

# **Innovationen in der Hochschullehre evaluieren. Herausforderungen für Methodik und Konzeption**

*Philipp Pohlentz*

*Otto-von-Guericke-Universität Magdeburg*

## **1 Evaluation in der Hochschullehre: Zwischen externer Kontrolle und interner Entwicklung**

Evaluation hat in der Hochschullehre in institutionalisierter Form seit den 1990er Jahren eine stärkere Bedeutung erhalten. Insbesondere im Zuge neuer Steuerungsmodelle, die auf die wettbewerbliche und stärker eigenverantwortliche Selbststeuerung von Hochschulen ausgerichtet sind („New Public Management“; vgl. zusammenfassend Lange 2008) haben Qualitätsmanagement, Akkreditierung und hochschulinterne (sowie -externe) Evaluationen Einzug in den Alltag des Wissenschaftsbetriebes gehalten. Mittlerweile ist in den Landeshochschulgesetzen die Durchführung von Evaluation als Mittel der Qualitätssicherung geregelt. Die Erwartungen an das, was Evaluation leisten soll, sind zugleich hoch und gehen vielfach über die Funktion der Qualitätssicherung hinaus. Vielmehr wird *Qualitätsentwicklung* erwartet, also eine Verbesserung von Prozessen und Praktiken auf der Basis von Evaluationsergebnissen. *Qualitätssicherung* hingegen meint eher die Einhaltung von Mindeststandards, die durch Evaluationen überprüft werden soll.

Hier wird ein Problem deutlich, welches die Evaluation im Hochschulbereich seit ihrer Einführung als kontroverse Diskussion begleitet hat: Mit der Durchführung von Evaluationen werden verschiedene, zuweilen in Konflikt miteinander stehende Ziele verfolgt. Evaluation als angewandte empirische (Sozial-)Forschung hat zum Ziel, Routinen der Durchführung von Projekten bzw. Programmen (wie

etwa Interventionen zur Verbesserung der Studienqualität bzw. ganze Studiengänge) zu verbessern, um deren Ergebnisqualität (Senkung des Studienabbruchs, Steigerung des Kompetenzerwerbs von Absolventinnen und Absolventen, etc.) zu erhöhen. Als Steuerungsinstrument hingegen wird Evaluation – teilweise zeitgleich – für die Kontrolle der Zielerreichung bei der Durchführung von Programmen eingesetzt. In diesem Sinne geht es ihr darum, Steuerungsentscheidungen – etwa zur Weiterführung von Programmen und Projekten – auf eine rationale Basis zu stellen und dadurch zugleich Transparenz zu schaffen und Legitimität für das Handeln von Organisation, also hier: Hochschulen herzustellen.

Die genannten Ziele von Evaluation<sup>1</sup> sind je für sich genommen ebenso valide, wie sie in Widerspruch zueinander stehen können. So stellt eine auf Kontrollfunktionen ausgerichtete Evaluation potenziell eine Bedrohung der Handlungs- und Entscheidungsautonomie der am jeweils evaluierten Projekt Beteiligten dar, insofern beispielsweise die Fortsetzung (oder auch die Art der zukünftigen Durchführung) dieses Projekts von einer auf der Evaluation aufsetzenden Steuerungsentscheidung abhängt. Entsprechend besteht das Interesse der „evaluierten Personen“, die – in solchen Fällen zumeist extern durchgeführte – Evaluation für die Verlautbarung von „Erfolgsmeldungen“ zu nutzen und das eigene Projekt gegenüber den Evaluierenden von der besten Seite zu zeigen. Dies könnte man auch so beschreiben, dass eine entsprechend durchgeführte Evaluation nur in sehr geringem Grade über Fehlertoleranzen verfügt. Gerade diese sind aber die Voraussetzung dafür, dass man aus Problemen, die eine retrospektive Evaluation sichtbar macht, mithin für die weitere, interne Entwicklung von Projekten und Programmen aus Evaluationen lernt.

Eine wichtige Aufgabe von Evaluation bzw. der sie durchführenden Evaluatoredinnen und Evaluatoren besteht daher zunächst darin, den Nutzungskontext und das Verwertungsinteresse der Evaluation und ihrer Ergebnisse zu klären und transparent zu machen. Wie gesagt: Die verschiedenen Ziele von Legitimations- bis zu Entwicklungsfunktionen sind je für sich genommen überaus valide. So ist es plausibel, dass beispielsweise die weitgehend öffentlich finanzierte Hochschullehre sich gegenüber der Öffentlichkeit für die Art und die Ergebnisse des

























---

<sup>1</sup> Neben die beschriebenen Funktionen der Entwicklung, Kontrolle und Legitimation, tritt üblicherweise noch das Ziel der Forschung, im Sinne eines Beitrags von Evaluationen, generalisierungsfähiges Wissen über die evaluierten Gegenstände herzustellen (vgl. Schmidt 2010).

Miteinsatzes in diesem Bereich rechtfertigen können muss. Zugleich ist für eine Profilierung zunehmend stärker wettbewerblich agierender Hochschulen erforderlich, dass sie sich – nicht zuletzt auf der Basis von Evaluationen – in ihrer autonomen Strategiebildung weiterentwickeln können. Allerdings ist gerade hierfür erforderlich, dass Evaluationsergebnisse, die Optimierungsbedarfe offenbaren, eben nicht in unmittelbaren Steuerungsentscheidungen münden, sondern Entwicklungsspielräume für eine langfristige Perspektive eröffnen.

Die unterschiedlichen Instrumente, die im Bereich der Hochschullehre für die Evaluation landläufig eingesetzt werden, wie etwa Peer Reviews oder studentische Lehrveranstaltungsevaluationen (LVE) sind in unterschiedlicher Weise für die Erreichung der mit den verschiedenen Nutzungsarten von Evaluation verbundenen Ziele geeignet. Schmidt (2010) fasst dies wie in der folgenden Tabelle 1 dargestellt, zusammen:

Tab. 1: Ziele und Instrumente der Evaluation der Hochschullehre, zit. n. Schmidt (2010)

	Kontrolle	Entwicklung	Legitimation	Forschung
Peer Reviews				
LVE				
Absolventenstudien				
Rankings				
Akkreditierung				
Hochschulstatistiken				

Demnach sind insbesondere Evaluationsverfahren, die im Sinne von fachnahen Begutachtungen im Rahmen von Peer Reviews sowie der Auswertung hochschulstatistischer Daten (beispielsweise zu Absolventen- und Studienabbruchquoten) durchgeführt werden, dazu geeignet, Entwicklungen im Sinne hochschulinterner Verbesserungen anzuschieben. Stärker ausgeprägt sind indes die Bezüge der meisten eingesetzten Verfahren zur Kontroll- und Legitimationsfunktion von Evaluation. Forschungsziele werden durch die eingesetzten Instrumente insgesamt in vergleichsweise geringem Maße erreicht.

Insbesondere durch den letztgenannten Befund lässt sich vermutlich erklären, warum Evaluation seit ihrer Institutionalisierung von Kontroversen im Wissenschaftsbetrieb begleitet war. Sie wurde (und wird) vielfach als wissenschaftsexternes Beurteilungsinstrument wahrgenommen, welches auf der Basis von ebenso wissenschaftsfremden wie in ihrer Validität bedrohten Kriterien zu unsachgemäßen Urteilen über die Qualität der Lehre kommt (Pohlenz 2009). Mit Schimank (2015) kann man die Reaktion vieler Hochschullehrender auf die Einführung von Evaluation als institutionalisiertes Steuerungsinstrument als eine Form der Identitätsbehauptung, mit der auf eine Identitätsbedrohung reagiert wird, interpretieren: Durch die Evaluation an sich sowie die mit ihr verbundenen Aufwände (etwa im Bereich der Dokumentationspflichten) wird die Handlungs- und Entscheidungsautonomie der Hochschullehrenden eingeschränkt. Da diese aber geradezu identitätsstiftend und konstitutiv für das Selbstverständnis als Wissenschaftlerin oder Wissenschaftler ist, werden die „Abwehrkräfte“ gegen die Bedrohung organisiert.

Im Feld der Evaluation von Studium und Lehre drückt sich dies vielfach in einer methodischen Kritik der eingesetzten Verfahren aus, die – in Abhängigkeit von der Art der jeweils eingesetzten Evaluationsverfahren und vor allem von den Nutzungsszenarien – durchaus berechtigt ist. Gerade mit Hinweis auf die Funktion von Evaluationen, Steuerungsentscheidungen begründen zu können, ist eine hohe methodische Elaboration notwendig, um Gerechtigkeitsprobleme, die aufgrund von unsachgemäßen Evaluationsverfahren und -ergebnissen entstehen können, zu vermeiden.

## **2 Hohe methodische Standards: Evaluationen als Wirkungsanalysen?**

Insbesondere wenn Evaluation als Kontrollinstrument für eine Leistungsbeurteilung und -sanktionierung (sei diese nun positiv oder negativ) eingesetzt wird, ist von ihr zu erwarten, dass die Urteile, die gefällt werden, auf der Basis belastbarer Befunde zustande kommen. So wird mit Blick auf studentische Lehrveranstaltungsevaluationen (sowie auch hinsichtlich anderer studentischer Befragungsdaten, etwa aus Absolventenstudien) vielfach über die Frage der Repräsentativität der Ergebnisse sowie über deren Validität und Quellen für deren Verzerrung diskutiert (z. B. Pohlenz 2009, Wolbring 2013).

Aber über diese Aspekte hinaus stellt sich die methodisch grundlegendere Frage nach den Möglichkeiten von Evaluationen, Wirkungen, im Sinne kausaler Zusammenhänge zwischen Inputs und Outputs zu analysieren. Wirkungsorientierte Fragestellungen stehen typischerweise im Mittelpunkt des Interesses bei anstehenden Entscheidungen über die Zukunft von Projekten und Programmen: „Sind die mit dem Projekt oder Programm beabsichtigten Wirkungen erreicht worden oder nicht? Hat die Einführung einer strukturierten Studieneingangsphase zur Senkung der Studienabbruchquote geführt? Haben die Studierenden die intendierten Kompetenzen erworben, so wie vom Studiengang beabsichtigt und durch diesen angeregt?“

Aus methodischer Sicht sind damit hohe Anforderungen an die Evaluation gestellt. Ihre Potenziale zur Analyse von kausalen Zusammenhängen im beschriebenen Sinne sind indes begrenzt. Dafür stehen der Durchführung von Evaluationen bestimmte Hindernisse im Wege.

- Voraussetzung wäre zunächst, dass die zwischen den zur Verfügung stehenden Ressourcen (Projektmittel, Programmbestandteile) bzw. Inputs einerseits und den darauf bezogenen intendierten Wirkungen (Projektziele) andererseits klare theoretische Zusammenhänge bestehen und dass die entsprechenden Wirkungszusammenhänge auch in der Modellierung des zu evaluierenden Gegenstandes bereits mitgedacht wurden, um entsprechend „Erfolgskriterien“ und Indikatoren entwickeln zu können. Dem steht beispielsweise entgegen, dass die durch ein Studienprogramm intendierten Lerneffekte sich zuweilen erst zeitlich stark verzögert einstellen und nicht mehr als unmittelbarer „Effekt“ der Teilnahme am Studienprogramm in Erscheinung treten.
- Ebenso schwierig ist es, für das theoretisch modellierte Wirkungsgefüge die entsprechenden Indikatoren zu entwickeln und diese eindeutig auf die verschiedenen Programmbestandteile (Studieneingangsphase, Teilnahme an Lehrveranstaltungen, Praktika, etc.) und deren vermutete Effekte zu beziehen. Es stellt sich mithin die Frage: „Welche Projektaktivität führt zu welchem Output? Wodurch entstehen bei den Absolventinnen und Absolventen die intendierten Qualifizierungen im Einzelnen?“

- Hochschullehre ist – wie viele andere Felder, auf denen durch Evaluationen Wirkungen ermittelt werden sollen – ein komplexer sozialer Prozess, die Lehrenden sind mit ihrem Lehrhandeln nicht allein ursächlich für studentischen Lernerfolg oder Misserfolg, vielmehr sind Studierende als „Ko-Produzenten“ sowohl ihres eigenen Lernerfolges, wie auch – über das Herstellen einer kommunikativen Situation – desjenigen der anderen Studierenden. Entsprechend vielschichtig ist das theoretische und empirische Modell, welches man für die Erklärung von Erfolg oder Misserfolg zugrunde legen müsste.
- Kausalanalysen zur Beziehung zwischen Inputs und Outputs erfordern experimentallogische Forschungsdesigns, in denen man „modellexterne“ Einflussgrößen (wie etwa im vorliegenden Fall die Auswirkung des sozialen Hintergrundes der Studierenden auf deren Lernerfolg) weitgehend kontrollieren kann. Hochschullehre findet jedoch im Feld der sozialen Welt Hochschule statt, so dass experimentallogische Forschungsdesigns, in denen die miteinander in Relation stehenden Variablen (und nur diese) beobachtet werden können, im Grunde kaum umsetzbar sind. Methodisch gesprochen besteht der „Goldstandard“ der Evaluationsforschung in Zufallsexperimenten (Randomised Controlled Trials, RCT), in denen eine Experimentalgruppe mit einer Kontrollgruppe im Zeitverlauf verglichen wird und dabei beide Gruppen in unterschiedlicher Weise der interessierenden bzw. zu evaluierenden Intervention ausgesetzt werden. Gegenstand der Evaluation ist dann der Vergleich bspw. des Studienerfolgs, der sich ggf. in den beiden Gruppen unterschiedlich darstellt, je nachdem, ob man der Intervention (strukturierte Studieneingangsphase) „ausgesetzt“ war oder nicht (z. B. Duflo 2007, Bortz & Döring 2002).

Bereits in den 1990er Jahren beschäftigte sich die Community der Evaluationsforschung mit methodischen Herausforderungen und Ansprüchen an eine entsprechend robuste Vorgehensweise, wie aus dem folgenden, längeren Zitat deutlich wird, welches zugleich handlungsleitend für die Entwicklung von Evaluationsdesigns ist:

„Vor der Entwicklung des Forschungsdesigns muss Klarheit über die Untersuchungsziele – bezogen auf einen definierbaren und empirisch abgrenzbaren Untersuchungsgegenstand – bestehen. Für die Dauer der Datenerhebung dürfen sich weder die Untersuchungsziele noch die wesentlichen Randbedingungen des Untersuchungsgegenstandes in unvorhersehbarer Weise ändern. Vor der Entwicklung des Forschungsdesigns müssen [...] begründete Vermutungen (Hypothesen) über die Struktur des Gegenstandes wie auch über die Zusammenhänge und Beziehungen zwischen dessen wesentlichen Elementen existieren, nach Möglichkeit in Form empirisch bewährter Theorien. Erst auf ihrer Basis kann ein Gültigkeit beanspruchendes Indikatorenmodell konstruiert, können geeignete Messinstrumente entwickelt, kann über problemangemessene Auswertungsverfahren entschieden werden. Der Forscher muss die Kontrolle über den Forschungsablauf haben, um die (interne und externe) Gültigkeit der Resultate weitgehend sicher zu stellen“ (Kromrey 1995, S. 323).

Mit Blick auf die landläufig eingesetzten Verfahren der Evaluation im Bereich von Lehre und Studium ist zu sagen, dass diese die hier formulierten methodischen Anforderungen in den meisten Fällen nicht einhalten. Das bedeutet jedoch nicht, dass Verfahren, die die gesetzten Standards „unterschreiten“, prinzipiell keine nützlichen Informationen erbrächten. Es ist allerdings klar zu benennen, dass sie nicht geeignet sind, Wirkungen zwischen Inputs und Outputs als kausale Zusammenhänge sichtbar zu machen.

Die Tatsache, dass Evaluationsdesigns im beschriebenen Sinne im Feld der Hochschullehre kaum umzusetzen sind, Evaluationsergebnisse (bzw. die Berichte darüber) vielfach aber implizieren, dass genau dies durch die entsprechende Evaluation geleistet wurde und dass gleichzeitig anhand von Aussagen über Wirkungsrelationen Leistungsbeurteilungen und -sanktionierungen vorgenommen werden, hat vermutlich im Wissenschaftsbetrieb das Unbehagen über Evaluation an sich über die Zeit verstärkt. Es stellt sich mithin die Frage, wie Evaluationen organisiert sein können, um die Akzeptanz der Betroffenen bzw. „Evaluierten“ zu sichern und Evaluationen zu einem nützlichen Steuerungsinstrument zu machen. Darüber wird am Ende des vorliegenden Beitrages zu diskutieren sein. Nachfolgend soll kurz auf alternative Herangehensweisen im Vergleich zur beschriebenen Experimentallogik eingegangen werden.

### 3 Evaluation als Testen von Hypothesen: Theoriebasierte Evaluationen

Die Alternative zu randomisierten Experimenten, die in der frühen Phase der Evaluation im Bereich von Hochschullehre typischerweise gewählt wurde, bestand (und besteht) vielfach in einer stärkeren Fokussierung von Prozessen.

„Measuring program outcomes requires a great deal of rigor; it’s demanding. When program evaluators discovered how difficult it was, many abandoned the attempt and decided to focus on process, which is generally more traceable” (Smith 1994).

Unter prozessorientierten Evaluationen ist zu verstehen, dass die Art der Durchführung von Programmen stärker in den Blick genommen wird, als dessen Ergebnisse. Dies spiegelt sich im Bereich der Hochschulbildungsevaluation in Verfahren wie der studentischen Lehrveranstaltungsevaluation wider. Hier wird vielfach der Lehrprozess zum Gegenstand gemacht, indem die Studierenden um Zufriedenheitsurteile zu bestimmten Aspekten, wie etwa dem Tempo des Lehrvortrages, der Strukturiertheit des Stoffes, etc. gebeten werden.

Der studentische Kompetenzerwerb als Ergebnis der Teilnahme an den Lehrveranstaltungen ist dabei deutlich weniger im Fokus. Dies ist gerade das Ergebnis der beschriebenen methodischen Limitationen, die Evaluation aus methodischer Sicht bei der „Diagnostik“ von Kompetenzen bzw. den Relationen von Inputs der Lehre und den Outputs in Form eben dieses Kompetenzerwerbs hat. Kompetenzbezogene Lehrveranstaltungsevaluationen gibt es entsprechend zumindest im deutschsprachigen Raum eher selten (Braun 2007).

Mit Blick auf den Anspruch an Evaluationen, Programmwirkungen sichtbar zu machen, lässt sich eine Brücke zu den oben beschriebenen experimentallogischen Ansätzen bauen. Mittels der Entwicklung eines theoretischen Modells über das Wirkungsgefüge von Projektressourcen und -aktivitäten sowie den von ihnen ausgehenden, beobachtbaren Resultaten lassen sich theoriebasierte Evaluationen realisieren. Diese beruhen auf der Formulierung von Hypothesen zu den vermuteten Wirkungen von Inputs auf intendierte (oder nicht-intendierte) Outputs. Dabei wird in den Mittelpunkt gestellt, dass die zu evaluierenden Projekte oder Programme an sich ein Geflecht von Wirkungsannahmen sind (z. B. bestehend in Aussagen und Entscheidungen zum Programmdesign, etwa in dem Sinne, dass



Projektarbeiten der Studierenden zur Förderung ihrer Soft Skills beitragen und entsprechend als curriculare Bestandteile von Studiengängen modelliert werden).

Diese Wirkungsannahmen können dann in Systeme von empirisch testbaren Hypothesen übersetzt und im Rahmen von Evaluationen bearbeitet werden. „Evaluation seeks to discover whether programmes work; programmes are theories. Therefore it follows that: evaluation is theory-testing“ (Pawson 2003, S. 472).

Auch ohne eine experimentallogische Wirkungsüberprüfung liegt die Stärke des Ansatzes theoriebasierter Evaluation (Giel 2013) darin, dass die Evaluierenden die Programmlogik umfassend in ihrer Komplexität erfassen und dies in der Modellierung von Wirkungsbeziehungen berücksichtigen. Entsprechende empirische Untersuchungen sind in hohem Maße informativ für die Programmdurchführenden, gerade weil nicht versucht wird, künstliche Laborsituationen zu schaffen, sondern umfassend die soziale Realität in der Formulierung von Wirkungshypothesen abzubilden.

Auch die in einem solchen Forschungs- bzw. Evaluationsparadigma einzusetzenden Instrumente und Methoden können variabel auf die zu beantwortenden Fragen abgestimmt werden und von Befragungen über nicht responsive Verfahren (Beobachtungen) bis zur Analyse von einschlägigen Dokumenten und Daten (Statistiken) sowie prozessgenerierten Daten (z. B. Analyse von Nutzerdaten aus E-Learningsystemen) reichen. Ziel ist es wie gesagt, die soziale Realität und die eingesetzten Praktiken eines Programms mit Blick auf die zu erreichenden Programmziele zu analysieren.

Im Zusammenhang mit der Evaluation von Lehrinnovationen, also neuen sozialen Praktiken und neuen Programmbestandteilen (etwa digitalen Lehr-/Lernmedien) stehen vermutlich gerade Fragen im Vordergrund, für die nicht so sehr experimentallogisch „abgesicherte“ Informationen relevant sind, als vielmehr eine theoriebasierte oder „theoretisierende“ Annäherung an den Gegenstand selber. Theoriebasierter Evaluation kommt in diesem Sinne eine entwickelnde Dimension zu: Indem empirisch testbare Hypothesen zu Zusammenhängen zwischen Programmbestandteilen die Funktionsweise von Programmen und Projekten aufzuklären helfen, kann eine so verstandene Evaluation zur weiteren Entwicklung dieser Programme beitragen.

Entsprechend einer solchen Zielsetzung würden unter Evaluationen bzw. theoriebasierte Evaluationen dann auch Praktiken aus dem Bereich der Aktionsforschung fallen, wie etwa beim „design-based research (z.B. Gess et al. 2014): Im Prozess der Durchführung eines Programms werden Annahmen zu dessen Funktionsweise getroffen und als empirisch überprüfbare Hypothesen formuliert. Die Ergebnisse werden unmittelbar in den Prozess der Projektdurchführung zurückgemeldet und diese dadurch modifiziert und verbessert. Die iterative Durchführung mehrerer Schleifen bis zur Erreichung eines Optimums führt zugleich zur systematisierten Bestandsaufnahme hinsichtlich relevanter Einflussgrößen und zur Generalisierung der darauf bezogenen Befunde.

Evaluation wird auf diese Weise zu einer auf die Lehre bezogenen Forschungspraxis, mittels derer Lehrende selbstreflexiv Fragen an die eigene Lehrpraxis stellen können. Auf diese Weise lassen sich vermutlich eher Distanzen zu den Zielen und Vorgehensweisen von Evaluation abbauen.

## 4 Evaluation und Lehrinnovation

Derzeit wichtigstes oder am meisten bestimmendes Thema im Bereich der Innovation von Lehre und Studium ist vermutlich deren fortschreitende Digitalisierung. Diese setzt auch Impulse für die Evaluation in der Hochschullehre.

Diese beziehen sich sowohl auf den Gegenstand der Evaluation, also die digitale Lehre an sich, wie auch auf die Innovationen, die im Bereich des Methodenrepertoires der Evaluation zu verzeichnen sind.

In der zuerst genannten Perspektive, also der Evaluation der digitalen Lehre ist die Frage nach neuen Qualitätsvorstellungen von Hochschulbildung zu stellen und diejenige danach, wie diese Qualitätsvorstellungen mittels digitaler Lehr-/Lernmedien erreicht werden (können). Konkret geht es dabei um die Individualisierung und Personalisierung von Lehr-/Lernprozessen, in dem Sinne, dass unter Nutzung digitaler Medien in höherem Maße als bisher eine Selbststeuerung von Lernprozessen realisiert werden kann. Evaluationsfragestellungen, die sich darauf beziehen, kreisen um die Frage nach den Determinanten erfolgreichen Lernens in digitalen Welten, um Gestaltungsmerkmale digitaler Lehr-/Lernprozesse sowie um die Frage nach dem Verhältnis zwischen Lehrenden und Lernenden (oder auch Lernenden untereinander).

Gerade dadurch, dass diese Gestaltungsmerkmale noch vergleichsweise neu sind und gleichzeitig die Entwicklung in Studium und Lehre derzeit stark beeinflussen, ist es vermutlich wichtig, jenseits methodischer „Orthodoxie“ zu den Möglichkeiten und Grenzen wirkungsorientierter Evaluation, zunächst die Entwicklungsprozesse und die Entwicklungen im „Ökosystem“ Hochschulbildung an sich zu beobachten und daraus Theorien über die Effekte neuer Lehr-/Lernformen auf das Lernen abzuleiten. Diese können ebenso dystopisch, wie optimistisch sein (vgl. Pohlenz 2018). Dystopisch in dem Sinne, dass die Individualisierung von Lernprozessen zur Vereinzelung der Lernenden führt und damit die soziale und kommunikative Funktion akademischen Lernens und der Wissenschaft an sich Schaden nimmt. Optimistisch in dem Sinne, dass durch neue Möglichkeiten digitaler Medialität von Bildung neue Räume für eine sozial breit gestreute Bildungsteilnahme eröffnet werden. Hier aus der Beobachtung von Prozessen zu einer Prognose zukünftiger Entwicklungen und daran anschließend, angezeigter Maßnahmen zu gelangen, ist eine wichtige Aufgabe der Evaluation von Hochschullehre.

## 5 Fazit

Evaluation als Mittel der Qualitätssicherung sowie Qualitätsentwicklung hat sich – begleitet von kontroversen Diskussionen über die Rolle, die sie für Steuerungsentcheidung spielen soll sowie hinsichtlich methodischer Fragen zur Validität ihrer Ergebnisse – in den vergangenen Jahren im Alltag der Hochschullehre etabliert.

Insbesondere im Feld von Innovationen der Hochschullehre, beispielsweise in Gestalt von digitalen Lehr-/Lernmedien oder der Individualisierung der Lehre, des Lernens und Prüfens, stellt sich die Herausforderung für Hochschulen, zukünftige Entwicklungen mit ihren Vorteilen und Risiken zu beurteilen. Evaluation kann hier ein hilfreiches Instrument sein, Prognosen über die Entwicklungen anzustellen und für konkrete Managemententscheidungen zu nutzen. Dabei sollten sich die jeweils Verantwortlichen der Limitationen bewusst sein, die aus methodischer bzw. methodologischer Sicht die Aussagefähigkeit von Evaluationsergebnissen einschränken. Zudem ist in der Anforderung, Transparenz hinsichtlich der Nutzungszwecke von Evaluationen – zwischen Legitimationsbeschaffung und interner Selbstdiagnose – zu schaffen und die entsprechenden potenziellen Konflikte zu lösen, eine wichtige Aufgabe für die Gestaltung Nutzen stiftender Evaluationen zu sehen.

## Literatur

- Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation*. Berlin u. a.: Springer.
- Braun, E. (2007). *Das Berliner Evaluationsinstrument für selbsteingeschätzte studentische Kompetenzen (BEvaKomp)*. Göttingen: V&R Unipress.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). *Using randomization in development economics research: A toolkit. Handbook of development economics*, 4, S. 3895–3962.
- Gess, C., Ruess, J., & Deicke, W. (2014). Design-based Research zur Verbesserung der Lehre an Hochschulen: Einführung und Praxisbeispiel. *Qualität in der Wissenschaft*, 1, S. 10–16.
- Giel, S. (2013). *Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen*. Münster u. a.: Waxmann.
- Kromrey, H. (1995). Evaluation. Empirische Konzepte zur Bewertung von Handlungsprogrammen und die Schwierigkeit ihrer Realisierung. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie*, 15, 4, S. 313–336.
- Lange, S. (2008): New Public Management und die Governance der Universitäten. dms – der moderne Staat. *Zeitschrift für Public Policy, Recht und Management*, 1, S. 235–248.
- Pawson, R. (2003). Nothing as Practical as a Good Theory. *Evaluation* 9, 4, S. 471–490.
- Pohlenz, P. (2009). *Datenqualität als Schlüsselfrage der Qualitätssicherung von Lehre und Studium*. Bielefeld: Universitätsverlag Webler.
- Pohlenz, P. (2018, im Erscheinen). *Evaluation von Lehre und Studium in der Hochschule 4.0.*, herausgegeben von Treack, T., & Deickmann, M. Berlin: DUZ-Verlag.
- Schimank, Uwe (2015). Identitätsbedrohungen und Identitätsbehauptung – Professoren in reformbewegten Universitäten. In V. von Groddeck, & S.M. Wilz (Hrsg.), *Formalität und Informalität in Organisationen*, S. 277–296. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Schmidt, U. (2010). Anmerkungen zum Stand der Qualitätssicherung im deutschen Hochschulwesen. In Pohlenz, P. & Oppermann, A. (Hrsg.): *Lehre und Studium professionell evaluieren: Wie viel Wissenschaft braucht die Evaluation?* Reihe Qualität-Evaluation-Akkreditierung, Bd. 3, S. 17–32. Bielefeld: Universitätsverlag Webler.

Wolbring, T. (2013). *Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität*. Frankfurt/Main: Campus.

## **Autor**

**Prof. Dr. Philipp Pohlenz**, Professur für Hochschulforschung und Professionalisierung an der Otto-von-Guericke-Universität Magdeburg. Arbeitsschwerpunkte: Studierenden-forschung, Hochschulentwicklung, Evaluationsforschung, Lehr-/Lernforschung.

E-Mail: [philipp.pohlenz@ovgu.de](mailto:philipp.pohlenz@ovgu.de)