

Margrit Stamm

Lehrerschaft zum Rapport!

erschienen in: Neue Zürcher Zeitung, 23. März 2000, 85.

Die Beurteilung des Unterrichts der Lehrkräfte durch ihre Schülerinnen und Schüler oder Studierenden, noch vor kurzer Zeit von nicht wenigen Betroffenen als ‚standeswidrig‘ bezeichnet, ist heute zu einem virulenten bildungspolitischen Thema geworden. Entsprechende Gesetze und Projekte haben in letzter Zeit so viel sozialen Druck entstehen lassen, dass die Bildungspraxis nicht darum herum kommt, sich dieser Fragestellung anzunehmen. Die Akzeptanz der Lehrkräfte ist dabei insbesondere dann vorhanden, wenn die Unterrichtsbeurteilung positiv ausfällt, in klasseninterner und somit privater Initiative durchgeführt werden kann und mögliche Konsequenzen ebenfalls in der Eigenverantwortung bleiben. Widerstände wachsen jedoch, wenn es um systematische, schulumfangende und von der Schulleitung initiierte Beurteilungen geht, die auch einen gewissen Anspruch auf empirische Wahrheit haben sollen. Im Fadenkreuz der Diskussionen und Kontroversen stehen in diesem Fall berechtigterweise die *Gültigkeit* von Unterrichtsbeurteilungen, also die Frage, ob Schüler oder Studierende Unterricht oder Lehre überhaupt adäquat beurteilen können und wozu denn letztlich eine solche Evaluation dienen soll.

Damit sind wir bereits mitten in der Problematik. Überblickt man nämlich die aktuelle Situation an den schweizerischen Bildungsinstitutionen, kann man un schwer feststellen, dass Unterrichtsbeurteilungen als Evaluationsinstrument zwar relativ häufig, meist jedoch unsystematisch eingesetzt werden, auf die *Durchführung* beschränkt bleiben und deshalb in vollkommener *Konsequenzlosigkeit* enden. So bleibt auch die vielleicht wesentlichste Frage unüberprüft und deshalb unbeantwortet, nämlich ob sich die Qualität von Lehre und Unterricht durch Unterrichtsbeurteilung tatsächlich verbessert und mit welchen Mitteln dies erreicht werden könnte. Gesamthaft sieht die Bilanz also düster aus: Aufwand enorm - Ertrag praktisch nichtssagend - Wirkungslosigkeit gross.

Die berechtigten Zweifel der Lehrkräfte

Ein Grund dafür mag darin liegen, dass die Anforderungen an die Qualität des Instrumentariums nicht gross sein müssen, solange es nur darum geht, den Dialog zwischen Lehrperson und Schülerschaft über den Unterricht herzustellen. Soll die Unterrichtsbeurteilung jedoch weiterführenden Zwecken dienen und beispielsweise auch zur Personalentwicklung eingesetzt werden, steigen die Anforderungen an Messgenauigkeit, Zuverlässigkeit und Objektivität des Instruments, aber auch die Widerstände und Ablehnungen der Lehrerschaft. Kritische Stimmen bezeichnen Unterrichtsbeurteilungen etwa als Popularitätstest von Lehrkräften, der von der Benotungsstrenge in Prüfungen, von der Themenschwierigkeit, der Hausaufgabenlast und nicht zuletzt von der fachlichen Begabung der Schülerinnen und Schüler, deren Geschlecht und (Studien-)Alter abhängt. Attraktive Dozierende

und insbesondere Frauen würden zudem besser beurteilt. Derartige Vorurteile sind ernst zu nehmen und auf der Grundlage empirischer Untersuchungen zu beantworten.

Das Problem der inkompetenten Lehrkräfte

Eine mit praktischen Konsequenzen verbundene Durchführung von Unterrichtsbeurteilungen bedingt, dass das Thema ‚ungenügende Leistungen von Lehrkräften‘ aus der Tabuzone herausgeholt werden und der darüber liegende Dunstschleier gelüftet werden muss. Obwohl eigentlich jede Bildungsinstitution mit diesem Problem konfrontiert ist, war bis heute kaum jemand genötigt, sich mit seinen typischen Komponenten auseinanderzusetzen. Systematische Unterrichtsevaluation bringt nun jedoch die unangenehme Aufgabe mit sich, dass sich Schulverantwortliche diese Problematik nicht weiter ignorieren können und sich bereits in der Planungsphase mit daraus resultierenden Konsequenzen und möglichen Entwicklungsschritten befassen müssen. An Fachhochschulen kommt der Druck der Studierenden dazu, die aus direkter ‚Kundenperspektive‘ und mit Verweis auf die Wettbewerbssituation zwischen den Fachhochschulen argumentieren und unmittelbare Auswirkungen ihrer Beurteilungen geradezu fordern.

Antworten aus der Forschung

Nun wäre es ein Einfaches, solche Probleme als Empfehlung zu Absage an die Lehr- und Unterrichtsbeurteilung durch Schüler oder Studierende zu interpretieren und – mit Verweis auf die bereits in den siebziger Jahren gemachten Erfahrungen – ganz auf das Urteil der Lernenden zu verzichten. Im Rahmen eines umfassenden Qualitätsmanagements ist dies jedoch aus zwei Gründen nicht möglich: Erstens fordern Gesellschaft und Politik, dass Leistungen von Bildungsinstitutionen (z.B. von Fachhochschulen in den Bereichen Ausbildung, Weiterbildung, Forschung und Dienstleistung) zugänglich gemacht und ihre Leistungsbereitschaft der Öffentlichkeit und den potenziellen Studierenden gegenüber kommuniziert werden soll. Zweitens *muss* eine wesentliche Unterrichtskritik von studentischer oder schülerbezogener Seite herkommen, da Lehrkräfte den Unterricht von Kollegen kaum beurteilen können, weil sie an deren Veranstaltungen nicht oder sehr selten teilnehmen. Weil den Schülern und Studierenden so wesentliche Bedeutung zukommt, sind Forschungsarbeiten wichtig, welche die wissenschaftliche Grundlegung von derartigen Unterrichtsbeurteilungen ermöglichen. Im deutschsprachigen Raum existieren dazu umfangreiche Arbeiten. Sie können Antworten auf Behauptungen geben, Schüler- oder Studenturteile seien nicht gültig (valide) und deshalb nicht brauchbar. Derart befürchtete Verzerrungseinflüsse scheinen zwar nach eigenen Untersuchungen und in Übereinstimmung mit einem Grossteil internationaler Studien unbegründet zu sein. So üben beispielsweise weder Alter, Geschlecht, Semesterzahl oder Vorbildung einen wesentlichen Einfluss auf die Beurteilung aus. Schüler- und Studenturteile sind in der Regel ausserordentlich differenziert und objektiv und die Beurteilungstendenz ist eindeutig positiv. Die Erfahrung zeigt, dass die Urteile in der Regel zu 80 % positiv und zu je etwa 10 % eindeutig negativ bzw. ausgezeichnet ausfallen. Wichtiger scheint indes der Umstand – und damit die bereits angesprochenen Befürchtungen – zu sein, dass die *Zuverlässigkeit* der Beurteilungen negativer eingeschätzt werden muss. In allen eigenen Untersuchungen zeigte sich, dass erstens die Veranstaltungen von verschiedenen Schüler- oder Studierendengruppen unterschiedlich beurteilt werden und somit nicht etwa das Fach, sondern die *Klasse* die stärkste Einflussgrösse dar-

stellte. Zweitens wurden *Dozentinnen*, welche an technisch orientierten Fachhochschulen unterrichten, signifikant schlechter beurteilt als ihre männlichen Kollegen.

Aber auch die Qualität des Instruments an sich, das ja meist als Fragebogen eingesetzt wird, hat einen Einfluss auf die Gültigkeit und Zuverlässigkeit der Ergebnisse. Glaubwürdige Resultate können nur mit einem Instrument erhoben werden, das *unterschiedliche Dimensionen* beinhaltet, also etwa zwischen Rahmenbedingungen, didaktischen Qualitäten der Lehrperson, Lehr- und Lernklima und Ergebnissen unterscheidet und auch die Bewertung des eigenen Verhaltens durch die Lernenden (Lerngewinn, Kompetenzerwerb) einschliesst. Falsch ist dabei indes, wenn pro Beurteilung lediglich ein Mittelwert berechnet, dieser als Qualitätsindikator interpretiert und darauf basierend sogar ein Ranking veranstaltet durchgeführt wird. Richtig ist hingegen, wenn lediglich Aspekte gleicher Dimensionen zu einem Mittelwert verrechnet werden. Da Mittelwerte allerdings immer als problematische Grössen anzusehen sind, muss die Heterogenität der Urteile immer mitberücksichtigt werden.

Völlig fehlerfreie und durch keine Störfaktoren beeinflusste Unterrichtsevaluationen sind nicht möglich, auch nicht auf der Grundlage der ausgereiftesten und theoriebasiertesten Modelle. Gerade deshalb darf Unterrichtqualität nie ausschliesslich über Studenten- oder Schülerurteile definiert werden. Vielmehr gilt es, sie um Selbsteinschätzungen der Lehrkräfte und *fachliche* Qualifizierungen durch Experten (nicht durch Laien) zu ergänzen. Ein Qualitätsmanagement, das seinen Namen verdient, deklariert den Unterricht zwar als Herzstück der Qualitätsbemühungen, richtet sein Augenmerk jedoch auch auf die Qualität der Bildungsinstitution als Ganzes inklusive ihrer curricularen Prozesse sowie auf die Nachhaltigkeit der Ausbildungsleistungen, die über Absolventen- und Abnehmerevaluationen erhoben werden können.

Fünf Schritte zur Wirksamkeit von Unterrichtsbeurteilung

Damit Unterrichtsbeurteilung durch Studierende oder Schüler nicht zu einem Etikettenschwindel und damit zu einer netten Aufmachung medienbewusster Schulen verkommt, sind verschiedene, die Wirksamkeit fördernde Schritte zu berücksichtigen: *Erstens* sind unmittelbar an die Beurteilung anknüpfende *Rückmeldegespräche* zwischen Lehrperson und Schüler bzw. Klasse als verbindlich zu erklären. Dozierende werden so angehalten, ihre Resultate im Sinne eines Feedbacks zu besprechen. Dies führt dazu, dass Verbesserungseffekte wahrscheinlicher werden, weil der Verpflichtungscharakter grösser wird. Dazu kommen *zweitens* darauf aufbauende *Qualifizierungsgespräche* mit der Schulleitung, welche Zielvereinbarungen und Konsequenzdiskussionen beinhalten. Diese wiederum sollen auf *Gratifizierungssysteme* für Unterrichtsoptimierungen zurückgreifen können und als Motivierungsmassnahmen eingesetzt werden. Kritisch beurteilte Dozierende sind darin einzubeziehen. *Drittens* ist ein Parallelangebot an Beratung und Unterrichtsbegleitung mit ausschliesslicher Unterstützungsorientierung zu etablieren, das *viertens* mit gezielten, bei ungenügenden oder kritisch beurteilten Unterrichtsleistungen als verpflichtend bezeichneten *Weiterbildungsangeboten* verbunden wird. *Fünftens* sind die Hauptakteure, die Schüler oder Studierenden, in die Überlegungen einzuschliessen: Sie sind gezielt in ihre verantwortungsvolle Aufgabe einzuführen und auf gerechte, objektive Urteile zu verpflichten. Um einer gewissen Evaluationsmüdigkeit – die sich beim kontinuierlichen Ausfüllen von Fragebögen

zwangsläufig ergibt – vorzubeugen, ist die Evaluation während des regulären Unterrichts und in einem Rhythmus von drei bis vier Semestern durchzuführen.

Last but not least setzen Bemühungen um einen qualitativ hochstehenden Unterricht jedoch an einem ganz anderen Ort ein, nämlich bei der Berücksichtigung der *pädagogischen Eignung* von Dozentinnen und Dozenten bei Einstellungs- und Berufungsverfahren. Indikatoren zur Unterrichtsqualität sollten deshalb auch die Frage umfassen, welche Bemühungen eine Schule tatsächlich unternimmt, um *methodisch-didaktisch* ausgewiesene Lehrkräfte zu gewinnen und sich nicht lediglich auf Forschungs- und Publikations- bzw. andere Reputationsleistungen auszurichten.

Chancen und Gefahren

Insgesamt birgt studentische und schülerbezogene Unterrichtsevaluation Chancen und Grenzen. Die *Chancen* liegen in erster Linie darin, dass die Diskussion über Lehr- und Unterrichtsqualität eröffnet und die Arbeit von Lehrerinnen und Lehrern an bestimmten Standards gemessen wird. Durch die stärkere Gewichtung des eigenen Unterrichts kann das Qualitätsbewusstsein gefördert und die Sensibilität der Lehrkräfte gegenüber methodischen Stärken und Mängeln erhöht werden. Werden zudem die wissenschaftlichen Forschungsergebnisse akzeptiert, liegen die Chancen darin, dass Dozentinnen und Dozenten lernen, mit positiven und negativen Resultaten zu leben, ohne diese einfach mit Argumenten wie mangelnde methodische Qualität des Instruments oder unzureichender Beurteilungsfähigkeit der Schüler oder Studierenden abzutun. Die *Gefahren* finden sich indes dort, wo Unterrichtsbeurteilung allein eine von oben verordnete Massnahme bleibt, ad hoc eingeführt, nicht bezüglich der Konsequenzen kommuniziert und zu den Qualitätszielen der Schule in Bezug gesetzt wird. So gehandhabte Unterrichtsbeurteilung bewirkt dann entweder gar nichts und bleibt Etikettenschwindel oder sie führt die ‚Lehrerschaft zu einem erniedrigenden Rapport‘, was einer reinen Dozentenbewertung gleichkommt, die als polemische Angelegenheit zur reinen Manipuliermasse verkommen kann.