Lehrkräfte mit Ergebnismeldungen aus Schulleistungsstudien umgehen. Abgerufen von FU Berlin <http://www.diss.fu-berlin.de/2007/252/#contents> am 14.04.2008.
- Schuler, H.** (2004). Leistungsbeurteilung - Gegenstand, Funktionen und Formen. In H. Schuler (Hrsg.), Beurteilung und Förderung beruflicher Leistung (S. 1-23). Göttingen: Hogrefe Verlag.
- Schulmeister, R.** (2002). Zur Komplexität Problemorientierten Lernens. In J. Asdonk, H. Kroeger, G. Strobl, K.-J. Tillmann & J. Wildt (Hrsg.), Bildung im Medium der Wissenschaft (S. 185-202). Weinheim: Deutscher Studien Verlag.
- Schulz von Thun, F.** (2004). Miteinander reden. 1. Störungen und Klärungen. Allgemeine Psychologie der Kommunikation. Hamburg: Rowohlt Verlag.
- Smither, J., Larsen, H. & Richmond, K.** (2005). The relationship between leaders personality and their reactions to and use of multisource feedback. Group and Organization Management, 30, 181-210.
- Sterzenbach, G.** (2004). Techniken des interkulturellen Umgangs. In J. Roth & C. Köck (Hrsg.), Interkulturelle Kompetenz. Culture Communication Skills. (S. 49 - 60). München: Bayerischer Hochschulverband.
- Tan, K.** (2007). Conceptions of self-assessment. What is needed for long-term learning? In David Boud & N. Falchikov (Hrsg.), Rethinking Assessment in Higher Education. Learning for the longer term (S. 114-127). London, New York: Routledge.
- Tan, K. H. K.** (2008). Qualitatively different ways of experiencing student self-assessment. Higher Education Research & Development, 27(1), 15 - 29.
- Taras, M.** (2002). Using Assessment for Learning and Learning from Assessment. Assessment & Evaluation in Higher Education, 27(6), 501 - 510.
- Thorndike, E.** (1913). Educational Psychology. (Vol. 1: The original nature of man). New York: Columbia University.
- Thorndike, E.** (1922). Psychologie der Erziehung. Jena: Fischer Verlag.
- Tinnefeld, T.** (2002). Prüfungsdidaktik. Zur Fundierung einer neuen wissenschaftlichen Disziplin - am Beispiel der modernen Fremdsprachen. Aachen: Shaker Verlag.
- Tremp, P. & Eugster, B.** (2006). Universitäre Bildung und Prüfungssystem – Thesen zu Leistungsnachweisen in modularisierten Studiengängen. Das Hochschulwesen, 5, 163-165.
- Walker, A.** (2001). British psychology students' perceptions of group-work and peer assessment. Psychology Learning and Teaching, 1(1), 28 - 36.

## Autoren



Andreas MÜLLER M.A. || Wissenschaftlicher Mitarbeiter ||  
Lehrstuhl für Allgemeine Pädagogik und Sozialisationsforschung ||  
Ludwig-Maximilians-Universität München ||  
Forschungsschwerpunkte: Wirkungsweise von Feedback, Folgen  
von Schulrückmeldungen aus Leistungsvergleichsstudien  
<http://www.psy.lmu.de/ape/Mitarbeiter/Andreas-Mueller.html>

[asmueller@edu.uni-muenchen.de](mailto:asmueller@edu.uni-muenchen.de)



Dr. Bernhard SCHMIDT || Wissenschaftlicher Assistent ||  
Lehrstuhl für Allgemeine Pädagogik und Bildungsforschung ||  
Ludwig-Maximilians-Universität München ||  
Forschungsschwerpunkte: Hochschuldidaktik, Bildungsverhalten  
und -interessen Älterer, Informelles Lernen  
<http://www.edu.lmu.de/apb/personen/dozent/schmidt/index.html>

[b.schmidt@lmu.de](mailto:b.schmidt@lmu.de)

**BEWERTUNGSBOGEN**

für: \_\_\_\_\_

<b>BEOBACHTUNGS-KRITERIEN</b>	<b>BEWERTUNGS-STANDARDS</b>	
<b>Vortrag</b>		
Blickkontakt	Alle Tn. werden erfasst	
Körpersprache (Gestik, Mimik, Positionierung im Raum)	Inhalt unterstützend; entspannt	
Sprechweise (Tempo, Lautstärke, Pausen)	Ruhig, deutlich	
Vortragsweise	Freies Sprechen	
Moderation	Auf Fragen wird eingegangen	
<b>Inhalt</b>		
Struktur/ Gliederung	Einleitung & Schluss (Zusammenfassung) vorhanden; 'roter Faden' gegeben	
Erklärung von Zusammenhängen	Zusammenhänge werden verdeutlicht	
Zielgruppenorientierung	Fachbegriffe werden erläutert; Vorwissen wird aufgegriffen; Inhalte dem aktuellen Wissensstand der Tn angemessen	
Beschränkung aufs Wesentliche	Inhalte aufs Wesentliche reduziert	
Beispiele	Inhalte anhand von Bsp. verdeutlicht	
Zeitmanagement	zeitl. Gewichtung passend; flexible Reaktion auf zeitl. Verzögerungen	
<b>Medien &amp; Methodik</b>		
Umgang mit Medien	Souverän	
Qualität der Visualisierungen (z.B. Folien)	Unterstützende und übersichtliche Veranschaulichung der wesentl. Inhalte	
Qualität des Skripts/ Handouts	Zusammenfassung der wesentlichen Inhalte; Quellen angegeben	
Medien & Methoden stimmig mit Inhalt	Medien & Methoden tragen zur besseren Verständlichkeit des Inhalts bei	