



GRUNDLAGEN: LEHREVALUATION

ABSTRACT

Dieser Grundlagentext behandelt das Thema Lehrevaluation in einem weiten Verständnis und geht über die reine Lehrveranstaltungsevaluation durch Studierende, etwa durch Fragebögen am Semesterende, hinaus. Auch im Zuge der Entwicklung neuer Lehrkonzepte und/oder beim Beforschen der eigenen Lehre kann zu verschiedenen Zwecken unterschiedlich evaluiert werden. Der Text klärt zunächst die grundlegende Bedeutung von „Evaluation“, stellt drei wesentliche Zwecke von Evaluationen vor und führt aus, welche Evaluationsgegenstände man im Kontext der Hochschullehre unterscheiden kann. Im Anschluss werden mögliche Bewertungskriterien und -maßstäbe, Evaluationsformen und Methoden zur Lehrevaluation dargestellt.

AUTOR:IN: Gabi Reinmann (HUL)

DATUM: Juli 2022

LIZENZ:



Creative-Commons Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International Lizenz

VORBEMERKUNG

Wer im Zusammenhang mit Hochschullehre das Wort Evaluation hört, denkt in der Regel an Lehrveranstaltungsevaluationen oder noch spezifischer an Fragebögen, die Studierende zum Semesterende ausfüllen, um eine Veranstaltung zu bewerten. In diesem Grundlagentext soll das Thema Lehrevaluation breiter behandelt werden. Auch wenn Lehrveranstaltungsbefragungen eine wichtige Rolle spielen, so erschöpfen sich die Möglichkeiten der Evaluation in der Hochschullehre darin nicht. Auch im Zuge der Entwicklung neuer Lehrkonzepte und/oder beim Beforschen der eigenen Lehre kann zu verschiedenen Zwecken unterschiedlich evaluiert werden.

1. EVALUATION IM KONTEXT VERSCHIEDENER BEWERTUNGSPRAXEN

Was bedeutet das Wort Evaluation?

Evaluation ist ein Begriff, den man auch in der Alltagssprache gebraucht, um auszudrücken, dass etwas *bewertet* wird. Evaluation im Kontext der Hochschullehre ist ebenfalls ein Akt der Bewertung. In einer verbreiteten Definition von Evaluation heißt es, Evaluation sei „the process of determining the merit, worth, or significance of things“ (Scriven, 1996, p. 158). Um Evaluation in wissenschaftlichen Feldern wie der Hochschullehre von anderen Bewertungspraxen abgrenzen zu können, müssen allerdings noch ein paar Merkmale hinzukommen: (a) eine generelle Systematik und Nachvollziehbarkeit im Vorgehen, (b) Klarheit in der Bestimmung des Gegenstands, der bewertet werden soll, (c) der Einsatz passender Methoden, um eine Datengrundlage zur Bewertung zu erlangen, sowie (d) die Formulierung von Bewertungskriterien und -maßstäben. In der Hochschullehre dienen Evaluationen dazu, Informationen einzuholen, um damit die Qualität von Hochschullehre zu bewerten. In diesem Zusammenhang steht der Begriff der Evaluation mit mindestens zwei weiteren Begriffen und Handlungspraxen in Verbindung: dem studentischen Feedback und dem institutionellen Qualitätsmanagement.

Was hat studentisches Feedback mit Evaluation zu tun?

Es gibt zahlreiche Wege, auf denen man als Lehrperson jenseits einer Evaluation mit den oben genannten Merkmalen Informationen dazu einholen kann, wie gut Hochschullehre ist oder funktioniert: über zufällige Gespräche mit Studierenden, über studentisches Verhalten bei einer Veranstaltung, über sichtbar werdende Leistungen Studierender im Zuge eines Lehrangebots etc. Studierende melden in vielen Situationen direkt oder indirekt zurück, wie gut eine Veranstaltung ist oder läuft. Diese nicht-systematische Form von studentischem Feedback ist mit der Lehrevaluation gewissermaßen verwandt und für die Lehrpraxis nicht minder wichtig als eine geplante Lehrevaluation. Allerdings ist der Übergang vom ungeplanten studentischen Feedback im Lehralltag zu Evaluationen im Prozess (formative Evaluation) fließend. Nicht selten werden formative Evaluationsverfahren auch als Feedback-Methoden bezeichnet.

In welcher Beziehung steht Evaluation zum Qualitätsmanagement?

Die Bewertung ist auch das verbindende Glied zwischen Evaluation und Qualitätsmanagement. In der Regel werden heute Strategien zur Sicherung und Weiterentwicklung von Qualität in der Lehre institutionell verankert. Lehrveranstaltungsevaluationen mittels Fragebögen sind ein kaum wegzudenkendes Element in solchen Systemen des Qualitätsmanagements, die aber darüber hinaus weitere Maßnahmen beinhalten wie Peer Reviews, Hochschulstatistiken, Absolventenstudien etc. (Pohlenz, 2018). Lehrevaluation ist stets Teil des Qualitätsmanagements in der Lehre, aber nicht jede Evaluation in der Lehre muss zwingend darin eingebunden sein.

2. EVALUATIONSZWECKE

In der Evaluationsliteratur ist man sich weitgehend darin einig, dass man mindestens drei Zwecke von Evaluationen unterscheiden kann, nämlich (a) Kontrolle und Legitimation, (b) Verbesserung und Weiterentwicklung sowie (c) Erkenntnis und wissenschaftlichen Fortschritt. Diese schließen sich untereinander nicht unbedingt aus; zudem sind Kombinationen möglich. Es ist aber empfehlenswert, sich jeweils über den *primären Zweck* einer Lehrevaluation klar zu werden, da dieser Einfluss auf nachfolgende Entscheidungen haben kann, wer was wie evaluiert.

Wann liegen Kontrolle und Legitimation als Evaluationszweck vor?

Im Alltagsverständnis liegt es relativ nahe, dass man mit einer Evaluation etwas kontrollieren oder legitimieren will: Man evaluiert, weil man überprüfen will, ob das, was als Ziel gesetzt oder versprochen wurde, auch erreicht werden konnte. Zentral organisierte Lehrveranstaltungsevaluationen dienen häufig der Kontrolle, ob Lehrveranstaltungen z.B. Mindeststandards erfüllen. Ein anderes Beispiel sind Lehrprojekte, für die man Finanzmittel bereitgestellt hat: Evaluationen können hier den Zweck verfolgen, zu erfassen und zu bewerten, wie erfolgreich Projekte waren und/oder ob sich Investitionen „gelohnt“ haben. Man benötigt die Ergebnisse, um weitere Entscheidungen zu treffen (z.B. über Erhalt, Ausbau oder Einstellung einer didaktischen Intervention). Mit Evaluationen unter dem Leitkonzept „Kontrolle/Legitimation“ leistet man vorrangig einen Beitrag zur Planungsrationalität. Adressaten solcher Evaluationen sind z.B. Hochschulleitungen oder Dekanate, gegebenenfalls Geldgeber/Fördereinrichtungen oder auch die (Hochschul-)Politik.

Was bedeuten Verbesserung und Weiterentwicklung als Evaluationszweck?

Verbesserung und Weiterentwicklung sind für die Lehrpraxis und hochschulische Qualitätsbestrebungen ein wesentlicher und häufiger Evaluationszweck. Um Lehre zu verbessern und weiterzuentwickeln, kann man die zugrundeliegenden Konzepte, ausgelöste Aktivitäten und/oder erreichte Ergebnisse evaluieren, um Ansatzpunkte für die weitere Arbeit an der Lehre zu finden. Evaluationszwecke dieser Art können Teil einer reflektierten Haltung in der Lehre werden und im Idealfall das individuelle Lehrhandeln prägen. Auch für Lehrprojekte sind Verbesserung und Weiterentwicklung wichtige Evaluationszwecke: Oft beginnt hier eine Evaluation schon in der Planungsphase und integriert einen Dialog mit den Beteiligten. Mit Evaluationen unter dem Leitkonzept „Verbesserung/Weiterentwicklung“ stehen neue Interventionen oder Problem-

lösungen für die Lehre im Fokus. Adressaten solcher Evaluationen sind unmittelbar die Lehrpersonen selbst, die Lehrangebote gestalten oder Lehrprojekte durchführen, mittelbar aber natürlich vor allem die Studierenden, an die sich die Lehre richtet und für die sie verbessert wird.

Wie können Erkenntnis und wissenschaftlicher Fortschritt Evaluationszweck werden?

Erkenntnis und wissenschaftlicher Fortschritt sind ein dritter Evaluationszweck in der Hochschullehre. In Forschungsprojekten, die Evaluationen zur Evaluationsforschung machen, sind Erkenntnis und hochschuldidaktischer Fortschritt der primäre Zweck. In Projekten, in denen Lehrpersonen ihre eigene Lehre beforschen (Lehrforschung oder Scholarship of Teaching and Learning) und dabei Evaluationen durchführen, verbinden sich mehrere Evaluationszwecke, nämlich: Verbesserung und Weiterentwicklung einerseits sowie Erkenntnis und wissenschaftlichen Fortschritt andererseits. Mit Evaluationen unter dem Leitkonzept „Erkenntnis/wissenschaftlicher Fortschritt“ erweitert man die wissenschaftliche Wissensbasis zur Hochschullehre. Adressaten solcher Evaluationen sind vor allem die wissenschaftlichen Gemeinschaften, die sich mit Hochschullehre beschäftigen, aber auch diejenigen, die Hochschullehre gestalten.

3. EVALUATIONSgegenstände

Was ist ein Evaluationsgegenstand?

Das, was man evaluiert, kann man – analog zu einem Forschungsgegenstand – als Evaluationsgegenstand bezeichnen. Nun könnte man einwenden, dass im Kontext der Hochschullehre doch klar sei, was der Evaluationsgegenstand ist: die Hochschullehre. Bei genauerer Betrachtung aber kann man sich auf sehr unterschiedliche Dinge bzw. Aspekte der Lehre bei einer Evaluation konzentrieren (Kerres, 2021): auf das Lehrangebot an sich (*Input*) und/oder die Bedingungen, unter denen das Lehrangebot gemacht wird (*Kontext*), auf die Aktivitäten im Zuge eines Lehrangebots (*Prozess*) und/oder die damit erzielten Ergebnisse (*Output*); auf die (den Ergebnissen) nachgelagerten Wirkungen bei den Studierenden (*Outcome*) und/oder längerfristige und übergreifende Einflüsse (*Impact*).

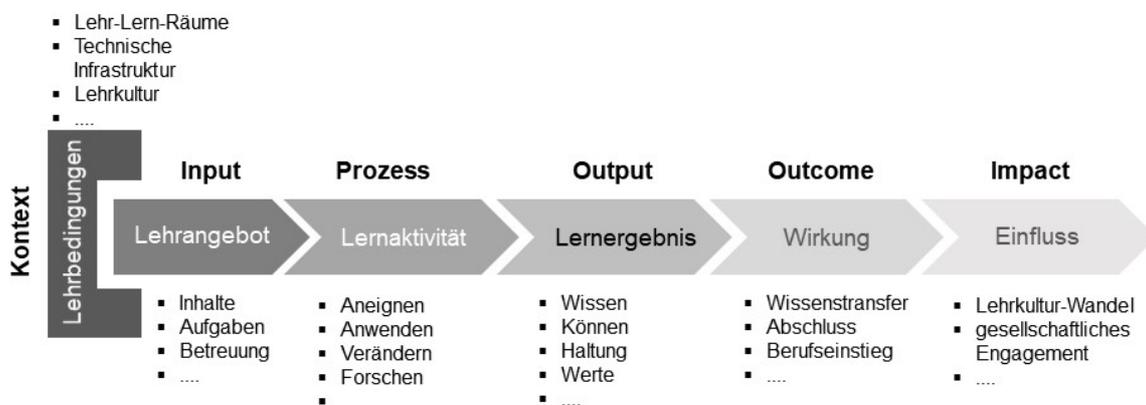


ABBILDUNG 1 Mögliche Ordnung von Evaluationsgegenständen (in Anlehnung an Kerres, 2021).

Eine solche Ordnung von Evaluationsgegenständen (andere sind ebenfalls möglich) orientiert sich an der Prozesslogik von Evaluationen im Rahmen von (Lehr-)Projekten. Je nachdem, welche Gegenstände jeweils im Blick sind, wird man auch unterschiedliche Fragen für die Evaluation haben und verschiedene Bewertungskriterien und -maßstäbe (siehe unten) heranziehen. Es ist nicht immer ganz einfach, diese Aspekte/Gegenstände untereinander abzugrenzen. Mitunter werden die hier erörterten Begriffe auch im Zusammenhang mit Bewertungskriterien behandelt. Letztlich aber sind für die Evaluationspraxis weniger die Details in der Begriffsverwendung entscheidend, als vielmehr der Anspruch, den *Blick* auf die Vielfalt dessen zu *schärfen*, was genau man (wozu) erfassen und bewerten will.

Was gehört zum Lehrangebot und zu Lehrbedingungen als Evaluationsgegenstand?

Gegenstand einer Evaluation kann ein Lehrangebot als solches sein: also die Inhalte und deren Darstellung, dazugehörige Aufgaben zur vertieften Auseinandersetzung, Betreuungsoptionen etc. Anders formuliert geht es hier um die konzeptionelle und materielle Seite eines Lehrangebots: um das, was man in das Lehr-Lerngeschehen bzw. in das potenzielle Bildungsgeschehen einbringt – daher auch die im Rahmen der Evaluation gebräuchliche Bezeichnung *Input*. Man kann diesen „Gegenstand“ als Lehrperson vorab und rückblickend selbst evaluieren (z.B. mit Kriterien-Rastern), dazu Expert:innen einladen oder darauf achten, dass in klassischen Befragungen von Studierenden im Anschluss an eine Lehrveranstaltung genau diese Kriterien erfasst und bewertet werden. Evaluieren kann man ebenso die Bedingungen, unter denen man ein Lehrangebot gemacht hat – kurz: den *Kontext*. Gemeint sind damit beispielsweise Lehr-Lernräume und deren Ausstattung, technische Systeme und Werkzeuge, die für das Lehrangebot bereitgestellt werden, aber auch immaterielle Aspekte wie die in einem Studienfach etablierte Lehrkultur und vieles mehr.

Was umfassen Lernaktivität und Lernergebnis als Evaluationsgegenstand?

Lehrevaluationen können auch die durch ein Lehrangebot angeregten Lernaktivitäten erfassen, um sie zu bewerten. Gemeint sein können damit weitgehend unsichtbare Aktivitäten wie die Aneignung von Wissen durch die Studierenden oder Aktivitäten der Anwendung von Wissen. Denkbar sind auch verschiedene Formen von Veränderung (Transformation) von Wissen durch Studierende bis hin zur Schaffung neuen Wissens etwa im Zuge forschenden Lernens. Es geht hier um *Prozesse* bzw. um die Dynamik im Lehr-Lerngeschehen, die in üblichen Lehrveranstaltungsevaluationen in der Regel keine nennenswerte Rolle spielen. Davon abzugrenzen sind Lernergebnisse – in der Sprache der Evaluation meist (etwas technisch klingend) als *Output* bezeichnet. Im Kontext der Hochschullehre sind damit vor allem Wissen, Können, Haltungen etc. gemeint; die Bezeichnungen variieren; heute ist oft von Kompetenzen die Rede – einem Begriff, bei dem man Wissen, Können und Haltungen wieder zusammenführt, was nicht *immer* sinnvoll ist (Reinmann, 2015). In klassischen Lehrveranstaltungsevaluationen erfährt man über den Output in der Regel nur etwas via Selbsteinschätzungen der Studierenden. Zur Evaluation von Lernergebnissen eignen sich entsprechend andere Evaluationsmethoden oftmals besser, z.B. die Analyse von Wissensartefakten in der Veranstaltung oder von Prüfungsergebnissen.

Wie sind Wirkung und Einfluss als Evaluationsgegenstand zu verstehen?

Im Zusammenhang mit Evaluationen hört man nicht nur von Output als Evaluationsgegenstand, sondern auch von Outcome und Impact. Die Begriffe Output, Outcome und Impact sind im Kontext der Hochschullehre nicht leicht voneinander zu unterscheiden. Unter *Outcome* versteht man Wirkungen als Folge von Lernergebnissen: Ein Beispiel für Outcome liegt vor, wenn Lernende das, was sie sich im Zuge eines Lehrangebots angeeignet haben, auch außerhalb der Hochschule anwenden (können). Allerdings ist ebenso denkbar, dass ein solcher Wissenstransfer bereits ein Lehr-Lernziel ist und zum angestrebten Output zählt. Abschlüsse und erfolgreiche Berufseinstiege wären weitere Outcome-Beispiele, die man evaluieren kann. Wenn man den Blick auf besonders langfristige und über die einzelne Person hinausgehende Einflüsse von Hochschullehre richtet, kann es mitunter sinnvoll sein, von *Impact* zu sprechen: Theoretisch ließe sich das gesellschaftliche Engagement von Studierenden oder Absolvent:innen infolge von Lehrangeboten zum forschenden Lernen oder Service Learning als Impact evaluieren. Denkbar wäre auch eine Impact-Evaluation in dem Sinne, dass man die Veränderung einer Lehrkultur infolge größerer Lehrprojekte erfasst und bewertet.

4. BEWERTUNGSKRITERIEN UND -MASSSTÄBE

Was versteht man unter Bewertungskriterien?

Da sich jede Evaluation durch den Akt des Bewertens auszeichnet, kann man die Frage „Was wird evaluiert?“ nicht nur mit dem Evaluationsgegenstand an sich beantworten. Man muss sich zudem Gedanken dazu machen, *in Bezug auf was genau* der Gegenstand bewertet werden soll: Man kann etwa ein digitales Werkzeug für die Lehre daraufhin analysieren, ob es von Lehrenden und Studierenden genutzt und akzeptiert wird oder ob es einen bestimmten Bedarf deckt. Die Einführung von Self-Assessments in einer Fakultät kann man daraufhin evaluieren, wie wirksam sie für die Verhinderung von Studienabbrüchen ist. Bei Lehrveranstaltungsevaluationen wird Lehre oft daraufhin bewertet, ob und inwieweit Lehrende Ziele transparent machen, strukturiert vorgehen, für Fragen zugänglich und erreichbar sind etc. Neue Lehrformate können daraufhin bewertet werden, wie gut sie ihre Ziele erreichen (Effektivität) und/oder wie angemessen das Verhältnis von Aufwand (der Angebotsgestaltung) und Nutzen (im Sinne von Wirksamkeit) ist (Effizienz). Nutzung, Akzeptanz, Lehrqualität, Bedarfsdeckung, Effektivität, Effizienz, Nachhaltigkeit und vieles mehr können Evaluationskriterien sein.

Wie kommt man zu relevanten Bewertungskriterien für die Lehre?

Bei zentral organisierten Lehrveranstaltungsevaluationen sind Evaluationskriterien über die Items eines Fragebogens meist vorgegeben. Manchmal kann man diese als Lehrperson aber auch anpassen oder ergänzen. Wer als Lehrperson eigene Evaluationen durchführt, kann hingegen selbst entscheiden, welche Kriterien für den jeweiligen Zweck relevant sind. In beiden Fällen sollten Kriterien sinnvoll sein und begründet werden können. Kriterien müssen stets zum Gegenstand passen, dem angestrebten Zweck der Evaluation entsprechen und prinzipiell für eine empirische Überprüfung geeignet sein. Bei der klassischen Evaluation von Lehrveranstaltungen

zieht man oft Kriterien heran, die typische Bedingungsfaktoren für erfolgreiches Lernen oder Bildung im Blick haben: etwa aktive Nutzung, Akzeptanz, Arbeitsaufwand, Merkmale eines Lehrangebots, die neu sind etc. Bei Evaluationen in Lehrprojekten können weitere Kriterien hinzukommen: etwa Bedarfsdeckung, Nachhaltigkeit, Effizienz. Bei Forschungsvorhaben in der Lehre können auch spezifischere Aspekte der Qualität des Lehrkonzepts, der Umsetzung bzw. Implementierung und Ähnliches relevant werden.

Was sind Bewertungsmaßstäbe?

Zur Bewertung von Maßnahmen benötigt man neben Bewertungskriterien immer auch Bewertungsmaßstäbe, denn: Die Frage, wie beispielweise Akzeptanz, Lernerfolg oder Wissenstransfer, Bedarfsdeckung oder Lehrkonzeptqualität zu interpretieren (also positiv, neutral oder negativ einzuschätzen) sind, lässt sich niemals absolut bestimmen. Vielmehr benötigt man *Bezugspunkte*, mit denen man erhobene Evaluationsdaten vergleichen kann. Solche Bezugspunkte zum Vergleich, um zu einer Bewertung zu kommen, nennt man Bewertungsmaßstäbe. Zur eigentlichen Bewertung in einer Evaluation kommt man also immer nur durch Vergleiche. Man kann zwei Gruppen von Vergleichen unterscheiden: (a) Man setzt man Ähnliches zueinander in Beziehung. Das kann man zeitlich tun (Wie war es früher oder vorher und wie ist es jetzt?) oder örtlich (Wie ist es woanders und wie ist es hier?) oder so, dass man verschiedene Interventionen miteinander vergleicht (Was leistet Intervention A im Vergleich zu B, C etc.?). (b) Man vergleicht gesetzte Ziele mit dem Erreichten (Wie waren die Ziele und welche davon sind erreicht worden bzw. wie weit hat man sich diesen angenähert?) oder setzt bestehende Standards als Maßstab (Wie lauten die Standards und wie gut werden diese erfüllt?).

5. EVALUATIONSFORMEN

Was ist mit Evaluationsformen gemeint?

Man kann Evaluationsformen nach unterschiedlichen Gesichtspunkten bestimmen und einteilen: Zwei bekannte und wichtige Unterscheidungen sind die zwischen formativer und summativer Evaluation sowie zwischen Selbst- und Fremdevaluation. Zudem kann man interne von externen Evaluationen unterscheiden sowie Evaluationen durch die Zielgruppe von solchen durch Expert:innen. Oft ist es so, dass der Evaluationszweck sowie der Evaluationsgegenstand bereits einen Pfad gelegt haben, auf dem man sich für eine Evaluationsform entscheidet: Besteht z.B. der primäre Zweck darin, ein größeres Lehrprojekt zu legitimieren, wird man sich für eine Fremdevaluation entscheiden, hat aber noch Spielraum in der Entscheidung für eine interne oder externe Evaluation. Verfolgt eine Evaluation den Zweck, eine Lehrveranstaltung zu verbessern, wird man eine interne Evaluation bevorzugen, sich auf eine Evaluation durch die Zielgruppe konzentrieren, aber vielleicht auch eine Selbstevaluation durchführen. Ein großes hochschuldidaktisches Programm, in das viel Geld investiert wurde, verlangt eher nach einer externen und Fremdevaluation und danach, Expert:innen zur Evaluation heranzuziehen. Ein kleines selbst finanziertes Projekt dagegen wird man intern evaluieren und sich dann für eine Selbst- oder Fremdevaluation entscheiden – ohne zusätzliche Expert:innen.

Wie unterscheiden sich **summative** von **formativen** Evaluationen?

Eine *summative* Evaluation lässt sich durch zwei Merkmale kennzeichnen: Sie findet *nach* Beendigung einer Lehrveranstaltung oder eines Lehrprojekts statt und bezieht sich in der Regel auf ein Lehrangebot (seine Prozesse/Ergebnisse), das bereits *fertig* vorliegt. Oft besteht der Zweck einer summativen Evaluation in der Kontrolle und/oder Legitimation, doch ein zusätzlicher Erkenntnisgewinn wissenschaftlicher Art ist nicht ausgeschlossen. Summative Evaluationen gelten als zusammenfassend, bilanzierend und ergebnisorientiert. Das Pendant zur summativen Evaluation ist die *formative* Evaluation. Diese hat ebenfalls zwei markante Merkmale: Sie findet *während* einer Lehrveranstaltung oder eines Lehrprojekts statt und bezieht sich in der Regel auf einen Gegenstand, der sich noch in *Entwicklung* befindet. Der Zweck einer formativen Evaluation ist in der Regel die Entwicklung und/oder Verbesserung; möglich ist aber auch ein zusätzlicher Erkenntnisgewinn wissenschaftlicher Art. Formative Evaluationen gelten als aktiv-gestaltend, kommunikationsfördernd und prozessorientiert (Stockmann & Mayer, 2010).

Wie unterscheiden sich **Selbst-** von **Fremdevaluationen**?

Bei der *Selbstevaluation* führen diejenigen Personen, welche ein Lehrangebot planen und/oder umsetzen, selbst die Evaluation durch, erheben also Daten, werten diese aus und ziehen daraus Folgerungen. Eine Selbstevaluation ist immer eine interne Evaluation. Trotzdem sind die beiden Begriffe (wie dies oft zu lesen ist) *nicht* gleichzusetzen, da sie sich auf verschiedene Dimensionen beziehen. Selbstevaluationen bieten sich innerhalb von formativen Evaluationen an, können aber auch integraler Bestandteil von Qualitätsmanagementkonzepten sein. Sie sind einerseits fehleranfällig, da man lieber eigene Erfolge als Misserfolge wahrnimmt. Andererseits können sie einen hohen Lern- und Veränderungseffekt auslösen. Bei der *Fremdevaluation* führen Personen die Evaluation durch, die nicht an der Planung und/oder Umsetzung des Lehrangebots oder Lehrprojekts beteiligt sind. Eine Fremdevaluation ist im Rahmen einer externen wie auch internen Evaluation möglich. Umgekehrt gilt aber, dass eine externe Evaluation immer eine Fremdevaluation ist. Fremdevaluationen gelten einerseits als objektiver, weil man die Fehleranfälligkeit von Selbstevaluationen vermeidet. Andererseits können sie Abwehrreaktionen auslösen. Man kann Selbst- und Fremdevaluationen aber auch miteinander verbinden.

Wie unterscheiden sich **interne** von **externen** Evaluationen?

Eine *interne Evaluation* ist eine Evaluation, bei der die evaluierende Person (oder das Evaluationsteam) aus der eigenen Hochschule stammt, in der z.B. ein Lehrprojekt umgesetzt wird. Das muss nicht zwingend mit einer Selbstevaluation zusammenfallen; möglich ist etwa, dass eine zentrale Einheit das Lehrangebot einer Fakultät evaluiert. Interne Evaluationen haben den Vorteil, dass der zeitliche und finanzielle Aufwand eher überschaubar bleibt, eine hohe Sachkenntnis zum Evaluationsgegenstand vorhanden ist und eine große Chance besteht, dass die Evaluationsergebnisse umgesetzt werden. Mögliche Risiken sind eine gewisse Betriebsblindheit, zu wenig Distanz und gegebenenfalls unzureichende Evaluationskompetenz. Eine *externe Evaluation* ist eine Evaluation, bei der die evaluierende Person (oder das Evaluationsteam) nicht zur eigenen Hochschule gehört, in der z.B. ein Lehrprojekt umgesetzt wird. Externe Evaluationen sind stets Fremdevaluationen. Sie haben den Vorzug, dass man Unabhängigkeit leichter

sicherstellen kann; gegebenenfalls erhöht man auch das Vertrauen in die Ergebnisse. Ungünstig kann sich auswirken, dass externe Evaluationsteams weniger (interne) Sachkenntnis haben, ihre Arbeit Abwehr hervorruft und Evaluationsergebnisse nicht anerkannt oder umgesetzt werden.

Worin liegt der Unterschied, wenn die Zielgruppe oder Expert:innen evaluieren?

Bei dieser Unterscheidung geht es um die Frage, wer bei der Evaluation einbezogen wird. Im Kontext der Hochschullehre sind das in der Regel vor allem die Studierenden, was naheliegend und zentral ist, weil es schließlich um ihr Lernen, ihre Bildung, ihre Abschlüsse etc. geht. Je nachdem, was der Evaluationsgegenstand ist, können aber auch andere Personengruppen bei Evaluationen einbezogen werden, etwa Vertreter:innen aus Wirtschaft und Gesellschaft, wenn es um Outcome oder Impact geht. Insbesondere bei Lehrprojekten, die Neuerungen in die Hochschullehre bringen sollen, wird nicht selten auf Expert:innen bei der Evaluation zurückgegriffen – sowohl bei formativen Evaluationen, um Entwicklungen zu verbessern, wie auch bei summativen Evaluationen, um Erfolge zu kontrollieren.

6. METHODEN ZUR LEHREVALUATION

Wie werden Evaluationsmethoden eingeteilt?

Leider werden in der Evaluationspraxis selten systematische Einteilungen von Methoden vorgenommen, die auch methodisch begründet sind. Meist unterscheidet man zwischen Methoden für formative und summative Evaluationen. In diesem Zusammenhang werden sogenannte *qualitative Methoden* (mitunter auch als Feedback-Methoden bezeichnet) in der Regel für formative Evaluationen empfohlen. Man kann diese beispielsweise danach ordnen, ob sie sich für kleinere, mittlere oder große Gruppen eignen; die unterschiedlich weit verbreiteten Methoden haben dann Bezeichnungen wie „Punkten“, „One Minute Paper“, „Änderungsfeedback“ und Ähnliches¹, wobei die genauen Bezeichnungen mitunter variieren. Für summative Evaluationen sind nicht selten zentral gestaltete Fragebögen vorgegeben, die als *quantitative Methoden* bezeichnet werden. Will man sich über diese praktischen Sammlungen, die ihren eigenen Wert haben, darüber hinaus einen umfassenderen Überblick über Möglichkeiten verschaffen, um eigene Veranstaltungen oder Lehrprojekte zu evaluieren oder Evaluationsmethoden in Lehrforschungsprojekten einzusetzen, ist es empfehlenswert, eine Systematik zu wählen, welche empirische Methoden generisch ordnet. Dazu kann man in einem ersten Schritt empirische Methoden zur Erhebung von solchen zur Auswertung von Daten unterscheiden. Im nächsten Schritt lassen sich verschiedene Formen der Erhebung und Auswertung differenzieren.

Was sind typische Methoden zur Datenerhebung in Evaluationen?

In den Bildungswissenschaften – und sicher nicht nur dort – gibt es eine Reihe gängiger Typen von Erhebungsmethoden: Man kann Daten erheben, indem man Phänomene *beobachtet*; dann

¹ Siehe z.B.: <https://www.hul.uni-hamburg.de/team-evaluation/sbnavi-lehrveranstaltungsevaluation/qualitative-feedbackmethoden.html>

hat man Beobachtungsdaten. Man kann Daten erheben, indem man Personen *befragt*; dann erhält man Befragungsdaten. Man kann verschiedene Testverfahren einsetzen, also beispielsweise *testen*, was Personen wissen; es resultieren dann *Testdaten*. Man kann aber auch Artefakte verschiedenster Art sammeln und *analysieren*, etwa schon Aufgabenbearbeitungen oder Problemlösungen von Studierenden im Veranstaltungsverlauf; man könnte das *Analysedaten* nennen. Innerhalb großer Methodengruppen wie Beobachtung, Befragung, Tests und Dokumenten- oder Artefakt-Sammlung und -Analyse kann man weiter differenzieren: Wird frei beobachtet oder befragt oder teilstrukturiert und vollumfänglich strukturiert? Man kann mit verschiedenen Instrumenten beobachten und testen; Befragungen kann man höchst unterschiedlich durchführen – schriftlich, mündlich, einzeln, in Gruppen. Auch bei Dokumenten- oder Artefakt-Analysen gibt es verschiedene Vorgehensweisen. Daraus ergibt sich ein facettenreiches Repertoire an Erhebungsmethoden, die man alle auch für Evaluationen in der Hochschullehre nutzen kann.

Was sind typische Methoden zur Datenauswertung in Evaluationen?

Wer Daten erhebt, muss diese natürlich auch auswerten. Dafür gibt es verschiedene Auswertungsmethoden. Numerische Daten wertet man in der Regel statistisch aus, und auch dazu gibt es zahlreiche spezifische Methoden. Es kommt allerdings auf die Art der Daten an, welche statistischen Auswertungsmethoden überhaupt in Frage kommen. Für die Auswertung verbaler oder anderer nicht-numerischer Daten gibt es ebenfalls zahlreiche Methoden. Manche qualitativen Auswertungsmethoden gehen *reduktiv* vor, versuchen also, die Menge und Komplexität der Daten auf einen Kern zu reduzieren. Andere qualitative Auswertungsmethoden konzentrieren sich im weitesten Sinne auf *Umstrukturierungen*; wiederum andere setzen eher darauf, *expansiv* zu arbeiten, das heißt: Das, was man erhoben hat, wird in der Auswertung weiter entfaltet. Wenn (im Rahmen von Evaluationen) von quantitativen und qualitativen Methoden die Rede ist, bezieht sich das also meist auf die Auswertungsmethoden, die aber wiederum mit der Art der erhobenen Daten zusammenhängen. Auch bei der Auswertung von Daten kann man alle genannten Möglichkeiten im Rahmen von Lehrevaluationen heranziehen.

LITERATUR

- **Kerres, M.** (2021). *Didaktik. Lernangebote gestalten*. Waxmann: Münster.
- **Pohlenz, P.** (2018). Innovationen in der Hochschullehre evaluieren. Herausforderungen für Methodik und Konzeption. In I. Gloinski, A. Browski, J. Gillen, S. Schanze & J. von Meien (Hrsg.), *Kohärenz in der universitären Lehrerbildung. Vernetzung von Fachwissenschaft, Fachdidaktik und Bildungswissenschaften* (S. 385–397). Potsdam: Universitätsverlag Potsdam.
- **Reinmann, G.** (2015). Lehrkompetenzen von Hochschullehrern: Kritik des Kompetenzbegriffs in fünf Thesen. In O. Hartung & M. Rumpf (Hrsg.), *Lehrkompetenzen in der wissenschaftlichen Weiterbildung. Theorie und Empirie Lebenslangen Lernens* (S. 17–36). Wiesbaden: Springer VS.
- **Scriven, M.** (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17, 151–161.
- **Stockmann, R. & Meyer, W.** (2010). *Evaluation. Eine Einführung*. Opladen: Budrich.