
Wirksamkeitsmessung von Hochschuldidaktik

Messmöglichkeiten und Anwendungsbeispiele hochschuldidaktischer Wirksamkeitsmessung

Immanuel Ulrich / Carmen Heckmann

Aktuell werden im Bund-Länder-Programm für bessere Studienbedingungen und mehr Qualität in der Lehre hochschuldidaktische Programme vieler deutscher Universitäten und Hochschulen auf 5+5 Jahre gefördert. Nach Abschluss der ersten fünf Jahre müssen die hochschuldidaktischen Maßnahmen in einer Wirksamkeitsmessung bestehen, wenn sie weitere fünf Jahre Förderung erhalten wollen. Dieser Beitrag berichtet über einige bekannte Wirksamkeitsstudien der hochschuldidaktischen Forschung, zeigt messmethodische Ansätze auf und beleuchtet Empfehlungen zur Generierung weiterer Messmethoden, die zur Evaluation der ersten fünf Jahre eingesetzt werden könnten.

Gliederung	Seite
1. Politische Ausgangslage	2
2. Stand der Forschung	3
2.1 Definition guter Hochschullehre	3
2.2 Empirie zu gutem Unterricht	5
2.3 Empirie zu guter Hochschullehre	6
3. Mögliche Formen der Wirksamkeitsmessung	9
3.1 Methodische Möglichkeiten in der Hochschuldidaktik	9
3.2 Mögliche Messebenen	11
4. Konkretisierung ausgewählter Formen der Wirksamkeitsmessung	14
4.1 Hochschuldidaktik und (Lehr-)Evaluation	14
4.2 Hochschuldidaktik und Videoaufzeichnung	16
5. Ausblick	19

1. Politische Ausgangslage

Empfehlungen des Wissenschaftsrates

Die Hochschuldidaktik kann in Deutschland auf eine lange Tradition zurückblicken (vgl. Berendt, 2005, S. 13 ff.). In den letzten Jahren ist sie verstärkt in den Fokus der Wissenschaft und Bildungspolitik geraten: so forderte der Wissenschaftsrat (2008, S. 65) in seinen dezidierten Empfehlungen zur Qualitätsverbesserung von Lehre und Studium: „Die Hochschulen benötigen nicht nur mehr, sondern auch besser qualifiziertes Lehrpersonal. Dazu bedarf es umfassender Qualifizierungsangebote für die Lehrenden (...)“. Diese Qualifizierungsmaßnahmen muss die Hochschuldidaktik leisten.

Jüngste politische Konsequenzen

Die Politik hat darauf reagiert und fördert seit 2011 mit ihrem *Bund-Länder-Programm für bessere Studienbedingungen und mehr Qualität in der Lehre* (Bundesministerium für Bildung und Forschung, 2010) universitäre Einrichtungen und Personen wie beispielsweise Mentoring-Programme, zusätzliche Lehrkräfte oder auch bestehende wie neu gegründete Zentren zur Hochschuldidaktik mit insgesamt bis zu 1,9 Milliarden Euro auf 5+5 Jahre. Die institutionelle Förderung der deutschen Hochschuldidaktik steigt hiermit massiv an. Auch eine Vielzahl von Programmen zur Förderung der Lehre wird im Rahmen dieser Förderung neu konzeptioniert, durchgeführt und evaluiert. So werden aktuell beispielsweise Mentoring, E-Learning- oder Peerteaching-Angebote an den Universitäten implementiert. Eine beispielhafte Umsetzung eines solchen Angebots ist beispielsweise das Programm „Peerteaching für Studienanfänger in Psychologie“ (Höhler, Horz, Hansen, Fabriz, Stehle, Heckmann, Ulrich & Secon, im Druck).

Wirksamkeitsmessung als notwendige Bedingung

Alle seitens des Bund-Länder-Programms geförderten Institutionen müssen sich nach fünf Jahren einer Evaluation stellen. Es werden nur jene Einrichtungen auf weitere fünf Jahre gefördert, welche neben der Umsetzung ihres Programms auch dessen positive Wirkung nachweisen können. Dieses Beispiel verdeutlicht die große Relevanz valider Formen der Wirksamkeitsmessung für verschiedene Maßnahmen, die durch öffentliche Gelder gefördert werden.

Um das Ziel der positiven Evaluierung erreichen zu können, müssen valide Formen der Wirksamkeitsmessung auch im Bereich hochschuldidaktischer Maßnahmen eingesetzt werden. Folgend soll der aktuelle Stand der Forschung in diesem Bereich aufgezeigt, ein Überblick über Verfahren, die zur Messung der Wirksamkeit zum Einsatz kommen gegeben und ein Ausblick auf mögliche Weiterentwicklungen angeregt werden.

2. Stand der Forschung

2.1 Definition guter Hochschullehre

Bevor eine Intervention konzeptioniert und durchgeführt, geschweige denn evaluiert wird, ist es nötig, den Zielzustand zu definieren – in der Hochschuldidaktik folglich das Ideal guter Hochschullehre:

Was ist gute Hochschullehre?

Hierzu liegt jedoch bislang – trotz zahlreicher Publikationen (vgl. z.B. Berendt, 2000; Koch, 2004; Ledić, Rafajac, & Kovač, 1999; Rindermann, 2010; Spiel & Gössler, 1999; Webler, 1991) – kein allgemein akzeptiertes Ideal vor, welches über Allgemeinaussagen wie beispielsweise „wir bieten gute Lehre an“ hinausgeht. Im Speziellen werden häufig unterschiedliche Aspekte guter Hochschullehre genannt. Dies liegt an der Vielzahl der Fachhintergründe und Positionen der Akteure – Studierende, Lehrende, Qualitätsbeauftragte etc. So hat ein geisteswissenschaftliches Seminar im Master hat (in Teilen) etwa andere Ideale guter Lehre als eine Einführungsvorlesung im Bachelor eines naturwissenschaftlichen Faches. Im ersteren Fall könnte beispielsweise ein ausgiebiger Diskurs den Idealtypus bilden, während im letzteren Fall womöglich die Vermittlung von grundlegendem Wissen im Fokus steht.

Zur Annäherung an eine Definition ist zunächst festzuhalten, was gute Hochschullehre *akteurs- und fachübergreifend* umfasst. Der klassische PDCA-Zyklus im Qualitätsmanagement (vgl. Deming, 1982; Wagner, 2007) mit seinen Phasen *Plan, Do, Check, Act* ist als Grundlage auf die Hochschullehre übertragbar (vgl. Abb.1):

Was umfasst gute Hochschullehre?

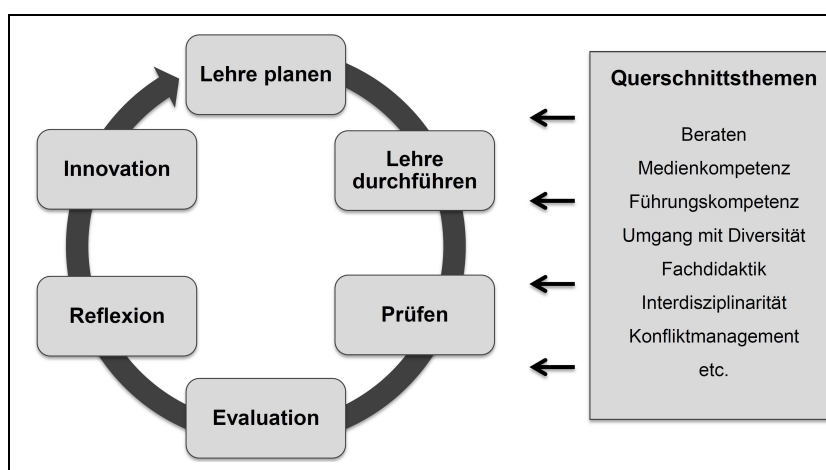


Abb. -1

Zyklus der Lehre (Hochschuldidaktik Hessen,

 2011)¹

Im abgebildeten Modell beginnt jede Lehrveranstaltung mit ihrer *Planung*, sie wird *durchgeführt* und *prüft* die Studierenden. Sie wird idealerweise auch von den Studierenden *evaluiert*. Der Lehrende *reflektiert* den Verlauf seines Kurses, auch mit Hilfe der Evaluationsergebnisse, und nimmt notwendige *Innovationen* in seiner Lehre vor. Dem zugeordnet sind Querschnittsthemen wie z.B. die *Beratung* oder der *Umgang mit heterogenen Studierendengruppen*, welche in mehreren oder gar allen Phasen relevant sind. Zu beachten ist dabei, dass die Lehre nicht nur der Vermittlung von (Fach-)Kompetenzen, sondern auch von Werten dient (z.B. Werten der jeweiligen Fachkultur, des wissenschaftlichen Arbeitens, der staatsbürgerlichen Verantwortung² etc.). Diese Ausführungen ergeben folgende fokussierte Definition:

Definition guter Hochschullehre

Gute Hochschullehre umfasst die professionelle Vermittlung von Wissen, Fähigkeiten, Kompetenzen und Werten, insbesondere im Rahmen des jeweiligen Faches. Gute Hochschullehre nutzt zur ihrer stetigen Optimierung die Standards des Qualitätsmanagements und die neuesten hochschuldidaktischen Forschungsergebnisse. Das Ziel guter Hochschullehre besteht letztendlich in der Ausbildung mündiger, kompetenter und wertgefestigter (Staats-)Bürger.

Hochschuldidaktische Umsetzung

Diese Definition muss nun in der hochschuldidaktischen Weiterbildung umgesetzt werden. Das Ziel hochschuldidaktischer Weiterbildung ist die Befähigung Lehrender zur Umsetzung guter Hochschullehre, gemäß dieser Definition. Die einzelnen Weiterbildungen können dabei wiederum anhand der PDCA-Phasen konzeptioniert werden (z.B. ein hochschuldidaktischer Workshop „Strategische Planung von Lehrveranstaltungen“ für die Phase „Plan“). Für die Umsetzung entsprechender Maßnahmen ist jedoch eine feinere Differenzierung notwendig: So bleibt zu konkretisieren, was gute Hochschullehre beinhaltet, etwa in der Phase der Durchführung, wenn der Lehrende beispielsweise insbesondere Grundwissen vermitteln möchte? Insbesondere zu diesen Teilbereichen kann die einschlägige Forschung Hinweise geben.

**Interdisziplinäre
schuldidaktik, a
alwissenschaftli**

¹ Dieser Zyklus für gute Lehre an der Hochschule wurde auf der Basis der Standards für die Lehrerbildung (Kultusministerkonferenz, 2004) in einer Arbeitsgruppe auf dem Treffen der hessischen Hochschuldidaktiker (7. November 2011, Frankfurt am Main) entwickelt.

² Die Vermittlung staatsbürgerlicher Verantwortung der Hochschullehre wird oftmals vernachlässigt, stellt aber eine Grundbedingung des humboldtschen Bildungsideals dar, auf welches sich die Hochschulen und Universitäten häufig berufen.

Die Nutzung wissenschaftlicher Erkenntnisse und Methoden im Bereich der Hochschuldidaktik steht vor einer Herausforderung, die sich aus der Unterschiedlichkeit beteiligter Disziplinen ergibt: Während in der Lehre und Hochschuldidaktik Akteure unterschiedlicher Disziplinen tätig sind, ist die hochschuldidaktische Forschung per Definitionem ein Feld der Sozialwissenschaften: die Wirksamkeitsmessung didaktischer Interventionen bei Menschen (Lehrenden wie Studierenden).

Für die notwendige und erfolgreiche Rezeption empirischer Ergebnisse für Nutzung in der praktischen Tätigkeit von Hochschuldidaktikern und nicht zuletzt für die Entwicklung eigener aussagekräftiger Forschungsansätze ist es wichtig, sowohl den aktuellen Stand einschlägiger Forschung zu kennen, wie auch gute methodische Kenntnisse zu haben. Über beides gibt dieses Kapitel Aufschluss: Zunächst wird der aktuelle Stand der Forschung dargelegt, anschließend werden mögliche Messmethoden erläutert.

2.2 Empirie zu gutem Unterricht

Es existiert eine breite, sozialwissenschaftliche, empirische Forschung zu gutem Unterricht an Schulen (vgl. z.B. Hattie, 2009; Klieme, 2006; Marzano, 1998, Walberg & Lai, 1999).

Umfassende Forschungsergebnisse

Getreu dem Grundsatz „What's good for the goose is good for the gander“ (Walberg, 1984, S. 24) werden einige Erkenntnisse der Schulforschung auf die Hochschuldidaktik übertragbar sein: So ist beispielsweise die Variable Feedback-Einsatz, für Schüler (vgl. Marzano, 1998) ebenso wie für Studierende, ein wichtiger Faktor für den Lernerfolg (vgl. Feldmann, 1989). Andere Aspekte guten Unterrichts werden auf die Hochschule adaptierbar sein: Zwar hat die Hochschule infolge des fortgeschrittenen Alters der Lernenden nicht denselben Erziehungsauftrag wie die Schulen, jedoch sind ethische Aspekte deshalb nicht belanglos: So ist den Studierenden beispielsweise der Wert des wissenschaftlich sauberen Arbeitens zu vermitteln.

Zugleich werden die relevanten Aspekte guten Unterrichts in einigen Modellen übersichtlich vereint (z.B. Helmke, 2012; Meyer, 2004). Auch diese sind beispielsweise durch die Vergleichbarkeit der Grundstrukturen von schulischer und universitärer Lehre auf die Hochschule übertragbar.

2.3 Empirie zu guter Hochschullehre

Wenige Studien

Zu guter Lehre an der Hochschule liegen bereits einige empirische Studien vor (z.B. Abrami, d'Appolonia & Rosenfield, 2007; Feldman, 1989; Murray, 1997), wenn auch in weitaus geringerem Umfang als zum Schulunterricht. Leider werden in diesen oftmals nur Einzelergebnisse berichtet (vgl. z.B. Brinko, 1993; Gold & Souvignier, 2000; Murray & Lawrence, 1980; Preiser, 2000). Was bislang fehlt, ist ein umfassendes, empirisch geprüftes Modell guter Hochschullehre, welches alle relevanten Aspekte inklusive ihrer Effektstärken umfasst. Anhand eines übersichtlichen Modells können die Wirkfaktoren guter Hochschullehre und der Stärke ihres Einflusses aufgezeigt werden. Dies gilt insbesondere für die spezifischen Effektstärken auf die gewünschten Resultate guter Hochschullehre: Wissen, Fähigkeiten, Kompetenzen, Werte oder wissenschaftliche Sozialisation. So kann durch die extrinsische Lernmotivation, stimuliert durch eine anstehende Klausur, wahrscheinlich das Wissen der Studierenden (zumindest temporär) erhöht werden, ob sie aber wissenschaftliche Werte vermittelt, ist eher fraglich. Erste Modelle, jedoch ohne empirische Einzelprüfung bzw. ausführliche theoretische Herleitung, finden sich unter anderem bei Rindermann (2010) oder Helmke und Schrader (2006).

Wie kann ein solches Modell guter Hochschullehre gestaltet sein?

- a Es sollte die relevanten Personengruppen (Lehrperson, Studierende) und Kontexte (Lehre eines spezifischen Faches an einer Hochschule) umfassen.
- b Es sollte differenzieren können zwischen den verschiedenen Ebenen von Qualität (vgl. Donabedian, 1966), welche auch im Hochschulkontext unterschieden werden (Braun, Gusy, Leidner & Hannover, 2008).
- c Es sollte alle relevanten Aspekte guter Hochschullehre umfassen, folglich alle empirischen Studien im Bereich Hochschullehre.

Abbildung 2 zeigt ein Modell guter Hochschullehre, welches diesen Kriterien gerecht wird. Aus Übersichtsgründen werden nur gewählte Unterkategorien (z.B. Motivation, Lernaktivitäten) dargestellt.

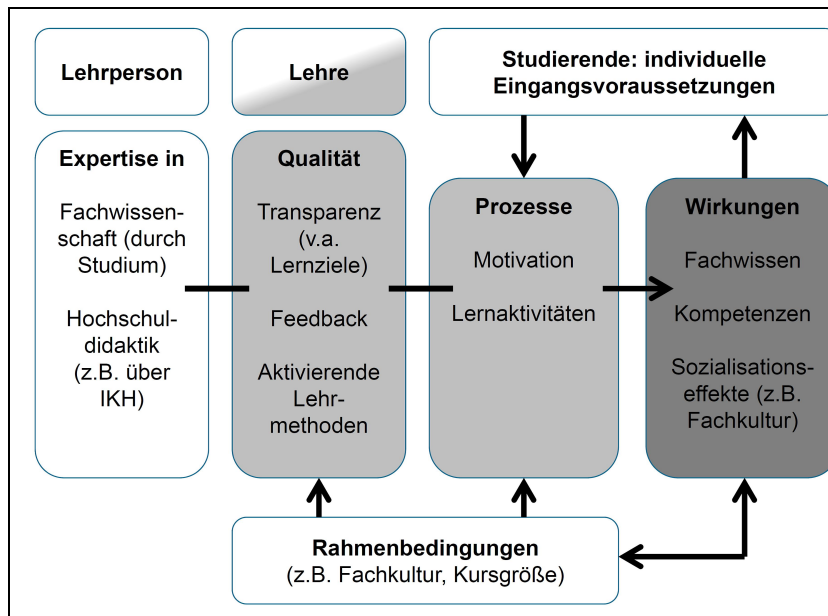


Abb. -2 Modell guter Hochschullehre (mit ausgewählten Unterkategorien), Arbeitsversion 1.4

Das Modell unterscheidet zwischen *Lehrperson* und *Studierenden*, welche in einem spezifischen Kontext, der *Lehre*, zusammentreffen. Die Lehre wirkt sich aufgrund wissenschaftlicher Fundierung (fachlich wie didaktisch!) auf die *Qualität* der Lehre aus, beeinflusst bei den Studierenden *Prozesse* wie Motivation und Lernaktivitäten, welche *Wirkung* zeigen, beispielsweise im Kompetenzaufbau. Zusätzlich beeinflusst werden diese Faktoren durch *Rahmenbedingungen* wie Fachkultur oder Studierendenzahlen.

Das Modell differenziert zwischen *Prozessen* (hellgrau hinterlegt), *Strukturen* (weiß) und *Ergebnissen* (dunkelgrau) guter Hochschullehre. **Strukturen** bedingen den Kontext guter Hochschullehre und sind nur mittelfristig durch die Beteiligten veränderbar. Sie umfassen Aspekte wie z.B. die Raumausstattung, die Anzahl, aber auch die Eingangsvoraussetzungen der Studierenden (z.B. Vorwissen, akademische Sozialisation). Besonders relevant für die Hochschuldidaktik ist der strukturelle Aspekt der „Lehrperson“, welcher dank Studium meist über eine fachliche, aber weniger eine hochschuldidaktische Expertise verfügt. Diese ist durch die Hochschuldidaktik zu gewährleisten, an der Goethe-Universität Frankfurt beispielweise über das Interdisziplinäre Kolleg Hochschuldidaktik (IKH).

Die **Prozesse** guter Hochschullehre sind unmittelbar durch die Beteiligten beeinflussbar. Sie umfassen sowohl die direkte Qualität der

Relevante Personen- gruppen und Kontexte

Ebenen der Qualität der Hochschullehre

Lehre (z.B. den Einsatz aktivierender Lehrmethoden und Feedback, vgl. Hattie, 2009), als auch die daraus resultierenden Prozesse der Studierenden (z.B. Lernaktivitäten). Welche Qualitäten die Hochschullehre umfasst (z.B. Einsatz und Transparenz der Lernziele, vgl. Hattie, 2009), hat die Hochschuldidaktik den Lehrenden zu vermitteln. Die Basis dessen muss die Empirie stellen.

Die **Ergebnisse** guter Hochschullehre umfassen das gesteigerte Wissen, die gestiegenen Kompetenzen und die veränderten Werte der Studierenden. Die Wirkungen ihrer Lehre sind für die Lehrperson nur indirekt beeinflussbar.

Nutzung empirischer Erkenntnisse

Auch bei suboptimaler Forschungslage muss sich die Hochschuldidaktik der (wenigen) empirischen Ergebnisse zu guter Hochschullehre bedienen. Wo keine Ergebnisse vorliegen, sollte (unter Vorbehalt der späteren Prüfung im Rahmen der selbst gegebenen Weiterbildungen) auf empirische Studien der Lehr-Lern-Forschung zurückgegriffen werden, die bisher meist in Schulen durchgeführt wurden. So nennt Hattie (2009) beispielsweise auf Basis seiner „Megaanalyse“ drei Aspekte als effektivste Grundlage der Qualität des Lehre: Transparenz (insbesondere der Lernziele), aktivierende Lehrmethoden und Feedbackinsatz. Folglich sind sie zentral im Modell genannt (vgl. Abb. 2).

Dieses vorgestellte Modell guter Hochschullehre ist ein Arbeitsmodell. In der Finalversion soll ein umfassendes, empirisch geprüftes Modell guter Hochschullehre resultieren, welches alle relevanten Aspekte inkl. ihrer Effektstärken umfasst (Ulrich, eingereicht).

Große Synergien zwischen hochschuldidaktischer Arbeit und Begleitforschung möglich

Neben der wissenschaftlichen Fundierung hochschuldidaktischer Arbeit ist die wissenschaftliche Begleitforschung *eigener hochschuldidaktischer Tätigkeit* erstrebenswert, da im Hochschulkontext Forschung und Lehre (leider) nicht im humboldtschen Gleichgewicht stehen, sondern der Forschung nach wie vor eine größere Wertschätzung zukommt. Daher kann es für die eigene Stellung als Hochschuldidaktiker im Hochschulkontext – gerade unter der Professorenschaft – vorteilhaft sein, nicht nur als (hochschuldidaktischer und studentischer) Lehrender, sondern auch als Forscher wahrgenommen zu werden.

Insbesondere Hochschuldidaktikern mit sozialwissenschaftlichem Hintergrund sei eine eigene wissenschaftliche Weiterqualifikation nicht nur im Lehrbereich (d.h. die eigene hochschuldidaktische Qualifikation als Trainer), sondern auch im Forschungsbereich (Promotion, publikationsbasierte Habilitation) nahegelegt: die „Durchführung“ sozialwissenschaftlicher, hochschuldidaktischer Studien erfolgt im Rahmen der eigenen Arbeit ohnehin. Eine begleitende Erhebung von Daten ist dabei relativ einfach möglich. Des Weiteren werden in der

deutschen Hochschullandschaft oftmals nur befristete „Qualifikationsstellen“ ausgeschrieben, welche eine Weiterqualifikation als Forscher vorsehen (und teils auch erzwingen).

Welche Forschungsmöglichkeiten bietet nun die Hochschuldidaktik? Das nächste Kapitel wird im Allgemeinen, das daran anschließende Kapitel im Speziellen darauf eingehen.

3. Mögliche Formen der Wirksamkeitsmessung

Die klassische Wirksamkeitsmessung (vgl. die (Lern-)Transferforschung zu Personalentwicklungsmaßnahmen aus der Wirtschaftspsychologie) orientiert sich zunächst am Subjekt der Maßnahme – hier dem didaktisch zu qualifizierenden Lehrenden. Zugleich sind alle potenziellen Nutznießer der Maßnahme Teil der erweiterten Prüfung – dies sind primär die Studierenden des jeweiligem Lehrenden, aber sekundär auch seine Kollegen: von einer Qualifikation profitiert nicht nur das jeweilige Teammitglied, sondern indirekt auch sein Team. Neben einer qualitativen Verbesserung der Lehre, wird auch die erforderliche kollegiale Unterstützung des Lehrenden in seiner Lehre weniger, der Lehrende hat durch seine didaktische Professionalisierung eine effektivere Vorbereitung seiner Lehre und somit mehr Zeit für andere (Team-)Aufgaben. Demzufolge sind (a) der Lehrende, (b) dessen Studierende wie auch (c) die Kollegen des Lehrenden übliche Zielgruppen hochschuldidaktischer Wirksamkeitsmessung.

**Studierende, Lehrende
und ihre Kollegen als
Zielgruppen.**

Zur Messung der Wirkung hochschuldidaktischer Maßnahmen empfiehlt sich des Weiteren eine Differenzierung verschiedener Ebenen, welche in Kapitel 3.2 erläutert wird. Zunächst folgt eine Übersicht zu den möglichen Messmethoden.

3.1 Methodische Möglichkeiten in der Hochschuldidaktik

Die Sozialwissenschaften bieten im Allgemeinen fünf verschiedene Messmethoden, welche in der Hochschuldidaktik eingesetzt werden können: Fragebogen, Beobachtungen, Interviews, Experimente und Inhaltsanalysen. Tabelle 1 führt zu jeder Methode in Abhängigkeit der jeweiligen Zielgruppen *eine* mögliche Umsetzung auf:

	Zielgruppen
--	--------------------

Antworten mit anschließender Quantifizierung. Neben der aufwändigen Auswertung, ähnlich wie bei Beobachtungen, ist die Gestaltung der Fragen von entscheidender Bedeutung.

Experimente sind vergleichsweise schwierig durchführbar, da einige Bedingungen (z.B. Randomisierung der Teilnehmer) in der anwendungsorientierten Hochschuldidaktik ethisch nicht vertretbar sind (z.B. liegt die Studienfachwahl alleine beim jeweiligen Individuum und kann nicht durch Wissenschaftler randomisiert werden). Aber es ist möglich, auf weniger strikte, quasi-experimentelle Messmethoden auszuweichen. Kapitel 4.1. stellt eine quasi-experimentelle Studie vor.

Inhaltsanalysen sind demgegenüber weitaus einfacher durchführbar (vgl. Mayring, 2002). Kritischer Punkt ist hier jedoch die Qualität des vorliegenden Materials. So ist beispielsweise fraglich, welche validen Schlüsse sich aus einem Veranstaltungsplan auf die Durchführung der Veranstaltung ziehen lassen.

Letztendlich sollte aber nicht der Aufwand, sondern das angestrebte Erhebungsziel die Wahl der optimalen Erhebungsmethode determinieren. Nach Möglichkeit sollte man einen Methodenmix einsetzen, welcher gleich auch zur Konstruktvalidierung genutzt werden kann (über Mixed Method Ansätze oder die Multitrait-Multimethod-Methode, vgl. Bortz & Döring, 2002, S. 202 ff.).

3.2 Mögliche Messebenen

Eine in der Lerntransferforschung übliche Differenzierung möglicher Messebenen ist die von Kirkpatrick (1960 // 1979), in welcher vier verschiedene Ebenen unterschieden werden (die Beispiele beziehen sich hier auf die Zielgruppe der Lehrenden als *Lernende* im Rahmen eines hochschuldidaktischen Weiterbildung):

1. *Reaktion*: dies umfasst z.B. die Akzeptanz, Zufriedenheit oder den Enthusiasmus der Lernenden auf das hochschuldidaktische Training, und wird meist unmittelbar im Anschluss daran erhoben.
2. *Lernen* (d.h. der subjektive und / oder objektive Lernerfolg): Zu erfassen sind Verbesserungen in den Kenntnissen und Fähigkeiten der Lernenden als Folge der Maßnahme.
3. *Verhalten* (Lerntransfer im Training selbst und v.a. in der Praxis danach): hier wird die Verhaltensänderung des Lehrenden gemessen, sowohl jene in der Lehre (z.B. bessere Didaktik) als auch jene am Arbeitsplatz (z.B. effektivere Vorbereitung auf die Lehre).

So ist die *Reaktion* der Teilnehmer unmittelbar am Trainingsende relativ leicht erfassbar. Auch lassen sich *Lerneffekte* über gute Testverfahren relativ einfach messen, wobei hier die Vergessenskurve über die Zeit mit berücksichtigt werden muss, weshalb eine „Follow-up-Messung“ nach einer gewissen Zeit ergänzt werden sollte. Des Weiteren zeigen sich im Gelernten weitaus größere Effekte als im tatsächlichen Transfer des Gelernten auf das eigene *Verhalten* (Arthur, Bennett, Edens & Bell, 2003; Alliger, Tannenbaum, Bennett, Traver & Shotland, 1997; Saks, 2002; Taylor, Russ-Eft & Chan, 2005). Denn dies setzt ein gelungenes Lerntansfermanagement voraus (vgl. Solga, 2011). Auf der Ebene der *Ergebnisse* sind Effekte noch schwerer zu messen, da neben der hochschuldidaktischen Intervention eine Vielzahl anderer Faktoren eine Rolle spielt (z. B. gruppenspezifische Effekte zwischen Lehrenden und Studierenden auf die Ergebnisse der Lehrevaluation).

Dies ist u.a. einer der Hauptgründe, weshalb nach einer Studie von Bassi, Benson und Cheney (1996) 94% der Trainings in Organisationen als Erfolgskontrolle die Ebene der Reaktion der Trainingsteilnehmer erfassen und andere Ebenen wie Lernen (34%), Verhalten (13%) und Ergebnisse (3%) relativ selten erfasst werden. Mehrheitlich wurden demnach nur Reaktionen zur Erfolgskontrolle erfasst. Hochschulen mit wissenschaftlichem Anspruch sollten jedoch ihre Evaluation über diese Ebene hinaus planen.

**Einseitige Fokussierung
auf Reaktionen der Trainingsteilnehmer**

In statistischen Kennwerten, konkret in Effektstärken³ (vgl. Bortz & Döring, 2002, S. 604), gesprochen, finden sich die größten Effekte beim Lernen ($d = 1.51$), gefolgt vom Verhalten ($d = 0.88$) der Trainierten. Die Einstellungen ($d = 0.37$) und die (Leistungs-)Ergebnisse ($d = 0.41$) der Trainierten weisen die geringsten Effekte auf (vgl. Holling & Liepmann, 2007). Die Ebene der Reaktion wurde hier nicht mit überprüft, da ihre Verknüpfung mit den anderen Ebenen als zu gering angesehen wurde.

Es empfiehlt sich folglich, (a) mehrere Zielgruppen und nicht nur Lehrende sowie (b) mehrere Ebenen und nicht nur die Reaktion zu erfassen. Die Reaktion der Lehrenden ist zwar eine wichtige Ebene, aber nur eine unter mehreren. Wie eine solche Erfassung auf mehreren Ebenen aussehen kann, wird das folgende Kapitel darlegen.

³ Effektstärke: Standardisierter Wert zum Vergleich der Wirksamkeit von Studien. Ein kleiner Effekt liegt zwischen 0.20 und 0.50, ein mittlerer Effekt zwischen 0.50 und 0.80, ein großer Effekt über 0.80 (vgl. Bortz & Döring, 2002, S. 604).

4. Konkretisierung ausgewählter Formen der Wirksamkeitsmessung

4.1 Hochschuldidaktik und (Lehr-)Evaluation

Exemplarisch soll im Folgenden eine Studie (Ulrich, 2013) vorgestellt werden, welche eine hochschuldidaktische Intervention mit einer (Lehr-)Evaluation verbindet.

Sie nutzte die Methode des *Fragebogens* bei den Zielgruppen der *Lehrenden* und *Studierenden*. Erhoben wurden *alle Ebenen* bei den Lehrenden sowie die Ebenen *Reaktion* und *Lernen* bei den Studierenden. In kleineren Maß wurden auch Beobachtungen eingesetzt, welche hier aber aufgrund ihres vergleichsweise geringen Anteils nicht berichtet werden.

Da es sich um eine Promotionsstudie handelt, sind die Fallzahlen vergleichsweise hoch und für hochschuldidaktische Alltagsarbeit in der Größe nicht immer realisierbar. Ihr messmethodisches Design ist jedoch auch bei kleineren Fallzahlen problemlos nutzbar, nur die Aussagekraft der Stichprobe sollte gesondert begründet werden.

Design der Studie

Die Studie nutzte ein sogenanntes „Mehrphasen“-Design, in welchem Daten zu fünf Messzeitpunkten (t1 bis t5) erhoben werden:

	t1	t2	t3	t4	t5
EG 1	O	X	O		O
EG 2	O		O	X	O

Abb. -3 Design der Studie (sog. „Mehrphasen“-Design)

Nachhaltigkeit messbar

Im Mehrphasen-Design resultieren Messungen nach der hochschuldidaktischen Intervention (in Abb. 3 als „X“ markiert) sowie in Phasen ohne Intervention davor und danach (in Abb. 3 als „O“ markiert). Die mehrfachen Messungen erlauben es, die Effekte der hochschuldidaktischen Intervention nicht nur auf der Ebene der unmittelbaren Reaktion der Teilnehmer, sondern auch auf später relevanten Ebenen des Lehrverhaltens und seiner Ergebnisse zu bestimmen. Die Messzeitpunkte liegen über drei Semester verteilt, die hochschuldidaktische Intervention lag in der vorlesungsfreien Zeit, die anderen Messungen am Ende der Vorlesungszeit.

Das Mehrphasen-Design hat nur teilnahmewillige Lehrende in der Stichprobe, wodurch der methodisch unsichere Vergleich zwischen Experimentalgruppe („motiviert“ Teilnahmewillige) und klassischen Kontrollgruppen („unmotiviert“ restliche Lehrende) vermieden wird. Im Mehrphasen-Design werden alle Lehrenden zu Beginn randomisiert auf zwei Experimentalgruppen (EG 1 und EG 2, vgl. Abb. 3). Die Lehrenden der zweiten Gruppe erhalten die hochschuldidaktische Intervention (in Abb. 3 als „X“ markiert) ein Semester später, und können so zuvor als Kontrollgruppe fungieren. Insgesamt nahmen 79 Lehrende teil.

Durch die Randomisierung können die häufigsten Konfundierungsquellen ausgeschlossen werden. Ein fundierter Vergleich von teilnahmebereiten Lehrenden kann sogar mit einfachen Mittelwertsvergleichen stattfinden (komplexere Verfahren wie z.B. Regressionsanalysen mit Kontrolle von Personenfaktoren etc. können auch angewandt werden).

Sicherer Ausschluss von Störfaktoren

Die Gruppe der Studierenden kann nicht randomisiert werden. Studienwahl, Fachsemester und auch Kurswahl sind durch die Hochschuldidaktik nicht kontrollierbar. Daher mussten hier methodisch anspruchsvollere Verfahren gewählt werden, um die potenziellen Störquellen (sog. „unbeobachtete Heterogenität“) in der statistischen Auswertung kontrollieren zu können. Es wurde hier das Verfahren der Mehrebenenanalyse gewählt (vgl. Field, 2009, S. 725 ff.).

Studierende nicht randomisierbar

An dieser Studie nahmen 79 Nachwuchslehrende aller Fachrichtungen teil, d.h. v.a. Doktoranden und einzelnen Post-Doktoranden. Die insgesamt 3.183 Studierenden, welche in den Lehrveranstaltungen dieser Lehrenden saßen, wurden ebenfalls befragt. Prinzipiell können alle potenziellen Lehrendengruppen (z.B. Professoren) oder auch fachspezifische Stichproben realisiert werden.

Stichprobengröße

Die Studie nutzte wenn irgend möglich etablierte Instrumente. Für die Gruppe der Lehrenden wurden teils etablierte Instrumente auf den Hochschulkontext adaptiert (Ulrich, im Druck), da dazu noch keine Fragebogen vorlagen. Erhoben wurde unter anderem:

Instrumente

- die Bewertung des Workshops: Zufriedenheit (Grohmann & Kaufeld, im Druck), Gesamtnote
- die Lernerfolge aus Sicht der Lehrenden: Wissen über gute Lehre, hochschuldidaktische Fähigkeiten (Ulrich, im Druck)
- das Verhalten der Lehrenden: Abfrage der umgesetzten hochschuldidaktischen Aspekte

- die Ergebnisse der Veranstaltung über die Lehrevaluation: Prozess der Lehrveranstaltung (Staufenbiel, 2000), Studentische Kompetenzen (Braun, Gusy, Leidner & Hannover, 2008).

Tabelle 3 zeigt die erhobenen Variablen, differenziert nach Zielgruppe und Ebene:

Ebene	Zielgruppen	
	Lehrende	Studierende
Reaktion	Bewertung des Workshops	Bewertung der Lehre
Lernen	Wissen Fähigkeiten	Lernerfolg (Kompetenzen)
Verhalten	Umgesetzte hochschuldidaktische Aspekte	--
Ergebnisse	Lehrevaluationsergebnisse	--

Tab. -3 Überblick der wichtigsten erhobenen Variablen je Ebene

Zu beachten ist, dass die Lehrevaluationsergebnisse sowohl auf der Ebene der Ergebnisse der Lehrenden als auch auf den Ebenen Reaktion und Lernen der Studierenden herangezogen wurden.

Ergebnisse

Insgesamt konnte die Studie zeigen, dass (a) die Lehrenden die hochschuldidaktische Intervention positiv bewerteten und (b) ihr Wissen und ihre Fähigkeiten durch die Intervention nachhaltig als besser einstufen. Zugleich setzten sie (c) einige hochschuldidaktische Aspekte (z.B. den erstmaligen Einsatz aktivierender Lehrmethoden) um. Die Lehrenden erhielten (d) bessere Lehrevaluationsergebnisse. Die Studierenden bewerteten die Lehre der weitergebildeten Dozierenden in der Lehrevaluation besser. Auch ihre Kompetenzen schätzten sie höher ein. Mehr Details finden sich bei Ulrich (2013).

4.2 Hochschuldidaktik und Videoaufzeichnung

Videoanalysen in der Unterrichtsforschung

In der schulischen Unterrichtsforschung werden videoanalytische Verfahren bereits seit vielen Jahren erfolgreich eingesetzt (vgl. Helmke, 2003, Kunter, 2005, Seidel & Prenzel, 2007). Videoanalysen werden hier als wichtiges Instrument zur Weiterentwicklung von Unterrichtsgeschehen angesehen, da Beobachtungen einen objektiven Zu-

gang zur Kompetenzmessung in der Lehrerbildung ermöglichen. Kompetenzprofile lassen sich auf diesem Wege durch tatsächlich gezeigtes Verhalten abbilden und analysieren. Die Möglichkeit, komplexe Prozesse auswerten zu können, spiegelt sich in einer höheren prognostischen und inhaltlichen Validität von Beobachtungsverfahren gegenüber Selbstbeobachtungsverfahren wider (Oser, Curcio & Dügeli, 2007). Für die hochschuldidaktische Lehr-Lern-Forschung wurden jedoch erst in den letzten Jahren einige Forschungsarbeiten publiziert und neu angeregt. Vorteile von Videoanalysen, wie die Möglichkeit zu direkter Beobachtung des gezeigten Verhaltens gegenüber einem Selbstbericht, und daraus folgender objektiverer Beurteilung und höherer Validität werden in diesem Forschungsgebiet bisher selten genutzt. Ausschlaggebend dafür sind ein vergleichsweise hoher personeller und finanzieller Aufwand, den Forscher betreiben müssen, um Videoanalysen durchzuführen zu können.

Den Herausforderungen, die videobasierte Forschung mit sich bringt, hat sich im Hochschulkontext in den letzten Jahren insbesondere die Arbeitsgruppe um Tina Seidel gewidmet. Wie Videoaufzeichnungen für die Qualifizierung und Wirksamkeitsmessung in der Hochschuldidaktik einsetzen werden können, soll hier exemplarisch dargestellt werden. Die enge Verknüpfung von Forschung und Anwendung wird anhand dieser Beispiele deutlich.

Seidel und Hoppert (2011) entwickelten ein Beobachtungsinstrument zur systematischen Erfassung der Gestaltung von Hochschulseminaren anhand von Videoaufzeichnungen. Zur Entwicklung des niedriginferenten Beobachtungssystems wurde ein zyklischer Prozess verfolgt. In Anlehnung an das Verfahren zur qualitativen Inhaltsanalyse (Mayring, 2002) wurden erste Videoaufzeichnungen gesichtet und diskutiert. Anschließend wurde ein Kategoriensystem entworfen, das in einem nächsten Schritt zur Auswertung weiterer Videosequenzen eingesetzt wurde. Dieser analytische Zirkel wurde fortgesetzt, bis die Kategorien stimmig waren und unter hoher Beobachterübereinstimmung (über 89%) eingesetzt werden konnten. Für die Videoaufzeichnungen wurden zwei Kameras aus unterschiedlichen Perspektiven eingesetzt. Auf diese Weise können sowohl die Lehrenden- als auch die Studierendenperspektive und insbesondere die Interaktionen aufgezeichnet und ausgewertet werden. Die Forschergruppe analysierte die Videoaufzeichnungen von Unterrichtseinheiten von 20 Lehrenden unterschiedlicher Fakultäten, um die Studierenden- bzw. Lehrendenfokussierung der Lehre analysieren und Hinweise auf eine fakultätspezifische Gestaltung von Hochschulseminaren erfassen zu können. Bei der Auswertung von Sprechanteilen, didaktischen Phasen, Arbeitsformen und -materialien konnte fachübergreifend eine Tendenz zu lehrendenzentrierter Unterrichtsgestaltung an der Hochschule gefunden werden.

Entwicklung eines videoanalytischen Instruments

Ein videounterstütztes Trainingsprogramm

Zur Unterstützung von Lehranfängern entwickelten Johannes und Seidel (2010) ein einjähriges Trainingsprogramm „Lehrqualifikation Basic“, in dem Lehranfänger über zwei Semester beim Lehreinstieg begleitet wurden. Die Lehrenden bekamen während ihrer Schulung ein Workshopangebot zur Erweiterung ihres didaktischen Methodenrepertoires, wurden in einer Lehrveranstaltungssitzung videografiert und erhielten ein einstündiges individuelles Feedback zu dieser Videoaufzeichnung (Johannes & Seidel, 2012). Darüber hinaus wurden ausgewählte Videosequenzen für eine Gruppenreflexion ausgewählt und in das Training integriert. Das in der Forschungsarbeit von Seidel und Hoppert (2011) entwickelte Beobachtungsinstrument kam in dieser Studie für Qualifizierungszwecke zum Einsatz. Hier wurden folglich Videoaufzeichnungen und -analysen gezielt zur Schulung von Hochschullehrkräften eingesetzt. Die Ergebnisse der Studie weisen auf Unterschiede zwischen selbsteingeschätzter und real beobachteter Lehrenden- bzw. Studierendenfokussierung hin und deuten außerdem darauf hin, dass durch videobasierte Trainings professionelle (Selbst-)Wahrnehmung von Lehrkräften geschult werden kann. Somit zeigen sich hier sehr anschaulich mögliche Verknüpfungspunkte videobasierter, hochschuldidaktischer Forschung und der (Weiter-)Qualifizierung von Hochschullehrenden auf.

Zukünftige Forschung mit Videoaufzeichnungen

In der hochschuldidaktischen Forschung kam bisher noch kein videoanalytisches Instrument über einen längeren Zeitraum zum Einsatz, um Verhaltensänderungen im Lehrverhalten als Anhaltspunkt für eine Entwicklung von Lehrkompetenz abzubilden und auszuwerten. Diese Aufgabe könnte Bestandteil weiterer Forschungsarbeiten mit dieser Methodik sein.

Am Interdisziplinären Kolleg Hochschuldidaktik (IKH) der Goethe-Universität Frankfurt hat sich eine Forschergruppe gebildet, die sich mit verschiedenen Methoden hochschuldidaktischer Lehr-Lern-Forschung widmet. Dort soll in den nächsten Jahren unter anderem ein Instrument zur videoanalytischen Messung der Entwicklung von Lehrkompetenzen Hochschullehrender konstruiert werden. Des Weiteren sollen ein Modell guter Hochschullehre (Ulrich, eingereicht), ein Modell zur Messung von hochschuldidaktischer Kompetenz, sowie ein Verfahren zur Diagnostik hochschuldidaktischer Kompetenz über die Bewertung standardisierter Videovignetten entwickelt werden. Videoanalysen werden hier demnach zukünftig in Kombination mit anderen Methoden eingesetzt. Die Bedeutung von videoanalytischer Wirksamkeitsmessung in der Hochschullehre soll somit gestärkt werden.

5. Ausblick

Dank der aktuellen politischen Rahmenbedingungen hat die deutsche Hochschuldidaktik die Möglichkeit, bis mindestens 2016 (zum Ende der ersten BLP-Förderungsperiode) umfassende didaktische Interventionen in der Lehre samt forschungsbegleitender Messung vornehmen zu können (und häufig auch zu müssen).

Bis dahin sollten die wissenschaftlichen Grundlagen hochschuldidaktischer Messungen weiter ausgebaut werden, z.B. über die Weiterentwicklung von Modellen guter Hochschullehre, der Durchführung erster Pilotstudien zu aufwändigeren Messmethoden und -ebenen.

Insgesamt lassen die bis 2016 anstehenden Messungen einen breiten Methodenmix in der Hochschuldidaktik erwarten, welcher sowohl die Lehre (in der die Methoden eingesetzt werden), als auch die hochschuldidaktische Forschung insgesamt befruchten wird.

Zugriffsdatum für alle elektronischen Quellen: 5.12.2012

Literatur

- [1] Abrami, P. C.; d'Appolonia, S.; Rosenfield, S. (2007): The dimensionality of student ratings of instruction. What we know and what we do not. In: Perry, R. P.; Smart, J. C. (Hrsg.): *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*. New York, S. 385–456.
- [2] Alliger, G. M.; Tannenbaum, S. I.; Bennett, W.; Traver, H.; Shotland, A. (1997): A meta-analysis of the relations among training criteria. In: *Personnel Psychology*, 2 (50. Jg.), S.341-358.
- [3] Arthur, W.; Bennett, W.; Edens, P. S.; Bell, S. T. (2003): Effectiveness of training in organizations. A meta-analysis of design and evaluation features. In: *Journal of Applied Psychology*, 2 (88. Jg.), S. 234-245.
- [4] Bassi, L. J.; Benson, G.; Cheney, S. (1996): The top ten trends. In: *Training & Development*, 11 (50. Jg.), S. 28-42.
- [5] Bates, R. (2004): A critical analysis of evaluation practice. The Kirkpatrick model and the principle of beneficence. In: *Evaluation and Program Planning*, (27. Jg.), S. 341-347.
- [6] Berendt, B. (2000): Was ist gute Hochschullehre? In: *Zeitschrift für Pädagogik*, (41. Jg.), S. 247-260.
- [7] Berendt, B. (2005): „Academic Staff Development“ im Kontext und zur Unterstützung des Bologna-Prozesses. Stellenwert und Stand hochschuldidaktischer Weiterbildung. In: Berendt, B.; Szczyrba, B.; Wildt, J. (Hrsg.): *Neues Handbuch Hochschullehre*. Berlin, L 2.2.
- [8] Berendt, B. (2011): Academic Staff Development als Teil von Qualitätssicherung und -entwicklung im Kontext des Bologna-Prozesses. Entwicklungen hochschuldidaktischer Aus- und Weiterbildung 2008-2011 in Deutschland. In: Berendt, B.; Szczyrba, B.; Wildt, J. (Hrsg.): *Neues Handbuch Hochschullehre*. Berlin, L 2.7.
- [9] Bortz, J.; Döring, N. (2002): *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. 3. überarbeitete Auflage. Berlin.
- [10] Braun, E.; Gusy, B.; Leidner, B.; Hannover, B. (2008): Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). In: *Diagnostica*, 1 (54. Jg.), S. 30-42.
- [11] Brinko, K. T. (1993): The practice of giving feedback to improve teaching. In: *Journal of Higher Education*, 5 (64. Jg.), S. 574-593.
- [12] Bundesministerium für Bildung und Forschung (2010): *Bund-Länder-Programm für bessere Studienbedingungen und mehr Qualität in der Lehre*. <http://www.bmbf.de/foerderungen/15440.php>
- [13] Deming, W. E. (1982): *Out of the crisis*. Cambridge, Mass.
- [14] Donabedian, A. (1966): Evaluating the Quality of Medical Care. In: *The Milbank Quarterly*, 3 (44. Jg.), S. 166-203.

- [15] Feldman, K. A. (1989): The association between student ratings of specific instructional dimensions and student achievement. In: *Research in Higher Education*, 6 (30. Jg.), S. 583-645.
- [16] Field, A. (2009): *Discovering statistics using SPSS*. 3. Auflage. London.
- [17] Gold, A.; Souvignier, E. (2000): Rückmeldegespräche nach studentischen Referaten. Ein Beitrag zur Verbesserung der Lehre? In: Krampen, G.; Zayer, H. (Hrsg.): *Psychologiedidaktik und Evaluation II. Neue Medien, Psychologiedidaktik und Evaluation in der psychologischen Haupt- und Nebenfachausbildung*. Bonn, S. 203–218.
- [18] Glüer, M. (2011): Informationen zur computergestützten Videodatenanalyse. <http://www.glueer.com/computergestuetzte-videodatenanalyse/>
- [19] Grohmann, A.; Kauffeld, S. (im Druck): Evaluating training programs: Development and correlates of the Questionnaire for Professional Training Evaluation (Q4TE). In: *International Journal of Training and Development*.
- [20] Hattie, J. A. (2009): *Visible Learning. A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Oxon/New York.
- [21] Helmke, A. (2012): Unterrichtsqualität. In: Horn, K.-P.; Kemnitz, H.; Marotzki, W. & Sandfuchs, U. (Hrsg.): *Klinkhardt Lexikon Erziehungswissenschaft*, Band 3. Bad Heilbrunn, S. 348-349.
- [22] Helmke, A. (2003): Unterrichtsevaluation: Verfahren und Instrumente. In: *Schulmanagement*, 1, S. 8-11.
- [23] Helmke, A.; Schrader, F.-W. (2006): Hochschuldidaktik. In: Rost, D. H. (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. 3. überarbeitete und erweiterte Auflage. Weinheim, S. 246-252.
- [24] Hochschuldidaktik Hessen. (2011). *Adaptation der Standards der Lehrerbildung der Kultusministerkonferenz 2004 auf den Hochschulkontext. Treffen der hessischen Hochschuldidaktiker (7. November 2011, Frankfurt am Main)*. (Unveröff. Manuskript).
- [25] Höhler, J.; Horz, H.; Hansen, M.; Fabriz, S.; Stehle, S.; Heckmann, C.; Ulrich, I.; Secon, R. (im Druck): Verbesserte fachliche und soziale Integration von Studienanfängern durch Peerteaching im Bachelorstudiengang Psychologie. In: Krämer, M. (Hrsg.): *Psychologiedidaktik und Evaluation IX*.
- [26] Holling, H.; Liepmann, D. (2007): Personalentwicklung. In: Schuler, H. (Hrsg.), *Lehrbuch Organisationspsychologie*. 4., aktualisierte Auflage. Bern, S. 345–383.
- [27] Johannes, C.; Seidel, T. (2010): Professionelles Lernen von Anfängern in der Hochschullehre – Erwartungen und Vorstellungen über Hochschullehre im Rahmen des Projekts LehreLernen. In: *Personal- und Organisationsentwicklung in Einrichtungen der Lehre und Forschung (P-OE)* 5, S. 31-40.
- [28] Johannes, C.; Seidel, T. (2012): Professionalisierung von Hochschullehrenden: Lehrbezogene Vorstellungen, Wissensanwendung und Identitätsentwicklung in einem videobasierten Qualifikationsprogramm.
- [29] Kirkpatrick, D. L. (1960 // 1979): Techniques for evaluating training programs. In: *Training and Development Journal*, 6 (33. Jg.), S. 78-92.
- [30] Klieme, E. (2006): Unterrichtsqualität, Schulqualität und Leistungsentwicklung im Fach Deutsch. Befunde der DESI-Studie. Symposium Deutschdidaktik. Weingarten, 18. September 2006 (Unveröff. Manuskript).

- [31] Koch, E. (2004): Gute Hochschullehre. Theoriebezogene Herleitung und empirische Erfassung relevanter Lehraspekte. Hamburg.
- [32] Kultusministerkonferenz (2004): Standards für die Lehrerbildung. Bildungswissenschaften.
http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- [33] Kunter, M. (2005): Multiple Ziele im Mathematikunterricht. Münster.
- [34] Ledić, J.; Rafajac, B.; Kovač, V. (1999): Assessing the quality of university teaching in Croatia. In: Teaching in Higher Education, 2 (4. Jg), S. 213-233.
- [35] Marzano, R. J. (1998): A theory-based meta-analysis of research on instruction. Aurora, CO.
- [36] Mayring, P. (2002). Einführung in die qualitative Sozialforschung. 5. Auflage. Weinheim.
- [37] Meyer, H. (2004): Was ist guter Unterricht? Berlin.
- [38] Mummendey, H. D.; Grau, I. (2008): Die Fragebogen-Methode. 5. Auflage. Göttingen.
- [39] Murray, H. G. (1997): Effective teaching behaviors in the college classroom. In: Perry, R. P.; Smart, J. C. (Hrsg.): Effective teaching in higher education. Research and practice. New York, S. 171-204.
- [40] Murray, H. G.; Lawrence, C. (1980): Speech and drama training for lecturers as a means for improving university teaching. In: Research in Higher Education, 1 (13. Jg.), S. 73-90.
- [41] Oser, F., Curcio, G.-P., Düggeli, A. (2007): Kompetenzmessung in der Lehrerbildung als Notwendigkeit – Fragen und Zugänge. In: Beiträge zur Lehrerbildung, 25, S. 14-25.
- [42] Preiser, S. (2000): Feedback nach Referaten als hochschuldidaktisches Instrument. In: Krampen, G.; Zayer, H. (Hrsg.): Psychologiedidaktik und Evaluation II. Neue Medien, Psychologiedidaktik und Evaluation in der psychologischen Haupt- und Nebenfachausbildung. Bonn, S. 187-202.
- [43] Rindermann, H. (2010): Kennzeichen guter Lehre. In: Spiel, C.; Schober, B.; Wagner, P.; Reimann, R. (Hrsg.): Bildungspsychologie. Göttingen, S. 410-415.
- [44] Saks, A. M. (2002): So what is a good transfer of training estimate? A reply to Fitzpatrick. In: The Industrial-Organizational Psychologist, 1 (40. Jg.), S. 29-30.
- [45] Seidel, T.; Hoppert, A. (2011): Merkmale von Lehre an der Hochschule. Ergebnisse zur Gestaltung von Hochschulseminaren mittels Videoanalysen. In: Unterrichtswissenschaft, 2, S. 154-172.
- [46] Seidel, T. & Prenzel, (2006): Stability of teaching patterns in physics instructions: findings from a video study. In: Learning and Instruction, 16, S. 228-240.
- [47] Solga, M. (2011): Management des Lerntransfers. In: Ryschka, J.; Solga, M.; Mattenklott, A. (Hrsg.): Praxishandbuch Personalentwicklung. Instrumente, Konzepte, Beispiele. 3., vollst. überarb. und erw. Auflage. Wiesbaden, S. 339-367.
- [48] Spiel, C.; Gössler, P. M. (1999): Allgemeingültige Qualitätskriterien von Lehre – eine Fiktion? Ergebnisse einer repräsentativen Studie an der Universität Wien. In: Zeitschrift für Hochschuldidaktik, 2 (23. Jg.), S. 72-95.

- [49] Staufenbiel, T. (2000): Fragebogen zur Evaluation universitärer Lehrveranstaltungen durch Studierende und Lehrende. In: *Diagnostica*, 4 (46. Jg.), S. 169-181.
- [50] Taylor, P. J.; Russ-Eft, D. F.; Chan, D. W. L. (2005): A meta-analytic review of behavior modeling training. In: *Journal of Applied Psychology*, 4 (90. Jg.), S. 692-709.
- [51] Ulrich, I. (2013): Strategisches Qualitätsmanagement in der Hochschullehre. Theoriegeleitete Workshops für Lehrende zur Förderung kompetenzorientierter Lehre. Wiesbaden.
- [52] Ulrich, I. (im Druck): Strategische wissenschaftlich fundierte universitäre Personalentwicklung zu guter Lehre und ihre Effekte bei Lehrenden und in der Studierendenevaluation: Messinstrumente. In: Krämer, M. (Hrsg.), *Psychologiedidaktik und Evaluation IX*.
- [53] Ulrich, I. (eingereicht): Model of good teaching in higher education. Relevant aspects and effect sizes. Paper presentation at the 15th Biennial Conference Earli 2013, Munich, Germany.
- [54] Wagner, K. (2007): Ausgestaltung von QM-Systemen auf Basis der ISO-9000-Reihe. In: Masing, W.; Pfeifer, T.; Schmitt, R. (Hrsg.): *Handbuch Qualitätsmanagement*. 5., vollst. neu bearb. Auflage. München, S. 173-205.
- [55] Walberg, H.J. (1984): Improving the productivity of America's schools. In: *Educational Leadership*, 8 (41. Jg.), S. 24.
- [56] Walberg, H. J.; Lai, J.-S. (1999): Meta-Analytic Effects for Policy. In: G. J. Cizek (Hrsg.): *Handbook of Educational Policy*. London, S. 419-453.
- [57] Webler, W.-D. (1991): Kriterien für gute akademische Lehre. *Das Hochschulwesen*, 6 (39. Jg.), S. 243-249.
- [58] Wissenschaftsrat (2008, 04. Juli): Empfehlungen zur Qualitätsverbesserung von Lehre und Studium. <http://www.wissenschaftsrat.de/download/archiv/8639-08.pdf>.

Informationen zum Autor und zur Autorin

Dr. **Immanuel Ulrich**, Diplom-Psychologe, ist wissenschaftlicher Mitarbeiter am Interdisziplinären Kolleg Hochschuldidaktik (IKH) der Goethe-Universität Frankfurt. Die Schwerpunkte seiner Tätigkeiten liegen in der Konzeption und Verknüpfung von wissenschaftlichen Kennzeichen guter Hochschullehre, hochschuldidaktischer Intervention und kompetenzorientierter (Lehr-)Evaluation. Daneben hat er (Lehr-)Evaluationen an natur-, geistes- und sozialwissenschaftlichen Fachbereichen implementiert, führt selbst hochschuldidaktische Workshops für Lehrende durch und gibt Lehrveranstaltungen für Studierende, v.a. in der pädagogischen Psychologie.

Carmen Heckmann, Diplom-Psychologin, ist wissenschaftliche Mitarbeiterin am Interdisziplinären Kolleg Hochschuldidaktik (IKH) der Goethe-Universität Frankfurt. Dort ist sie in der hochschuldidaktischen Weiterbildung der Lehrenden der Goethe-Universität, mit Schwerpunkt auf den geisteswissenschaftlichen Fächern, tätig. In ihrer Forschung beschäftigt sie sich insbesondere mit der Entwicklung von Lehrkompetenzen, Videoanalysen und Kriterien guter Hochschullehre.

Kontakt:

Interdisziplinäres Kolleg Hochschuldidaktik
Goethe-Universität Frankfurt
PEG
Grüneburgplatz 1
60323 Frankfurt am Main

Immanuel Ulrich Tel. 069 / 798 – 28321
Email: ulrich@psych.uni-frankfurt.de
Web: www.immanuel-ulrich.de

Carmen Heckmann Tel. 069 / 798 – 23259
Email: heckmann@paed.psych.uni-frankfurt.de