

Was ist gute Lehre?

Gedanken zur ›Evaluation‹ von Lehrveranstaltungen an einer Universität

Da kommt mal wieder eine so genannte Lehrrevaluation auf uns zu. Muss das sein? Es sieht so aus: Unsere Universität hat eine [Evaluationsordnung für die Fachbereiche \(EvalO\)](#). Vermutlich gilt sie auch für deren Rechtsnachfolger: die Fakultäten. Die Ordnung umfasst 15 Paragraphen, verfasst in schönstem Behördendeutsch. Für uns Lehrende ist die „studentische Veranstaltungsbeurteilung“ zuständig (§§ 11–14).

Da heißt es unter anderem:

(2) Die studentische Lehrveranstaltungsbeurteilung dient der Qualitätssicherung und -verbesserung der Lehre im Fachbereich. Sie soll den Studierenden ermöglichen, systematisch und anonym die aus ihrer Sicht vorhandenen Stärken und Schwächen einzelner Lehrveranstaltungen zu benennen und so Verbesserungen in der Lehre zu initiieren bzw. zur Verstetigung guter Lehrveranstaltungskonzepte beizutragen. Den Lehrenden sollen damit zugleich empirisch fundierte Kenntnisse über Stärken und Verbesserungspotentiale ihrer Lehre aus Sicht der Studierenden zur Verfügung gestellt werden.

Wie haben die sich damals wohl vorgestellt? Hat man davon gehört oder gelesen, dass diese „empirisch fundierte Kenntnisse“ irgendwo zu einer Verbesserung geführt hätten. Ich kann mir auch schwer vorstellen, wie das gehen soll: Da kriegen wir zwei oder drei Kurven an die Hand, eine davon sind wir selbst. Wir freuen uns, wenn wir in der richtigen Richtung vom Rest der Welt abweichen – ganz gleich, ob die Abweichung statistisch signifikant ist. Wir freuen uns jedenfalls, wenn wir im grünen Bereich der Skala sind – ganz gleich, ob das nicht vielleicht für Erziehungswissenschaftler an allen deutschen Universitäten der Fall ist. Und wenn wir uns auf der falschen Seite finden, dann haben wir schnell viele gute Erklärungen dafür zur Hand.

Rückblick und Beobachtungen

Die Evaluation von Lehrveranstaltungen an einer Universität ist nicht so neu wie ihr Name: Ich erinnere mich sehr gut daran, wie am Schwarzen Brett des Bonner Instituts für Erziehungswissenschaft – es war wohl im WS 1968/69 – ein Paar DIN A4 Blätter hingen, auf denen die Lehrveranstaltungen unserer Professoren kritisiert wurden – ihr Inhalt wohl gemerkt, die Form spielte keine Rolle. Wir jungen Assistenten freuten uns darüber, zugegeben, denn so ähnlich wie die Studenten-Autoren sahen wir Vieles auch. Und ich erinnere mich recht gut daran, dass mein Assistent, als ich dann selber einer von solchen Professoren war, zu meiner ersten Vorlesung im Sommer 1973 einen der damals kursierenden Fragebogen zur Veranstaltungskritik mitbrachte. Den haben wir adaptiert; und ich habe ihn seither über viele Jahre hin immer wieder einmal benutzt: in Hannover, in Siegen; bei angehenden Lehrern und bei Hauptfachstudenten; zu Beginn von deren Studium und in der Mitte.

Später kam die Sache in Schwung: Es gab standardisierte Fragebögen, HILVE habe ich einmal ausprobiert. Es gab sogar einen Fragebogen unseres Wissenschaftsministeriums – nebenbei ein ausgemachter Schwachsinn, der auch gleich danach in der Versenkung verschwand; ausprobiert habe ich selbst den. Ein ganzer Ordner voll mit Fragebögen, Berechnungen, Ergebnis-Folien dokumentiert meine und in gewisser Weise die Geschichte der Evaluation von Lehrveranstaltungen.

Was dabei herausgekommen ist? Nichts – nichts was ich nicht schon in Hannover gelernt und ohnehin gewusst hätte. Wichtiger aber: Ich habe mich natürlich bemüht, an meinen, mir durchaus bekannten Schwächen zu arbeiten. Das war und ist der Sinn der des inzwischen fest etablierten und gut ausgebauten Systems der Evaluation. Nur: Ausweislich meiner Daten hat sich über die ganze Zeit hin bei mir nichts Wesentliches verändert, allen Bemühungen zum Trotz. Sollte das nur an mir liegen?

Nebenbei: Einmal konnte ich einen Kollegen dafür gewinnen, dass wir beide den meinen alten Fragebogen einsetzten. Es hat uns erstaunt, wie gut die Ergebnisse die Unterschiede der beiden Personen abbildeten – die uns allerdings bereits zuvor durchaus bekannt waren.

Welche Folgerungen habe ich aus meinen Erfahrungen gezogen?

- Aktionismus, gemäß dem Evaluation *als solche* bereits als Beitrag zur Verbesserung proklamiert wird, ist wertlose Augenwischerei oder, in der Politik und auch innerhalb unserer Universität: Legitimation von hochschuldidaktischem Nichtstun.
- Die individuelle Rückmeldung der Wahrnehmungen der Teilnehmer einer Lehrveranstaltung zu *dieser* Lehrveranstaltung kann Leerstellen und Fehler bewusst machen.
- Der Glaube, die individuelle Rückmeldung führe *ohne Weiteres* zu einer Verbesserung von Lehrveranstaltungen, ist Selbstbetrug. Wird er Lehrenden versprochen, damit sie sich auf Befragungen einlassen, ist das vorsätzliche Täuschung.
- Eine individuelle Rückmeldung bedarf vermutlich der Ergänzung durch eine mehr oder weniger professionelle *Supervision*, etwa durch Kommunikation im Kollegium oder durch professionelles Coaching.
- Alle mir bekannten Methoden der Evaluation gehen auf die *Form* der Veranstaltung, allenfalls auf die Strukturiertheit der Inhalte; eine Evaluation des Inhalts ist Jedem selbst überlassen.

Zur Supervision: Der besagte Assistent hat mich in meine Vorlesung begleitet. Danach haben wir Stunden verbracht, in denen wir über die Veranstaltung diskutiert haben. Ich höre, dass heute die Chefs einer Arbeitsgruppe mit ihren Mitarbeitern absprechen, was die in ihren Veranstaltungen lehren (sollen), vielleicht auch darüber sprechen wie man das am besten macht. Von wechselseitigen Besuchen mit nachgängigen Beratungen habe ich nichts gehört – was natürlich nicht heißen muss, dass es das nicht gibt. Aber das in der Ausbildung von Lehrern übliche und in der Evaluation des Unterrichts dieser Novizen sehr erfolgreiche Institut der *Mentoren* hat keine Entsprechung in der Universität.

Eine professionelle Supervision kann man vergessen. Abgesehen davon, dass sie zu teuer ist: Wo soll man die ganzen Profis herkriegern? Ein KOSI soll dem Mangel abhelfen; da ist ganz am Ende der hochschuldidaktischen Exerzitien so etwas vorgesehen. Es mag ja sein, dass das was bringt. Aber KOSI selbst ist eher Dokument eines Mangels, als dass es ihn behöbe. Und vor allem ist das Programm für Adepten gedacht, nicht aber für die Etablierten, die keineswegs schon deswegen besser unterrichten können, weil sie etabliert sind.

Es ist geradezu ein Widerspruch in sich selbst, dass Erziehungs-, Bildungs- oder ähnliche Wissenschaftler zwar darüber belehren und schreiben, was gute Schulen und gute Lehrer sind, aber ihren eigenen (Hochschul-)Lehrer-Nachwuchs einfach so naturwüchsig heranwachsen lassen. Es muss ja nicht so formalisiert sein, wie in meiner Jugend in der gerne geschmähten Ordinariuniversität: teilnehmen an, beobachten von, Hilfsarbeiten in den Lehrveranstaltungen der Chefs und dann ganz langsam die Übernahme von eigenen, zunächst noch unter ihrer Verantwortung. Aber gar nichts von alledem ist ziemlich suboptimal.

Zu den Inhalten: Da passiert nichts. Jede, Jeder kann, ggf. nach ein paar klärenden Worten mit dem fürs Fach Verantwortlichen, diese ohnehin – Jeder kann machen, was er oder sie für richtig halten. Mit den Kollegen ins Gespräch kommen, eine Diskussion über das, was wir die Studenten lehren, ist nicht institutionalisiert. Es gibt keine verbindliche Auseinandersetzung, bei der man wechselseitig zur Kenntnis nimmt, bedenkt, auch mit Außenkriterien abgleicht – und zu Übereinkünften kommt derart, dass in Veranstaltungen mit demselben Titel auch einigermaßen Übereinstimmendes gelehrt wird.

Das sei auch gar nicht möglich, wird eingewendet. Die Pädagogik sei, anders als alle anderen Wissenschaften, standpunktbezogen und demnach vielgestaltig. Muss das heißen, dass man selbst an *einer* Universität darauf verzichtet, sich in Richtung auf einen klar umschriebenen Bestand an Wissen zu verständigen, den man in der Lehre vermittelt?

Die Studenten. In allen Äußerungen zur Evaluation erscheinen die Studenten als Kunden, die die Qualität eines Produkts beurteilen. Bestenfalls werden sie als Experten angesehen. Wofür? fürs Machen von Veranstaltungen? nein, sie sind und bleiben Kunden, Abnehmer eines vorgefertigten Produkts. Mich wundert, dass sie sich das so fraglos gefallen lassen, wie das offensichtlich der Fall ist, und nur darauf achten, dass man „Studierende“ sagt. Dass die Evaluationstheoretiker sich auch nur Gedanken gemacht hätten, welches die *Rolle von Studenten im Ganzen des akademischen*

Unterrichts ist, davon habe ich nichts gehört. Moniert bei uns etwa Jemand, dass Eva Lo die Studenten, die Studierenden gar, als Kunden geradezu *definiert*?

Genauer: Zu den Formen der Lehre, zur Arbeit als Arbeit können Studenten als Experten durchaus etwas sagen, sind sie doch wesentlich an dieser Arbeit beteiligt. Vermutlich deswegen rufen die Evaluationsinstrumente auch diese Aspekte ab. Aber ob die Ergebnisse der Arbeit – sachlich! – etwas taugen? Wenn Studenten das beurteilen könnten, brauchten sie kein Studium mehr. Das gilt erst recht für die Qualität der Gegenstände, die ihnen ihre Lehrer zur Bearbeitung vorgeben. Die können nur Leute beurteilen, die von der Sache was verstehen: Kollegen.

Damit bin ich – *wegen der Studenten*, um derentwillen der ganze Aufwand getrieben wird – wieder beim Kollegium.

Fazit zum gegebenen Anlass

Ich frage die Studenten in meiner Veranstaltung lieber selbst, und ich frage nach Sachen, die mir tatsächlich weiter helfen, zum Beispiel zur Arbeitsbelastung. Das Departement lässt einschätzen:

Den Zeitaufwand bezüglich der Vor- und Nachbereitung in dieser Veranstaltung empfinde ich insgesamt als angemessen.

– als gäbe es keine verbindlichen Vorgaben zu *work load* und *credit points*. Ich frage:

Die Aufgaben waren im Mittel in 2,5 Stunden – gut zu erledigen; kaum zu schaffen; nicht zu bewältigen (n.b. + 1,5 Stunden die Vorlesung selbst; ergibt bei 15 Vorlesungen ein workload von 60 Stunden; und das entspricht 2 CP).

Mit dem, was dabei rauskommt, kann ich was Handgreifliches anfangen: reduzieren oder beibehalten etwa. Oder: *An der Vorlesung finde ich gut ...; Ich wünschte mir...* – Das geht auf den Inhalt, und auch damit kann ich etwas anfangen.

Einmal die Einwände gegen das Überstrapazieren der Statistik beiseite gelassen: Das Verfahren des Departements lässt Vergleiche einzelner Lehrveranstaltungen mit allen übrigen zu; und die Fragebögen werden wohl auch so ausgewertet: Ich gegen den Rest des Departments. Aber was soll schon der Vergleich einer Vorlesung im Audi Max mit den kleinen Seminaren des Hauptstudiums oder den Übungen zur Vorbereitung aufs Praktikum? wenig sinnvoll. Dabei gäbe es Sinnvolleres. Da haben wir zum Beispiel zwei zu meiner parallele *Einführungen in die Erziehungswissenschaft*. Mich würde sehr interessieren, ob und, wenn ja, wie die drei sich unterscheiden. Wenn überhaupt etwas, dann könnte ich aus einem solchen Vergleich etwas lernen, der Vergleichbares vergleicht. Also werde ich die *empirisch fundierte(n) Kenntnisse über Stärken und Verbesserungspotentiale* meiner Vorlesung gerne Jeder oder Jedem zur Verfügung stellen, der Interesse daran hat; übrigens auch die meiner Privatevaluation.

Nebenbei: Die Teilnehmer meiner Vorlesung konnten die letzteren Ergebnisse ein paar Tage nach Schluss der Befragung in Moodle einsehen. [Eva Sys](#) kann schneller rechnen als ich. Ob der Fragebogen des Departements wohl so schnell rüberkommt, dass ich den Studenten noch rechtzeitig zeigen kann, was sie von unserer Veranstaltung halten?

[Peter Menck](#)