II.2.5
Lehrevaluation und Qualitätsentwicklung

Avelina Lovis-Schmidt und Heiner Rindermann

Abstract | Lehrevaluation dient dazu, die Qualität der Lehre messbar zu machen und die Bedingungen sowie Merkmale ihres Erfolges zu erfassen (Diagnostik). Die Durchführung von Lehrevaluation leistet zudem einen Beitrag zur Verbesserung der Lehre (Qualitätsentwicklung). Es werden im Modell der Lehrqualität verschiedene Dimensionen unterschieden, die Lehrerfolg beeinflussen: die Lehrperson (z. B. Verhalten, Kompetenzen, Methodeneinsatz), Studierende (z. B. Kompetenzen, Motivation, aktive Veranstaltungsteilhabe) und Rahmenbedingungen (z. B. Besuchszahl, Lehrdeputat, Veranstaltungsart). Die empirische Forschung zeigt eine ausreichende Validität studentischer Lehrveranstaltungsevaluationen. Um aber positiv veränderungswirksam zu sein, bedarf es weiterer Maßnahmen. Lehrevaluation als Bestandteil von Qualitätsentwicklung sollte deshalb durch Angebote von Beratung, Weiterbildung und Training ergänzt werden.

Stichworte | Hochschullehre, Lehrqualität, Qualitätsverbesserung, Evaluation, Studienqualität

Einleitung: Lehre und Lernen an Hochschulen und die Entwicklung ihrer Qualität

Lehrveranstaltungen an Hochschulen verfolgen das Ziel, Studierenden Wissen zu vermitteln, sodass diese damit verstehend umgehen können. Zudem sollen Einstellungen und Motivationen in eine positive Richtung gelenkt werden (etwa Interesse am Studienfach) und schließlich kommt die Förderung allgemeiner und je nach Studiengang spezifischer praktischer Kompetenzen hinzu (etwa Literatur suchen, Referate halten, Texte schreiben, mit anderen über kontroverse Themen diskutieren, eine Untersuchung durchführen). Insbesondere in der deutschen Tradition ist die Lehre an eine humanistische, aufklärerische und die Fähigkeit zum Selberdenken als Ziel verfolgende Bildungsidee geknüpft (Hubig/Rindermann 2012; Müller 1802/1990).

Lehrevaluation dient der Bestimmung der Lehrqualität im Allgemeinen, Lehrveranstaltungsevaluation der Qualität einzelner Veranstaltungen (oft wird begrifflich die Kurzform „Lehrevaluation" für beides verwendet). Im günstigen Fall ermöglicht Evaluation Qualitätsentwicklung, welche Feedback-Mechanismen und Angebote für die Professionalisierung von Lehrenden beinhaltet. Lehrevaluation kann auch in der Hochschulgovernance genutzt werden, beispielsweise in Entwicklungsgesprächen mit Lehrpersonen oder bei der Akkreditierung.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Besonderheiten der Hochschullehre: Ein multidimensionales Modell von Lehrqualität und Lehrerfolg

Lehre an Hochschulen, abgeleitet vom alten gotischen Wort „laisjan" mit der Bedeutung „wissend machen", beschreibt einen bildenden Prozess, der speziell auf wissenschaftliche Inhalte ausgerichtet ist und sich am fortgeschrittenen Kompetenz- und Wissensniveau der erwachsenen, mehrheitlich jungen Studierenden orientiert. Funktionales Ziel ist es, die Studierenden auf komplexe berufliche Anforderungen vorzubereiten. Während das ‚Unterrichten' in der Schule oder ‚Erziehen' im Kindergarten einer pädagogischen Ausbildung bedarf, weisen Lehrpersonen an der Hochschule üblicherweise (ausschließlich) eine forschungsorientierte Ausbildung mit fachlichem Schwerpunkt auf. Der Besuch hochschuldidaktischer Kurse ist freiwillig und war zumindest bis vor etwa einem Jahrzehnt nicht üblich – trotz Benennung ‚Hochschuldozent' (→ II.2.2 Professionalisierung (in) der Hochschullehre und Hochschuldidaktik). Dies führt mitunter zu einer Diskrepanz zwischen Fach- und Lehrkompetenz (Reinmann 2015), was sich auch darin zeigt, dass Studierende und Dozierende verschiedener Fachkulturen unterschiedliche Qualitätsmerkmale von Lehre nennen (Lovis-Schmidt et al. 2020).

Die Qualitätsmerkmale verteilen sich auf verschiedene Dimensionen, die im *Modell guter Lehre* aufgegriffen werden (s. Abb. II.2.5.1; Rindermann 2009). Neben den *Lehrenden* gehören zu *Lehrerfolg* ebenso die *Lernenden* und *Rahmenbedingungen*. Auf der Zielseite *Lehrerfolg* ist zwischen einer guten Veranstaltung selbst und Änderungen in Merkmalen der Studierenden zu unterscheiden. Das hier verwendete Modell ist eines der wenigen empirisch getesteten Modelle von Lehrqualität (siehe Übersicht in Rindermann 2009), welches den Lehrerfolg ganzheitlich unter Berücksichtigung verschiedener Bedingungsfaktoren versteht. Es bietet Anwendungsbezug durch den Fragebogen zur Evaluation von Lehrveranstaltungen an Hochschulen HILVE-II (Rindermann 2004).

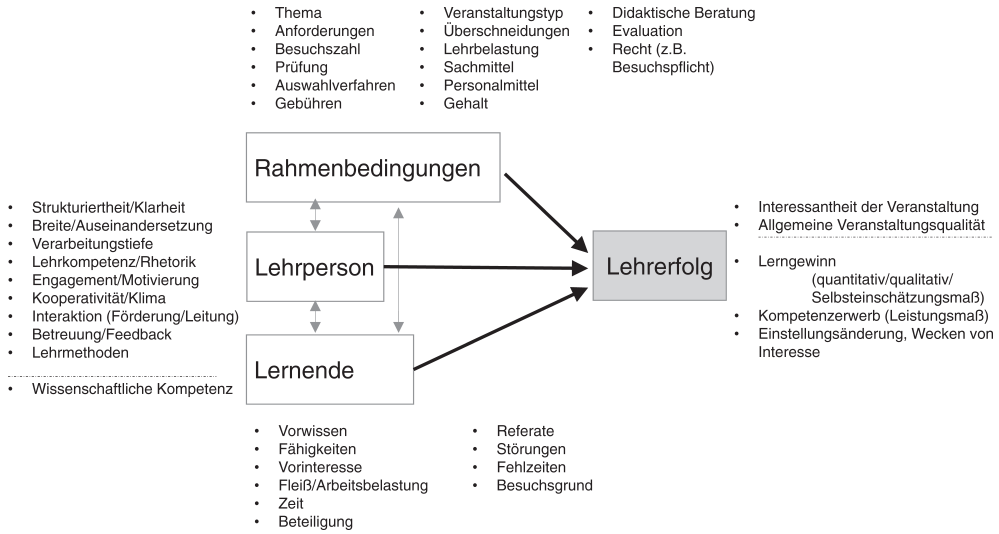
Lehrpersonen tragen zum Gelingen von Lehre bei, indem sie lernförderliche Instruktionsmethoden anwenden. Verschiedene Arbeiten beschäftigen sich mit der Frage, *welche* Merkmale der Lehrperson zu guter Lehre beitragen (Reinmann 2015; Hattie 2023; Wolbring et al. 2013).

Auch die Art der *Interaktion* zwischen Lehrperson und Studierenden nimmt auf den Erfolg von Lehre Einfluss. Studierende sind interessiert an anregenden, vielfältigen Lehrmethoden wie Fragenstellen, Diskussionen, Quiz und Rollenspielen (Bochmann 2018; Lovis-Schmidt et al. 2020). Diese Methoden steigern die studentische Motivation und verlängern die Lernzeit (Bochmann 2018). Relevant für gute Lehre sind weiterhin Bezüge zur Praxis und Verwendung von erläuternden Beispielen. Verarbeitungstiefe soll u. a. eine kritische Analysefähigkeit anregen, etwa durch Betrachtung von Inhalten aus verschiedenen Perspektiven (Lovis-Schmidt et al. 2020). Verarbeitungstiefe könnte durch eine größere Diversität, was weltanschauliche und politische Orientierungen unter Lehrenden betrifft (z. B. Frisby et al. 2023; Hommel 2024), zusätzlich gesteigert werden.

Der Beitrag der *Lernenden* ist ein weiterer wichtiger Bedingungsfaktor für Lehrerfolg. Dazu zählen: Vorwissen und Fähigkeiten (spezifisches Vorwissen, allgemeine Intelligenz, nichtkognitive Studierfähigkeiten und andere für Studium und Beruf relevante Kompetenzen wie soziale Fähigkeiten), Fleiß, Zeit (für das Studium, wenig andere Belastungen), Interesse am Studium und seinen Themen (Vorinteresse), Disziplin und Selbstdisziplin (wenig Störungen und Unruhe), Anwesenheit in Lehrveranstaltungen und veranstaltungsinterne Beteiligung (in der Regel dadurch Erhöhung der Lernzeit), Grund

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Abb. II.2.5.1: Multidimensionales Modell guter Lehre



Quelle: H. Rindermann

der Teilnahme (Interesse vs. Prüfung) und Referatsqualität (z. B. in Seminaren, z. B. Rindermann 2004).

Schließlich sind *Rahmenbedingungen* für den Erfolg der Lehre relevant (auch *Strukturqualität* genannt, siehe auch Rindermann 2016). Dazu gehören das Thema (Schwere, Relevanz, Anregungsgehalt), die Höhe der Anforderungen (Umfang, Geschwindigkeit, inhaltliche Schwere), der Veranstaltungstyp (in Seminaren gibt es mehr studentische Beteiligung) und die Zahl der Teilnehmenden (problematisch bei über 30 bis 40 Personen in einem Seminar), inhaltliche Überschneidungen zu anderen Veranstaltungen (Überlappungen, positiv Anknüpfungspunkte) und die Existenz einer anschließenden Prüfung (erhöht Motivation und Fleiß). Auch äußere Kontextfaktoren sind relevant wie die Existenz und Qualität von Auswahlverfahren an der Institution (Numerus clausus-Beschränkung erhöht studentische Kompetenzen und sichert angemessene Besuchszahlen), Finanzen, beispielsweise Gebühren und Förderung (erhöhen Finanzmittel, steigern den Verpflichtungscharakter, Stipendien erlauben Finanzschwächeren das Studium), Sachmittel (wie für Literatur, Kopien, Beamer, Notebooks), Personalmittel (etwa für Tutorien und studentische Hilfsmittel), Lehrbelastung (z. B. an deutschen Hochschulen höher als in den USA; Griffith/Altinay 2020), Gehalt der Lehrkräfte (führt zu einer positiven Selektion und Motivation) sowie Angebot didaktischer Beratung, Weiterbildung und Evaluation (Messung und Kommunikation von Lehrqualität).

Lehre soll nicht nur Qualität als Merkmal an sich aufweisen, sondern auch positive Wirkungen nach sich ziehen (z. B. Rao 2020). Gute Lehre führt vor allem zu *Lernen*; die Studierenden haben danach mehr Wissen und dies kann – valider als durch Selbsteinschätzungen – insbesondere durch Leistungstests festgestellt werden (bspw. im Prä-Post-Design). Neben Erwerb von deklarativem Wissen wird auch der Erwerb von *Kompetenzen* in verschiedenen Bereichen (Anwendung von Wissen, Verste-

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

hen, Problemlösen, je nach Fach auch soziale, emotionale und motorische Kompetenzen) angestrebt (→ II.1.7 Kompetenzorientierung). Schließlich ist es oft Ziel, *Einstellungen* zu verändern, z. B. Interesse an Themen des Studiums, an der Wissenschaft, am selbständigen Denken, an Inhalten der späteren Berufspraxis einschließlich ethischer Orientierungen.

Die skizzierten Elemente stehen in Wechselwirkung, beeinflussen sich also gegenseitig. Beispielsweise erhöht eine hohe Veranstaltungsqualität das studentische Interesse und die Anwesenheit (z. B. Bochmann et al. 2019; Wolbring 2013), und motivierte wie kompetente Studierende erhöhen die Lehrqualität (Rindermann 2016). Die Qualität der Rahmenbedingungen verändert den Lehrerfolg direkt und indirekt über eine Einflussnahme auf Lehrende und Studierende (Rindermann 2016).

Einsatz von Lehrevaluationen zur Entwicklung von Qualität

Im Rahmen von Qualitätsentwicklung spielt Lehrevaluation (Lehre inkl. Curricula, Betreuung von Abschlussarbeiten etc.) eine wichtige Rolle. Daneben haben sich, laut einer bibliometrischen Studie von Steinhardt et al. (2017), auch andere Cluster im Qualitätsmanagement in den letzten 30 Jahren etabliert: 1) "Assessment-Cluster", welches sich mit der Bewertung von Lernergebnissen beschäftigt, 2) "Quality-Cluster", welches die allgemeine Qualitätssicherung mithilfe von Richtlinien und Programmen umfasst, eng damit verbunden 3) "Quality-Management-Cluster", welches sich speziell auf die Planung und Verbesserung von Qualitätsprozessen konzentriert und zuletzt wie oben bereits erwähnt, 4) "Student-Evaluation-of-Teaching-Cluster" unter Einbezug von Studierendenfeedback.

Schritte der Lehrevaluation

(Lehr-)Evaluation beginnt zunächst mit einer *Festlegung der Untersuchungsgegenstände* und der *Bestimmung der Qualitätsmerkmale* wie etwa den studentischen Lernerfolg. Darauf folgt die *Auswahl oder Entwicklung geeigneter Evaluationsinstrumente*, sowohl *quantitativ* mittels Bewertungsskalen als auch *qualitativ* durch offene Fragen. Eine *klare Zielsetzung* definiert die Ziele der Evaluation wie die Identifizierung von Stärken und Schwächen oder die Anpassung von Lehrinhalten. Eine hohe und vor allem repräsentative Teilnehmerzahl ist relevant (*Stichprobenplanung*); dazu trägt die Zusicherung von Anonymität bei. Eine *zeitnahe Auswertung der Ergebnisse* ermöglicht ein schnelles Feedback und die Einleitung möglicher Maßnahmen. Das *Feedback* an die Lehrenden sollte mit Stärken beginnen, dann verbesserungswürdige Bereiche identifizieren und mit der Anregung spezifischer Maßnahmen enden. Die *Implementierung von Verbesserungen* sollte gesichert werden. Langfristig ist Lehrevaluation als ein Element in institutionelle Qualitätsentwicklung zu integrieren.

Akteure der Lehrevaluation

Durch systematische Rückmeldungen (von Studierenden, seltener Selbsteinschätzungen der Lehrperson und Fremdraterurteile) kann eine Informationsquelle geschaffen werden, die dabei hilft, Lehre

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

zu verbessern und Lernerfolg zu erhöhen. Der Evaluationsprozess sollte dabei die Reflexion über die eigene Lehre und – falls Ergebnisse besprochen werden – den Dialog zwischen Lehrenden und mit Lernenden über die Lehre erleichtern. Lehrende können sich (in einem vertrauensvollen Verhältnis) auch gegenseitig evaluieren und Feedback geben (z. B. im Rahmen von Microteaching; Hattie 2023). Die Einbindung aller Beteiligten in den Evaluationsprozess könnte eine offene Lehrkultur und kollaboratives (zusammenarbeitendes) Lernen fördern.

Mögliche Probleme der Lehrevaluation

„*One need not talk with faculty very long to be aware of their concern about possible biases in student ratings*" (Cashin 1995: 6).

Die Hochschulforschung zu Lehrevaluation widmet sich schon seit geraumer Zeit der Frage, ob Studierende imstande sind, die Qualität von Lehre zu beurteilen. Hinzu kommt die Frage, wie Lehrevaluation einen positiven Effekt auf Lehrqualität haben kann. Folgende Überlegungen werden dazu geäußert:

(1) *Absentismus reduziert die Aussagekraft von Lehrevaluation* (und verschlechtert Lernergebnisse): Studentische Lehrevaluationen können durch selektives Besuchsverhalten von Veranstaltungen verzerrt werden, wenn etwa am Ende des Semesters nur noch eine kleine Gruppe, zusammengesetzt aus vermutlich eher motivierten Studierenden, präsent ist (*dropout bias*; Hinz 2012; Rehm 2023; Wolbring 2013). Besser wäre es, auch die (wohl eher kritischeren) Meinungen derjenigen zu berücksichtigen, die den Kurs vorzeitig verlassen haben. Eine Lösung besteht in der Durchführung von Lehrevaluationen nicht am Ende, sondern in der Mitte des Semesters (Klinkhammer/Keller 2022). Eine kontinuierliche (digitale) Evaluationsmethode ermöglicht zum Beispiel ‚Gamification' unter Umsetzung spielerischer Wettbewerbsaufgaben, welche die studentische Teilhabe und Motivation steigern (Raichle 2016). Abgesehen vom Problem der Evaluationsverzerrung verkürzt Absentismus auch die Lernzeit und erschwert eine Tiefenverarbeitung des Lehrinhalts (Bochmann et al. 2019; Rehm 2023). Absentismus ist unter anderem eine Folge behördlicher Vorgaben in den letzten 20 Jahren (Anwesenheit darf nicht mehr kontrolliert werden).

(2) *Studentische Aussagen können durch Biasvariablen verzerrt sein:* Unter *Bias* ist zu verstehen, dass Merkmale, die nicht für die Qualität der Lehre stehen, das Urteil der Studierenden beeinflussen. Allerdings ist nicht immer klar, was Ursache und was Wirkung ist: In eigenen Studien (Rindermann 2009; 2016) zeigten sich bessere Lehrevaluationsergebnisse bei höherer *Popularität* der Lehrkraft (r=.57), interessanterem *Thema* (r=.43, HILVE-I, r=.19, HILVE-II) und Interesse als *Besuchsgrund* (r=.28). Alle diese Merkmale können aber durch das Verhalten von Lehrkräften beeinflusst werden, etwa bei besserer Lehre höhere Popularität, Besuch einer Veranstaltung wegen der Lehrkraft und Wahrnehmung des Themas als interessant.

Merkmale der *Lehrperson* könnten dazu beitragen, dass studentische Lehrevaluationen verzerrt sind, bspw. indem sie weitaus besser beurteilt wird, wenn sie sympathisch oder attraktiv erscheint (nach einer Übersicht von Wolbring 2013: 120). Dies kann u. a. mit dem *Halo-Effekt* erklärt werden, wonach ein (dominierendes) Merkmal der Person die Wahrnehmung von anderen relevanten Merkmalen be-

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

einflusst (Cannon/Cipriani 2024). Attraktivität führt aber auch zu mehr Aufmerksamkeit und kann so Lernen fördern („Produktivitätseffekt", Wolbring 2013). Damit würden Studierende mehr lernen, der Einfluss des Merkmals wäre aber „unfair", weil es wenig durch die Lehrperson selbst veränderbar ist. Auch *milder benotende Lehrpersonen* erhalten im Schnitt etwas bessere Evaluationsresultate (Berezvai et al. 2021; nach einer Übersicht von Wolbring 2013: 120).

Falls Korrelationen mit Geschlecht und Hierarchieposition beobachtet werden, müssen diese nicht zwingend für Verzerrungen stehen, können doch zwischen Männern und Frauen und Professoren und Mittelbauangehörigen Unterschiede in der Lehre bestehen (Benton/Cashin 2012). Mittelbauangehörige bspw. unterscheiden sich weniger in Alter und Habitus von Studierenden und unterrichten mehr in Seminaren, was in Kommunikationsaspekten zu besseren Bewertungen führt. Gut unterrichtende Lehrpersonen werden mit der Zeit beliebter und ihre Veranstaltungen intensiver besucht, wodurch Lehrkompetenz Interesse während der Veranstaltung für verschiedene Themen wecken kann. Ein weiteres oft untersuchtes Merkmal ist die *Entertainmentqualität*, wonach unterhaltsame (aber teilweise inhaltsleer unterrichtende) Lehrpersonen bessere Evaluationsresultate bekommen. Allerdings ist Aufmerksamkeitsstimulation (ein Aspekt von Entertainment) auch ein Merkmal guter Lehre (Rindermann 2009: Tabelle 9.6).

Auch *Merkmale der Lernenden* selbst können die Lehrevaluation beeinflussen, bspw. Persönlichkeitsfaktoren wie Gewissenhaftigkeit oder Tendenzen im Antwortverhalten, z. B. Tendenz zur Mitte. Nach eigener Forschung (Rindermann 2009: Kapitel 9.3; Rindermann 2016) korrelieren *biographische Variablen* (Alter, Geschlecht, Abiturnote etc.) wenig mit Lehrevaluationsresultaten, durchaus aber *Interesse*. Erhebt man spezifischer studentisches *Vorinteresse* („Ich habe mich schon *vor* dem Kurs sehr für die Themen interessiert."), sinkt die Korrelation mit eingeschätzter Lehrkompetenz, was darauf schließen lässt, dass Vorinteresse weniger eine Verzerrungsvariable zu sein scheint als Interesse eine Folge guter Lehre.

Es besteht weiterhin die Befürchtung (Klinkhammer/Keller 2022), Studierende würden Lehre und Lehrkräfte nur global positiv oder negativ beurteilen. In der Tat sind im Vergleich studentischer Einschätzungen mit Dozentenselbsturteilen erstere stärker von einem Generalfaktor der wahrgenommenen Lehrqualität geprägt (alle Aspekte einer Lehrveranstaltung werden stärker einheitlich gut oder schlecht bewertet). Trotz eines gewissen Generalfaktors kann aber die Multidimensionalität studentischer Urteile festgestellt werden (z. B. Marsh 2007; Rindermann 2009).

Rahmenbedingungen betreffen mögliche Verzerrungen durch gegebene Umstände der Lehre. Beispielsweise wird angenommen, dass Studierende solche Veranstaltungen besser bewerten, die sie freiwillig besuchen oder deren Themen sie vorab schon interessieren.

(3) *Geringe Nützlichkeit von Lehrevaluation*: Oft wird befürchtet, die Durchführung von studentischer Lehrevaluation würde wenig zur Verbesserung der Qualität der Lehre beitragen. Die Resultate der Forschung sind recht klar (Marsh 2007: 358ff.; Rindermann 2009: Tabelle 11.1; Rindermann 2023): Studentische Lehrevaluation ohne unterstützende Maßnahmen wie Beratung, Training oder Weiterbildung hat nur einen geringen Effekt; Lehrevaluation mit Beratung oder Trainings zeigt dagegen deutlich positive Effekte.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Wie lässt sich die Validität studentischer Lehrevaluation bestimmen?

Drei unterschiedliche Methoden zur Bestimmung der Validität lassen sich unterscheiden:

(a) *Korrelation mit Biasvariablen:* Angenommene Biasvariablen werden miterfasst. Finden sich Korrelationen, die für Verzerrungen stehen, versucht man, diese statistisch zu kontrollieren (bspw. sehr hohe Besuchszahlen im Seminar).

(b) *Zusammenhang mit Fremdeinschätzungen*: Wenn studentische Urteile mit denen kompetenter anderer Personen übereinstimmen, spricht dies für ihre Validität. Es zeigte sich, dass studentische Urteile mit denen von Beobachtern oder Kollegen übereinstimmen ($r_{S\text{-}B}$=.53)[1], weniger jedoch mit Dozentenselbsteinschätzungen ($r_{S\text{-}D}$=.29), und am geringsten sind die Korrelationen zwischen Dozierenden und externen Beurteilern ($r_{D\text{-}B}$=.19, Metaanalyse Feldman, 1989). Die Ergebnisse wurden in eigenen Studien bestätigt ($r_{S\text{-}B}$=.57, $r_{S\text{-}D}$=.28 und $r_{D\text{-}B}$=.29, Rindermann 2004).

(c) *Zusammenhang mit Maßen des Lernerfolgs*: Lernerfolg, gemessen über Leistungstests, ist ein wichtiges und objektives, aber auch distales, weiter entferntes Kriterium. Es ist ein indirekter Indikator des Lehrerfolgs, da Prüfungsleistung nicht nur von Lehrqualität abhängt, sondern auch von Eingangs- und Rahmenbedingungen wie der Relevanz von Prüfungsresultaten, die den Fleiß der Studierenden beeinflussen. Metaanalytisch betrachtet gehen Lehrevaluationsergebnisse mit Leistungsmaßen im mittleren Bereich einher (r=.33, d'Apollonia/Abrami, 1997, mit um die r=.20 geringer bei Uttl et al. 2017). Die Zusammenhänge fallen höher aus bei Reliabilitätskorrektur (r=.47, d'Apollonia/Abrami, 1997, bzw. je nach Mittelbildung bis zu r=.41, Clayson 2009).[2]

Lehrevaluation mit Beratung von Lehrpersonen ergänzen

Beratung und Weiterbildung für Lehrpersonen an Hochschulen verfolgen die Ziele, Lehrqualität und studentisches Lernen zu steigern. Beratung beinhaltet didaktische Aspekte, die auf der Lehrevaluation beruhen, bspw. Vielfalt an Lehrmethoden, interaktionsbezogene Komponenten wie Kommunikation, Gesprächsführung und Studierendenorientierung sowie personenbezogene Komponenten wie Perspektivwechsel und Auftreten (für einen Überblick zu Beratung, siehe Schubert et al. 2019 oder etwas spezifischer Böss-Ostendorf/Senft 2018). Beratung ist im mittleren Bereich[3] erfolgreich für die Qualitätssteigerung studentischer Evaluation (d=0.69, Metaanalyse von Penny/Coe 2004). Nach einer Studie von Hampton und Reiser (2004) zeigen sich positive Effekte nicht nur in der eingeschätzten Lehrqualität, festgemacht an Lehrevaluationsergebnissen, sondern ebenso in Leistungstestergebnissen, wenn auch im Vergleich etwas geringer (Lehre d=0.65 und Leistung d=0.46).

1 Nach Fisseni (2004) kann man bei Korrelationen zwischen r=.40 und .60 von mittlerer Validität sprechen, die damit für studentische Urteile gegeben ist.

2 „reliabilitätskorrigiert" heißt, dass (geringe) Reliabilität beachtet wurde und der „wahre" Zusammenhang zwischen Merkmalen berechnet wird – als ob die Reliabilität perfekt wäre (auch genannt: Minderungskorrektur, Attenuationskorrektur).

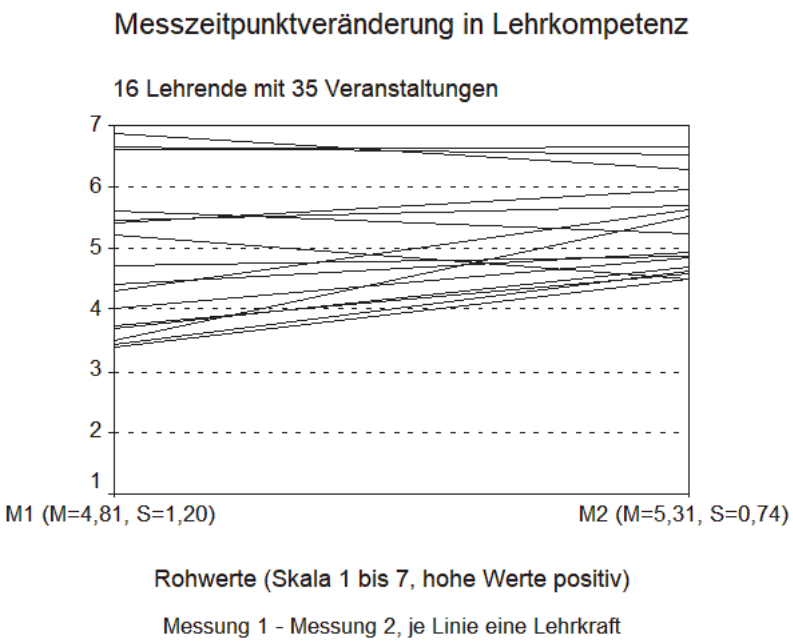
3 Nach den gängigen Interpretationsschwellen (d=0.20 ist kleiner Effekt, d=0.50 mittlerer und d=0.80 großer) ist dies ein mittlerer Effekt.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Im hochschuldidaktischen Kontext ist es heutzutage üblich, Beratungen nicht nur von extern durchführen zu lassen, sondern auch in strukturierten und auf Vertrauen basierenden Einheiten im Kollegium Lehre zu reflektieren und auf Grundlage von Feedback zu verbessern (auch *Peer-Beratung* genannt, siehe Rohr et al. 2016). Der Ansatz der Peer-Beratung wird in der effektiven Methode des Micro-Teachings aufgegriffen (Hattie 2023). Laut dieser Methode wird die Lehrqualität mithilfe von Videosequenzen aus der Veranstaltung eingeschätzt und dann offenes und ehrliches Feedback unter gleichberechtigten Kollegen ausgetauscht, was nachhaltig auch zu einer Verbesserung der Lehre führt.

Grenzen des Beratungsansatzes sind aber auch erkennbar: Unterschiede zwischen Lehrkräften bleiben als Rangreihen trotz Intervention stabil (Rindermann/Kohler 2003). Lehrkräfte verbessern zwar ihre Lehre bzw. die Lehrevaluationsresultate (Mittelwert *M* steigt von 4.81 auf 5.31) und die absoluten Unterschiede werden kleiner (Standardabweichung *S* oder *SD* sinkt von 1.20 auf 0.74). Die relativ Schwächeren bleiben aber die relativ schwächer Beurteilten und bei zuvor schon engagierten und guten Lehrkräften gibt es kaum Veränderungen (Bochmann 2018; s. a. Abb. II.2.5.2 hier). Da es Ziel ist, die Lehrkompetenz der schwach beurteilten Personen zu verbessern, ist dies aber ein insgesamt guter Effekt.

Abb. II.2.5.2: Veränderungen durch Beratung (d=0.94) versus Stabilitäten relativer interindividueller Unterschiede (r=.80), Mittelwerte M, Standardabweichungen S



Quelle: H. Rindermann

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Soll im Rahmen von Qualitätsentwicklung über Lehrevaluation die Lehrqualität verbessert werden, ist Lehrevaluation mit Beratung, Training oder Weiterbildung zu ergänzen.[4] Liegen Defizite in Rahmenbedingungen vor (Rindermann 2016), welche die Lehrqualität beeinträchtigen, müssen diese angegangen werden. Aus funktionalen wie ethischen Gründen ist es nicht sinnvoll, Evaluation durchzuführen, ohne Personen und Institutionen dabei zu unterstützen, Qualität zu verbessern.

Fazit

Lehrevaluationen spielen eine wichtige Rolle bei der Verbesserung der Lehrqualität an Hochschulen. Durch das Feedback von Studierenden können Lehrende Stärken identifizieren und Schwachstellen erkennen und dann gezielte Maßnahmen zur Verbesserung der Lehre einleiten. Gute Lehre zeichnet sich aus durch klare Struktur, Verarbeitungstiefe, Verwendung verschiedener didaktischer Methoden, angemessene Anforderungen und Aktivierung der Lernenden. Neben der Vermittlung von Wissen fördert sie Lernen, Verstehen und Anwenden (und entsprechende Kompetenzen) sowie soziale, emotionale, berufsbezogene, praktische und motorische Fähigkeiten. Auch die Merkmale der Lernenden wie kognitive, soziale und persönlichkeitsbezogene Voraussetzungen beeinflussen den Lehr- und Lernprozess. Lehrqualität erfordert nicht nur das Engagement der Lehrenden und Lernenden, sondern auch eine geeignete Ausgestaltung des Curriculums und der hochschulischen Rahmenbedingungen einschließlich Studierendenauswahl, Lehrdeputat und Gehalt. Um ein möglichst vollständiges Bild der Lehre zu erhalten, sollten zusätzlich zu studentischen Einschätzungen und quantitativen Bewertungen qualitative Methoden und Peer-Reviews einbezogen werden. Lehrevaluation ist ein Baustein in einem Konzept der Qualitätsentwicklung und sollte durch Maßnahmen wie Beratung oder Weiterbildung begleitet werden. Letztendlich sollten Lehrevaluationen Teil eines breiteren kollaborativen Ansatzes sein, der auf kontinuierlicher Weiterentwicklung basiert.

Literaturempfehlungen

Knödler, Elisa (2019): Evaluation an Hochschulen: Entwicklung und Validierung eines verhaltensbasierten Messinventars zur studentischen Lehrveranstaltungsevaluation. Heidelberg/Berlin: Springer. *Das Buch adressiert neben der Frage nach der Operationalisierung von Lehrqualität auch die der Validität studentischer Lehrevaluation. Knödler entwickelte ein verhaltensbasiertes Messinventar, welches in dem Buch anhand verschiedener Teilstudien vorgestellt wird.*

Wolbring, Tobias (2013): Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität. Frankfurt a. M.: Campus. *Tobias Wolbrings Dissertation untersucht Einflüsse der Notengebung und Dozierendenattraktivität auf studentische Lehrveranstaltungsevaluationen sowie individuelle Unterschiede in der Urteilsstrenge der Studierenden und Selektionseffekte. Er betont die hochschulpolitischen Implikationen dieser Ergebnisse.*

4 Effekte von Beratung sind besser untersucht als die von Training oder Weiterbildung. Wir gehen davon aus, dass Training oder Weiterbildung (als umfassendere Maßnahmen als Beratung) ähnlich wirksam sind wie Beratung. Wenn aber Training oder Weiterbildung sich etwa nicht auf relevante Merkmale von Lehrqualität beziehen und zu weit weg von den jeweiligen Unterrichtsformen und Fachgegebenheiten sind, könnte diese Annahme nicht richtig sein.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Noller, Jörg/Beitz-Radzio, Christina/Kugelmann, Daniela/Sontheimer, Sabrina/Westerholz, Sören (2019): Methoden in der Hochschullehre: Interdisziplinäre Perspektiven aus der Praxis. Heidelberg/Berlin: Springer. *Das Buch untersucht, welche Lehrmethoden im breiten Fächerspektrum der Hochschulen besonders geeignet sind und präsentiert bewährte Ansätze zur Qualitätsentwicklung und Verbesserung der Hochschullehre.*

Literaturverzeichnis

Benton, Stephen L./Cashin, William E. (2012): Student ratings of teaching: A summary of research and literature. Manhattan, Kansas: The IDEA Center, IDEA Paper 50.

Berezvai, Zombor/Lukáts, Gergely D./Molontay, Roland (2021): Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. In: Assessment & Evaluation in Higher Education 46(5), 793–808. DOI: 10.1080/02602938.2020.1821866.

Bochmann, René (2018): Einsatz didaktischer Methoden, Interaktionsverhalten und Wirksamkeit pädagogisch-psychologischer Beratungen von Lehrenden an deutschen Hochschulen. Dissertation. TU Chemnitz.

Bochmann, René/Roepke, Avelina L./Reiher, Monique/Rindermann, Heiner (2019): Mangelnde Anwesenheit in Vorlesungen: Eine fächerübergreifende Einschätzung von Studierenden in Deutschland. In: Die Hochschullehre 5, 201–222.

Böss-Ostendorf, Andreas/Senft, Holger (2018): Einführung in die Hochschul-Lehre: Der Didaktik-Coach. Stuttgart: utb.

Cannon, Edmund/Cipriani, Giam Pietro (2024): Quantifying halo effects in students' evaluation of teaching. In: Assessment & Evaluation in Higher Education 49(1), 66–71. DOI: 10.1080/02602938.2023.2180484.

Cashin, William E. (1995): Student ratings of teaching: the research revisited. Manhattan: Center for Faculty Evaluation and Development (IDEA No. 32).

Clayson, Dennis E. (2009): Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. In: Journal of Marketing Education 31, 16–30. DOI: 10.1177/0273475308324086.

d'Apollonia, Sylvia/Abrami, Philip C. (1997): Navigating student ratings of instruction. In: American Psychologist 52, 1198–1208. DOI: 10.1037/0003-066X.52.11.1198.

Feldman, Kenneth A. (1989): Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. In: Research in Higher Education 30, 137–194.

Fisseni, Hermann-Josef (2004): Lehrbuch der psychologischen Diagnostik. Göttingen: Verlag für Psychologie.

Frisby, Craig L./Redding, Richard E./O'Donohue, William T./Lilienfeld, Scott O. (Hg.). (2023): Ideological and political bias in psychology. Cham: Springer.

Griffith, Andrew/Altinay, Zeynep (2020): A framework to assess higher education faculty workload in U.S. universities. In: Innovations in Education and Teaching International 57(3). 1–10. DOI: 10.1080/14703297.2020.1786432.

Hampton, Scott/Reiser, Robert (2004): Effects of a theory-based feedback and consultation process on instruction and learning in college classrooms. In: Research in Higher Education 45, 497–527. DOI: 10.1023/B:RIHE.0000032326.00426.d5.

Hattie, John (2023): Visible learning, the sequel: A synthesis of over 2,100 meta-analyses relating to achievement. London: Routledge.

Hinz, Arnold (2012): Lehr- und Unterrichtsevaluationen durch Studierende und Schüler mittels Ratingskalen. Valide und nützlich oder verzerrt und schädlich? In: Journal für Psychologie 20(3), 1–26.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Hommel, Bernhard (2024). Dealing with diversity in psychology: Science or ideology? In: Perspectives on Psychological Science 19(3), 558–563. DOI: 10.1177/17456916241236170.
Hubig, Christoph/Rindermann, Heiner (2012): Bildung und Kompetenz. Göttingen: Vandenhoeck & Ruprecht.
Klinkhammer, Dennis/Keller, Katrin (2022): Evaluationsansätze für digitale Hochschullehre. Holger Angenent/Jörg Petrie/Tatiana Zimenkova (Hg.), Hochschulen in der Pandemie. Bielefeld: transcript, 264–279. DOI: 10.25656/01:24614.
Lovis-Schmidt, Avelina/Bochmann, René/Reiher, Monique/Baumeister, Antonia/Rindermann, Heiner (2020): Merkmale guter Vorlesungen: Studierende und Dozierende im fachkulturellen Vergleich. In: Die Hochschullehre 6, 402–412. DOI: 10.3278/HSL2027W.
Marsh, Herbert (2007): Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In: Perry, Raymond/Smart, John (Hg.), The scholarship of teaching and learning in Higher Education: An evidence-based perspective. Dordrecht: Springer, 319–383. DOI: 10.1007/1-4020-5742-3_9.
Müller, Ernst (1802/1990): Gelegentliche Gedanken über Universitäten von J. J. Engel, J. B. Erhard, F. A. Wolf, J. G. Fichte, F. D. E. Schleiermacher, K. F. Savigny, W. v. Humboldt und G. F. W. Hegel. Leipzig: Reclam.
Penny, Angela/Coe, Robert (2004): Effectiveness of consultation on student ratings feedback: A meta-analysis. In: Review of Educational Research 74, 215–253. DOI: 10.3102/00346543074002215.
Raichle, Nico (2016): Quests, Raids. Level Up ... Game Over?! Erfolgsfaktoren von Gamification in der Hochschullehre. In: Pfau, Wolfgang/Baetge, Caroline/Bedenlier, Svenja Mareike/Kramer, Carina/Stöter, Joachim (Hg.), Teaching Trends 2016. Digitalisierung in der Hochschule. Münster: Waxmann, 221–234. DOI: 10.25656/01:18901.
Rao, Nalam (2020): Outcome-based education: An outline. In: Higher Education for the Future 7(1), 5–21. DOI: 10.1177/2347631119886418.
Rehm, Michaela (2023): Recht auf schlechte Lehre? Zur Debatte um studentische Partizipation. In: Forschung & Lehre 30(2), 114–116.
Reinmann, Gabi (2015): Lehrkompetenzen von Hochschullehrern: Kritik des Kompetenzbegriffs in fünf Thesen. In: Egger, Rudolf/Wustmann, Cornelia/Karber, Anke (Hg.), Forschungsgeleitete Lehre in einem Massenstudium, Lernweltforschung. Wiesbaden: Springer, 55–75. DOI 10.1007/978-3-658-08869-9_2.
Rindermann, Heiner (2004): HILVE-II in einer computerbasierten Form mit Normen, individualisierter, ergebnisabhängiger Rückmeldung, mit Interpretationshilfen und Beratungsvorschlägen zur Verbesserung der Lehre und mit automatisierter Auswertung und Ergebniszustellung über EvaSys. Lüneburg: Electric Paper.
Rindermann, Heiner (2009): Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Landau: Empirische Pädagogik.
Rindermann, Heiner (2016): Lehrveranstaltungsevaluation an Hochschulen: Der Einfluss der Rahmenbedingungen auf Qualität von Lehre und Ergebnisse von Lehrevaluation. In: Großmann, Daniel/Wolbring, Tobias (Hg.), Evaluation von Studium und Lehre: Grundlagen, methodische Herausforderungen und Lösungsansätze. Wiesbaden: Springer VS, 227–262. DOI: 10.1007/978-3-658-10886-1_7.
Rindermann, Heiner (2023): Lehr- und Unterrichtsqualität. In: Betz, Tanja/Feldhoff, Tobias/Bauer, Petra/Schmidt, Uwe/Schmidt-Hertha, Bernhard (Hg.), Handbuch Qualität in pädagogischen Feldern: Diskurse. Theoretische Grundlagen. Empirische Beiträge. Kritische Einwürfe. Wiesbaden: Springer, 1–19.
Rindermann, Heiner/Kohler, Jürgen (2003): Lässt sich die Lehrqualität durch Evaluation und Beratung verbessern? Überprüfung eines Evaluations-Beratungs-Modells. In: Psychologie in Erziehung und Unterricht 50, 71–85.
Rohr, Dirk/den Ouden, Hendrik/Rottlaender, Eva-Maria (2016): Hochschuldidaktik im Fokus von Peer Learning und Beratung. Weinheim: Beltz Juventa.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Schubert, Franz-Christian/Dirk, Rohr/Zwickel-Pelzer, Renate (2019): Beratung: Grundlagen, Konzepte, Anwendungsfelder. Berlin/Heidelberg: Springer.

Steinhardt, Isabel/Schneijderberg, Christian/Götze, Nicolai/Baumann, Janosch/Krücken, Georg (2017): Mapping the quality assurance of teaching and learning in higher education: The emergence of a specialty? In: Higher Education 74(2), 221–237. DOI: 10.1007/s10734-016-0045-5.

Uttl, Bob/White, Carmela/Wong Gonzalez, Daniela (2017): Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. In: Studies in Educational Evaluation 54, 22–42. DOI: 10.1016/j.stueduc.2016.08.007.

Wolbring, Tobias (2013): Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität. Frankfurt a. M.: Campus.

https://doi.org/10.5771/9783748943334-335 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY