

Verbesserung der Lehre durch Selbst-bestimmte Evaluation  
(Improvement of Teaching through Self-directed Evaluation; ITSE)

Bericht über einen Modellversuch an der Universität Konstanz<sup>1</sup>

Georg Lind<sup>2</sup>

---

<sup>1</sup> Gefördert durch Mittel des *Ausschuss für Lehre und Weiterbildung* (ALW) der Universität Konstanz.

<sup>2</sup> Anschrift des Autors: Prof. Dr. Georg Lind, Dipl.-Psych., Naturwissenschaftliche Sektion,  
Fachbereich Psychologie, Universität Konstanz, 78457 Konstanz. Email: Georg.[Lind@uni-konstanz.de](mailto:Georg.Lind@uni-konstanz.de) .  
Projekt-Webseite: [www.uni-konstanz.de/itse/](http://www.uni-konstanz.de/itse/) .

## Theoretischer Hintergrund des Projekts

Das Konzept der Selbst-Evaluation geht von einem dynamischen Qualitätsentwicklungsmodell aus, das die Analyse und Bewertung von Lehrmethoden in den Mittelpunkt stellt. Er geht von der wissenschaftlich abgesicherten Erkenntnis aus, dass Lehrfähigkeit erlernt werden kann und eine statische Einteilung von Lehrenden in gute und schlechte sich nicht empirisch begründen lässt, sowie von vielfachen Befunden, dass eine statische Qualifizierung von Lehrenden oder Institutionen in der Regel mit einer Absenkung des Leistungs- und Lernniveaus im ganzen Spektrum einhergeht. Die als "schlecht" etikettierten Lehrenden verlieren dadurch oft die Motivation, sich zu verbessern; vor allem zeigt ihnen diese Art der "Evaluation" keinen Weg zur Verbesserung auf. Aber auch für die als "gut" ausgezeichneten Lehrenden stellen solche "Auszeichnungen" selten einen Anreiz dar, ihre Lehrmethoden zu verbessern, da sie damit riskieren, "Fehler" zu machen und in die Kategorie "schlecht" zu rutschen (Kohn, 1999).

Evaluation kann unter bestimmten Bedingungen die Qualität der Lehre und die Motivation der Lehrenden und Lernenden entscheidend verbessern helfen, nämlich a) wenn für den Lehrenden und die Hochschule Nutzen und Ertrag in einem günstigen Verhältnis stehen, b) die Evaluation in eine Qualitätsentwicklung eingebettet ist und ihre Methoden selbst bestimmten Mindeststandards genügen und c) wenn sie auf wissenschaftlichen Erkenntnissen über Lehr-Lernprozesse basiert. Der Bericht über den ersten Projektabschnitt wird dies ausführlicher begründen und eine Reihe von stützenden Analyseergebnissen vorlegen.

### *Selbst-Evaluation der Lehreffektivität: Ein paar Definitionen*

*Evaluation* bedeutet hier die Bewertung von empirischen Wirksamkeitsstudien anhand von klar definierten Kriterien im Rahmen übergreifender einschlägiger Zielsetzungen und wissenschaftlicher Erkenntnisse. Gewöhnlich wird hierzu auch noch – zu Recht – die Bewertung der *Effizienz*, also der *Relation von Aufwand und Ertrag* gezählt, was hier aber – aus Kapazitätsgründen – ausgeklammert bleiben muss.

*Selbst-Evaluation* bedeutet hier, dass ein Hochschullehrer<sup>3</sup> seine Lehre *selbst* evaluiert, das heißt die Maßnahmen und Methoden wählt, die er evaluieren will, die Ziele und Methoden dieser Evaluation bestimmt, über die resultierenden Daten verfügen kann und die Interpretation der Daten selbst vornimmt, ohne dass die Pflicht besteht, die Ergebnisse seinem Arbeitgeber oder der Öffentlichkeit zu offenbaren (um sich dafür Lob oder Bestrafung einzuhandeln). Selbst-Evaluation bedeutet hier nicht die *subjektive Selbsteinschätzung* von Hochschul-

---

<sup>3</sup> Wenn hier die generische Form 'Hochschullehrer' verwendet wird, die sich mit der männlichen deckt, so sind damit natürlich immer auch weibliche Hochschullehrerinnen gemeint.

lehrern bezüglich der Qualität oder Wirksamkeit ihrer Lehre. Subjektive Selbsteinschätzungen sind zwar einfach durchzuführen und billig (weswegen sie sich auch einer großen Beliebtheit erfreuen) und können ohne Zweifel auch ein wichtiges Element von Evaluation darstellen. Sie können aber aus offenkundigen Gründen keine wissenschaftlich fundierte Selbst-Evaluation ersetzen. Selbst-Evaluation und Selbsteinschätzung haben methodisch kaum Gemeinsamkeiten, dienen meist sehr verschiedenen Zielsetzungen und genügen unterschiedlich stark ausgeprägten Ansprüchen an die wissenschaftliche Qualität. Man sollte daher vermeiden, Selbsteinschätzungen von Lehrenden als ‘Selbstevaluation’ oder ‘Eigenevaluation’ zu bezeichnen, wie das offenbar häufig der Fall ist.<sup>4</sup>

Immer dann, wenn eine Evaluation von außen angeordnet, in der Durchführung vorgeschrieben, mit der Auflage der Veröffentlichung und mit Sanktionen verbunden ist, spricht man besser von *Fremd-Evaluation*. (Man kann wohl davon ausgehen, dass bereits die Offenlegung eine Art Sanktion darstellt oder mit sich zieht, indem sie meist unweigerlich Lob oder Tadel bzw. Belohnungen oder Bestrafungen mit sich zieht, wie sie in anderen Ländern bereits praktiziert werden und auch bei uns angekündigt sind.<sup>5</sup> Im Rahmen einer Fremd-Evaluation werden zwar ‘Eigenevaluationen’ (d.h., Selbsteinschätzungen) der Lehrenden und der Hochschulen angesprochen. Man darf aber davon ausgehen, dass sich im Kontext einer Fremd-Evaluation, bei der immer auch öffentliche Herabwürdigung oder gar materielle Nachteile drohen, a) niemand selbst der Ineffektivität bezichtigt wird und b) Außenstehende jedwede Selbsteinschätzung als irrelevant abtun werden.

Natürlich kann Fremd-Evaluation eine wichtige Rolle in der Qualitätsentwicklung des Bildungssystems und der Lehre spielen und ist dafür eigentlich auch unverzichtbar. Um diesen Zweck zu erfüllen, muss auch sie aber wissenschaftlich anerkannten Standards genügen (Lind, 2004), die zum Beispiel Sanktionen (auch unbeabsichtigte Sanktionen, wie jede, die mit einer bestimmten Art der Veröffentlichung einhergehen) ausschließen, da solche Sanktionen in der Regel zu einer zu großen “Unschärfe” der Evaluation führen mit weitreichenden negativen Folgen für das Bildungssystem.<sup>6</sup> Sanktionen jeglicher Art führen offenbar immer dazu, dass bei den Betroffenen (wozu nicht nur Hochschullehrer zählen, sondern auch Rektoren, Minister und private Hochschulentwicklungs-Firmen) mehr und mehr die materiellen Folgen für die eigene Person als die Qualität der Hochschule und der Lehre in den Vordergrund rücken.

---

<sup>4</sup> Siehe u.a. das Universitätsgesetz des Landes Baden-Württemberg, § 4a, Abs. 2: “Die Arbeit der Hochschulen in Forschung und Lehre .... soll durch Eigen- und Fremdevaluation regelmäßig bewertet werden... die Ergebnisse der Bewertungen sollen veröffentlicht werden.” Auch: Expertenkommission (2000).

<sup>5</sup> Kohn (1999); Witte & Schreiter (2002).

<sup>6</sup> Amrein & Berliner (2002); Kohn (2000).

In dem vorliegenden Projekt wird durchweg von Evaluation in dem oben dargelegten Sinn gesprochen, also nicht in den diversen Bedeutungen, die der Evaluation in der Öffentlichkeit gegeben werden. Wir beschränken den Gegenstand hier auch auf Selbst-Evaluation. Gleichwohl sind viele der Methoden und Techniken, die hier zur Sprache kommen, auch für die Fremd-Evaluation verwendbar, wenn diese sich auf die Verbesserung der Hochschule und der Lehre konzentriert und nicht anderen, extrinischen Zielen (mit-)dient, wie der Belohnung oder Bestrafung von Personen und Institutionen, wie das offenbar bei den meisten in der Öffentlichkeit diskutierten Evaluationsprojekten der Fall ist.<sup>7</sup>

Wissenschaftliche Evaluation, also Evaluation, die nach wissenschaftlich anerkannten Standards arbeitet, muss sich selbst der Analyse und Bewertung (also der Evaluation) unterwerfen, um die eigenen Standards weiter zu entwickeln, um gegebenenfalls Fehlentwicklungen erkennen zu können, die die Evaluation selbst zu einer Schadensquelle werden lassen. Schlechte Evaluation kann mehr Schaden anrichten, als ihre Unterlassung. Der Schaden einer schlechten Evaluation kann katastrophale Ausmaße annehmen, wenn keine Vorkehrungen getroffen werden, um Fehlentwicklungen rechtzeitig zu erkennen.

### *Warum Evaluation?*

Wissenschaftliche Hochschulevaluation, also Bewertung von Lehr-Maßnahmen und -methoden ist eine unverzichtbare Voraussetzung für deren Verbesserung. Nur so kann man erfahren, ob eine Maßnahme oder Methode in der Praxis auch wirklich die Effekte haben, die man sich von ihnen aufgrund von theoretischen Überlegungen und empirischer Forschung unter Laborbedingungen erhofft.

Angefangen von einfachen didaktischen Techniken, über die Seminargestaltung, Lehrmaterialauswahl, Studienordnung, Mittelverteilung auf Hochschulebene und gesellschaftsweit bis hin zu Prüfungsordnungen und Entscheidungen der Gesetzgeber über die Ausbildung, Lehrbelastung und Besoldung von Hochschullehrern, all das ist meist mit der *Absicht* verbunden, die Lehre zu verbessern, damit Studierende möglichst gründlich und umfassend auf die Bewältigung der vielfältigen Aufgaben vorbereitet werden, mit denen Menschen generell und Hochabsolventen im Besonderen konfrontiert sind bzw. Zukünftig konfrontiert sein werden.

Evaluation ist überall ein Thema, wo soziale Ressourcen (Geld, Personal, Gebäude etc.) ef-

---

<sup>7</sup> Zu dieser Schlussfolgerung gelangt man, wenn man solche Projekte auf ihre Ziele hin befragt. Entweder es werden nur Sanktionen (in Form von Mittelzuweisungen oder -kürzungen) diskutiert oder sie werden meist gar nicht genannt, was wohl kaum der Fall wäre, wenn sie sich klar und ausschließlich auf die Verbesserung der Hochschule beziehen würden.

fektiv und effizient eingesetzt werden sollen: in Medizin, Wirtschaft, öffentlichen Verwaltung etc. Die Evaluation verschiedener Bereiche des Bildungssystems hat mit "PISA", der weltweiten Evaluation nationaler Bildungsangebote für die unter 15-Jährigen, auch öffentliches Medieninteresse geweckt: Was leisten unsere Schulen und Hochschulen? Haben wir die richtigen Lehrer, die richtigen Strukturen oder die richtigen Lehrpläne und Lehrmethoden?

Ob und inwieweit die Maßnahmen und Methoden, die man beschlossen hat oder die man in für Reformen vorschlägt, auch wirklich erreicht werden, ist immer fraglich. Meist stützen sich solche Beschlüsse und Vorschläge auf theoretischen Schlussfolgerungen, die in der Praxis noch kaum erprobt sind, oder auf kleinen Modellversuchen, die sich bei der Ausdehnung auf größere Einheiten als unwirksam oder sogar als kontraproduktiv erweisen können, oder auf Laborstudien unter kontrollierten Bedingungen, die in einem realen Umfeld unkontrollierbar sind. Oder die Maßnahmen und Methoden werden gewählt, weil sie sich im praktischen Kontext in einem anderen Land oder zu einer anderen Zeit bereits gut bewährt zu haben scheinen.

All dies ist legitim und oft auch der einzig erfolgversprechende Weg, um festgestellte Defizite zu beheben oder gewollte Verbesserungen einzuleiten. Aber die Gefahr ist groß, dass die gewählten Maßnahmen und Methoden keine der erhofften Verbesserungen erbringen oder sie sich am Ende sogar als Fehler herausstellen.

Aber entgegen einer weit verbreiteten Meinung sind nicht alle Methoden der Evaluation gleich gut geeignet, um diese wichtige Funktion zu erfüllen, was von vielen "Evaluatoren" leider übersehen wird. Dabei hat sich die Evaluationsforschung im Laufe der letzten Jahrzehnte zu einem profunden Wissenschaftszweig entwickelt, der klare Standards für eine nutzbringende und faire Evaluation erarbeitet hat.<sup>8</sup>

### **Projektbeschreibung**

Als Evaluation bezeichnet man die gezielte Überprüfung der Wirksamkeit (Effektivität) und Leistungsfähigkeit (Effizienz) bestimmter Maßnahmen und Methoden (zum Beispiel in der Hochschullehre) anhand klarer Zielvorgaben. Evaluation ist dort besonders sinnvoll, wo Qualitätsentwicklung betrieben wird (siehe auch ISO 9000-9004 und das Konzept des "Total Quality Managements" TQM).<sup>1</sup> Eine Evaluation ohne klare Entwicklungsziele kann dagegen eher negative Auswirkungen haben. Ohne Ziele fehlen der Evaluationsforschung die Kriterien für die Anlage und Durchführung einer Evaluation sowie für deren Bewertung. Das gilt für die Wirtschaft ebenso wie für das Bildungssystem oder andere Bereiche (Weinert, 2001;

---

<sup>8</sup> Siehe u.a. Feldman & Newcomb (1970), Dohmen (1971), Popham (1974), Smith (1986), McLaughlin & Phillips (1991), Pascarella & Terenzini (1991), Sanders (1994), Kromrey (1994), Hage, N.e. (1996), Balk (2000), Rindermann (2001).

Deming, 1994). McKeachie (1997) nennt neben unklaren Evaluationskriterien noch sechs weitere Problembereiche von unsachgemäß durchgeführter Evaluation:

1. Studenten wie Professoren lernen leichter und sind motivierter, wenn sie selbst aktiv werden. Eine häufige Evaluation kann ihre Lernmotivation untergraben (S. 1219).
2. Vergleichende Bewertungen wirken sich eher negativ als positiv aus. Die überdurchschnittlich Guten werden kaum zu besserer Leistung angeregt und die Unterdurchschnittlichen oft noch mehr demotiviert. "A solution for [ . . . ] these problems is better feedback [ . . . ] targeted at specific problems." (S. 1223)
3. Wenn der Lernerfolg auf abfragbares Faktenwissen reduziert wird, leiden viele andere Kompetenzen, die für das Leben und den Beruf unerlässlich sind.
4. Ob und wie viel ein Studierender lernt, liegt nicht allein in der Macht des Hochschullehrers, sondern hängt von vielen Faktoren ab, auf die er oder sie keinen Einfluss hat.
5. Sanktionen führen nicht automatisch zu besserer Lehre. Wüssten die Lehrenden, die für schlechte Lernleistungen zu Rechenschaft gezogen werden, wie man effektiv lehrt, wären sie nicht zur Zielscheibe von Sanktionen geworden.

Angesichts dieser Probleme hielt es der Wissenschaftsrat (1996) "grundsätzlich für erforderlich, das Evaluationsverfahren sowie das hierbei angewandte Instrumentarium einer kontinuierlichen, übergreifenden Beobachtung zu unterziehen".

#### *Mindestanforderungen an gute Evaluation*

In der Evaluationswissenschaft haben sich einige Mindestanforderungen für gute Evaluation herausgebildet, die unabdingbar, wenn auch nicht unbedingt ausreichend sind. Das ITSE-Projekt versucht diese Standards umzusetzen.

- a. *Einbettung in die Qualitätsentwicklung der Lehre*: Das ITSE-Projekt stützt sich auf das erklärte Ziel, die Lehre zu verbessern, um bei den Studierenden ein besseres Lernverhalten zu erreichen. Als Aktionsebene werden hier die "Hochschullehrer" angesprochen, die den größten unmittelbaren Einfluss auf das Qualitätsniveau der Lehre haben.
- b. *Zielvalidität*: Die Fragen im ITSE-Fragebogen wurden so ausgewählt, dass sie eine

möglichst hohe Validität der Dateninterpretation erlauben. Es werden möglichst klar formuliert und – bei komplexen Sachverhalten – gut erprobte Frage- und Testinstrumente eingesetzt.

- c. *Lernqualität als Kriterium*: Im ITSE-Projekt ist Lernen bzw. die *Qualität der Lernergebnisse* das zentrale Kriterium der Evaluation. Dies erscheint umso wichtiger als es hierzu nur wenig Evaluationsforschung gibt.
- d. *Input-Output-Kontrolle*: Um möglichst eindeutige Hinweise auf die Gründe für die (mangelnde) Effektivität und Effizienz von Lehrveranstaltungen zu erhalten, wird im ITSE-Programm eine Mehrfacherhebung unterstützt. Da eine Einmalerhebung kaum Hinweise auf Ursachen mangelnder Lehreffektivität gibt, ist als Minimum eine zweimalige Erhebung pro Semester vorgesehen. Es können jedoch mühelos Zwischenerhebungen und Nachfolgerhebungen integriert werden.
- e. *Sanktionsfreiheit*: Evaluation haben offenbar auch oder gerade dann eine fördernde Wirkung auf die Qualität der Lehre, wenn sie frei von Sanktionen und Belohnungen sind (Balk, 2000). Studien haben immer wieder gezeigt, dass Hochschullehrer wie auch Schullehrer an sich selbst hohe Ansprüche stellen, was die Qualität und Effektivität der Lehre angeht und “extrinsischer” Motivation nicht bedürfen. Sanktionen und Belohnungen scheinen sogar den breiten Grundvorrat intrinsischer Motivation systematisch zu verringern und genau das Gegenteil von dem zu bewirken, was beabsichtigt ist, nämlich eine Verschlechterung der Lehrleistung und Lehreffektivität (Deming, 1994; Kohn, 1999; Frey, 2003).<sup>2</sup> Im ITSE-Projekt liegt der “Gewinn” für den Lehrenden in der Erfahrung von didaktischer Wirksamkeit.
- f. *Transparenz und Mitwirkung*: Um die Einhaltung der oben aufgeführten Mindeststandards sowie die Umsetzung der Befunde in praktische Lösungen zu sichern, ist das ITSE-Projekt völlig transparent angelegt.

## 2. *Effektive Lehre als Ziel der Evaluation*

Das Ziel hochschulischer Ausbildung besteht in der Vermittlung umfassender Lebens- und Berufskompetenzen, die den Hochschulabsolventen für die unterschiedliche Bewältigung von Aufgaben und Funktionen befähigen soll (Bargel & Framhein, 1976; Bargel et al., 1982). Lehrziele müssen also viel umfassender als bislang verstanden – und evaluiert – werden, nämlich als die sich gegenseitig bedingende Verschränkung von fünf Lernqualitäten, die eine psychologische Einheit bilden: Aufnehmen, Verstehen, Handeln, Fühlen und Bewerten. Oft steht in der hochschulischen Lehre nur das “Aufnehmen” (und Weitergeben) von Wissen im

Vordergrund. Für eine voll entwickelte akademische Persönlichkeit, die in Wirtschaft und öffentlicher Verwaltung einmal eine verantwortungsvolle Stellung einnehmen soll, bedarf es der Ausbildung vielfacher Kompetenzformen: Verstehen, Handeln, Fühlen und moralische Urteilen.<sup>3</sup>

Solche Lehr- und Lernziele des Hochschulstudiums sind nicht immer einfach zu lehren und auch nicht immer einfach zu evaluieren, weswegen die meisten Evaluationsverfahren sich auf einfach zu erhebende Teilaspekte oder indirekte Indikatoren beschränken. In diesem Projekt sollen direkte Indikatoren für den Lernzuwachs bei den Studierenden ermittelt und ein Test für das komplexe Konstrukt moralische Urteils- und Diskursfähigkeit eingesetzt werden, der in Konstanz im Rahmen eines SFB-Projekts entwickelt (Lind, 2004/im Druck) und in verschiedenen internationalen Projekten zur Evaluation von Bildungsinstitutionen eingesetzt wurde (Lind, 2002).

Bei diesen fünf Lernqualitäten handelt es sich nicht um ein Schichtenmodell, wie das im oft benutzten Begriff der "Basisfähigkeiten" anklingt, sondern um ein Aspektmodell, in dem sich alle Prozesse gegenseitig stützen und bedingen (oder, im negativen Fall, gegenseitig behindern). Aufmerksamkeit ist eine Voraussetzung für alles Lernen; ohne die Aufnahme von Informationen durch die Lernenden kann kein Lernen stattfinden. Aber, wie viele Untersuchungen der Lehr- und Lernforschung zeigen, mangelnde Aufmerksamkeit ist auch von dem Niveau der angesprochenen Lernqualität abhängig. Reine "Wissensvermittlung" ermüdet, weil sie an dem Bedürfnis der Lernenden nach Verstehen, Handeln, Fühlen und Bewerten vorbeigeht. Studierende lernen wesentlich engagierter und effektiver, wenn sie all ihre Lernbedürfnisse respektiert sehen, auch ihre emotionalen, sozialen und moralischen.

Das Projekt will einen Beitrag zur Umsetzung dieses breiten, vernetzten Lernbegriffs in der Lehrevaluation leisten und Instrumente für die Erfassung inhärenter Qualitätsmerkmale von guter Lehre entwickeln, die, wie der Wissenschaftsrat (1996) anmerkt, bislang "nicht verfügbar sind".

### *3. Realisierung des Projekts "Verbesserung der Lehre durch Evaluation"*

Auf der Grundlage des oben dargestellten Ziels, die Qualität bzw. Die Effektivität der Lehre zu fördern, wurde vom Antragsteller das ITSE-Projekt vorgeschlagen, das im Antrags sowie im Bericht zum ersten Projektabschnitt ausführlicher dargestellt wird.

Im Wesentlichen soll mit dem ITSE-Projekt jedem an Lehrverbesserung interessierten Hochschullehrer eine Internet-gestützte Infra-Struktur für die Durchführung und Auswertung eines eigenen Evaluationsvorhabens zur Verfügung gestellt werden. Das Ziel des



vorliegenden Projekts ist, diese Infrastruktur möglichst einfach handbarbar, preiswert und vielseitig einsetzbar zu gestalten, so dass im Prinzip jeder Lehrende und jede Lehrende sie nutzen können. Als positiver Nebeneffekt, so stellt sich heraus, wird dem Nutzer des ITSE-Programms die Teilnehmerverwaltung seiner Kurse erleichtert, indem die Anmeldung und Dokumentation der studentischen Leistungen durch ITSE weitgehend automatisiert wird.

Das ITSE-Projekt richtet sich an Hochschullehrer, die im Projektzeitraum bestimmte Aspekte ihrer Lehre evaluieren und evtl. Neues auszuprobieren planen.

Die Teilnehmer am ITSE-Projekt sollen vor allem vor Beginn relativ intensiv betreut, um sie mit den Möglichkeiten des Programms vertraut zu machen und ihnen zu helfen, Evaluationsideen umzusetzen.

Alle Daten der teilnehmenden Hochschullehrer und Studierenden werden vertraulich behandelt. Die Befragungen der Studierenden zu den Lehrveranstaltungen erfolgen völlig anonym. Personenbezogene Daten (die Listen der Kursteilnehmer und der Projektteilnehmer) werden so gespeichert, dass keinerlei Verbindungen zu Befragungs- und anderen Evaluationsdaten hergestellt werden können.

Am Ende des 2. Projektabschnittes soll das gesamte Programm dem einzelnen Dozenten zur selbständigen Benutzung überlassen werden, so dass ein maximaler Vertrauensschutz gewährleistet ist. Alle Komponenten des ITSE-Programms sind kostenlos und können daher jedem Interessenten kostenlos überlassen werden.

#### *4. Beobachtungsinstrumente*

In der Basiserhebung und der Schlusserhebung des ITSE-Programms werden teilweise verschiedene, spezifische Fragen gestellt, und teilweise identische, um Wirkungseffekte erfassen zu können.:

##### *Basiserhebung:*

- Allgemeine Merkmale der Teilnehmer (Fach- und Studiensemester, abgelegte Prüfungen, abgeschlossene Kurse, spezielle Lernvoraussetzungen, Praktika etc.)
- Lernvoraussetzungen (Kursmotivation, Lernstile, Berufswunsch, Noten etc.)
- Erwartungen an den Kurs
- Objektive Lerntests zur Überprüfung allgemeiner Lernziele (z.B. Urteilsfähigkeit) und speziellen Lernzielen des Kurses

### *Schlusserhebung:*

- Subjektive Beurteilung des Lernzuwachses in Bezug auf die eigenen Erwartungen und in Bezug auf sonstige Lernzuwachsen.
- Rückmeldung über bestimmte didaktische Innovationen des Kurses (z.B. 4stündig, elektronische Medien, Portfolio, Kleingruppenarbeit etc.)
- Objektive Lerntests. Zunächst kommt ein vielfach erprobter Test zur Erfassung der moralischen Urteils- und Diskursfähigkeit zum Einsatz (Lind, 2002). Später sollen alternativ oder ergänzend auch Tests zur Erfassung anderer Lernqualitäten (siehe oben) eingesetzt werden, die sowohl fachübergreifend wie auch fachspezifisch sein sollen.

### **Schluss und Ausblick**

Das ITSE-Projekt hat sich in den qualitativen Erprobungen mit einigen Kollegen gut bewährt. Da der Finanzrahmen des Projekts sehr eng war, wies das ITSE-Programm im Einsatz noch einige Schwächen auf. Als positiv wurde aber vermerkt, a) dass die Studierenden das Programm gut annahmen und b) dass der Aufwand für die Lehrenden praktisch nicht ins Gewicht fiel.

Mit einem Folgeprojekt sollen die Voraussetzung geschaffen werden, dass auch die Rückmeldung von Ergebnissen an die (dazu berechtigten) Lehrenden und Studierenden automatisch über das Internet erfolgt. Ebenso soll der größere Teil des Erhebungsinstruments von den Lehrenden nach eigenen Wünschen und Kriterien gestaltet werden können, wofür aber Beratung durch das Expertenteam des Projekts zur Verfügung gestellt werden soll. Neben dem individuellen Angebot soll es auch möglich sein, dass nach Abschluss von einzelnen Methodenevaluationen andere Lehrende (anonymisierten) Einblick in die Ergebnisse erhalten. Damit soll zur Nachahmung angeregt und Vertrauen in diese Art der Lehrevaluation geschaffen werden.

#### a) Ausbau des ITSE-Programms

*Datenanalyse:* Zur Zeit können im ITSE-Programm vom Dozenten Freiantworten nach Kurs geordnet abgerufen werden sowie die Daten in einem allgemein kompatiblen Format (Textdatei mit Strichpunkt-Trennzeichen) zur Weiterverarbeitung in einem gängigen Statistikprogramm ausgegeben werden. Damit sind bereits vielfältige Anwendungen möglich. Jedoch besitzt nur ein Teil der Dozenten Zugriff auf das erforderliche Wissen, um die Daten zu verarbeiten und anspruchsvolle Analysen selbst durchzuführen. Es war daher von Anfang vorgesehen, eine Reihe von Standardmethoden zur Analyse der anfallenden ITSE-Daten vom Programm "auf Knopfdruck" berechnen zu lassen. Auf Grund der begrenzten Mittel des Projekts konnte das jedoch nicht realisiert werden. Priorität beim Programmieren hatten

bislang die Datensicherheit und Entlastung der teilnehmenden Lehrenden. Momentan bauen wir einfache statistische Analysen und graphische Darstellungen in das ITSE-Programm ein.

Nachdem für das Vertrautmachen mit der neuen Technologie der Internet-basierten Selbst-Evaluation eine weitgehend deckungsgleiche Fragestellung mit einem gering variierten Fragebogen ausreichend gute Dienste leistete, steigt bei allen Beteiligten nun das Interesse für eine stärker den individuellen Bedürfnissen angepasste Fragestellung und ein dementsprechend individuelleres Evaluationsdesign. Das soll ermöglicht werden durch eine Automatisierung der Erstellung und Verwaltung von Fragen und Test, so dass evaluationswillige Dozenten ihre Befragungsinstrument einfach und schnell anpassen und aus fertigen Vorlagen zielgenau zusammenstellen können. Im laufenden Projekt konnten bereits konzeptionelle Vorarbeiten und Designstudien getätigt werden. Die programmtechnische Umsetzung muss dem zweiten Projektabschnitt vorbehalten bleiben.

Für diese Arbeiten konnte mit Kristian Scholz eine qualifizierte Hilfskraft gewonnen werden, mit der diese doch sehr anspruchsvollen Programmierziele erreichbar erscheinen.

#### b) Beratungsangebot

Aus dem Kreis der Tester der ITSE-Evaluation wurde der Wunsch laut, neben der technischen Hilfe auch mehr Beratung bei der Präzisierung von Evaluationszielen, bei deren Umsetzung in Fragen und Tests sowie bei der Interpretation und praktischen Anwendung der Ergebnisse zu erhalten. Wir haben versucht, diesen Wünschen schon jetzt nachzukommen. Aus Kapazitätsgründen konnte das aber nur in einem bescheidenen Umfang geschehen. Im Fortsetzungsprojekt soll/muss dies ausgiebiger geschehen, um den Nutzen des ITSE-Programms für die Qualitätsverbesserung der Lehre voll auszuschöpfen zu können.

Das Beratungsangebot soll im Rahmen des Projekts vom Autor erbracht werden, der vielfältige Erfahrungen und seine Expertise als Bildungsforscher und Pädagogischer Psychologe einbringen kann. Dies wäre ihm wegen der hohen Lehrverpflichtungen aber nur möglich, wenn er durch eine geprüfte Hilfskraft darin unterstützt und von anderen Aufgaben teilweise entlastet würde.

## Literatur

- Amrein, A. L. & David C. Berliner, 2002: High-stakes Testing, Uncertainty, and Student Learning. *Education Policy Analysis*, 10(18) <http://epaa.asu.edu/epaa/v10n18/>
- Balk, M. (2000). *Evaluation von Lehrveranstaltungen: Die Wirkung von Evaluationsrückmeldungen*. Frankfurt: Lang.
- Balsler (2004). Online Delphie-Studie "Wie werden Evaluationsprojekte erfolgreich?" ... mehr
- Baltes, P.B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, 11, 145-171.
- Berliner, D.C. & Biddle, B.J. (1995). *The manufactured crisis. Myths, fraud, and the attack on America's public schools*. Reading, MA: Addison-Wesley.
- Bracey, G.W., (2002). *The war against America's public schools. Privatizing schools, commercializing education*. Boston: Alyn & Bacon. ISBN 0-321-08073-4 ... mehr
- Centrum für Hochschulentwicklung (2002). *Einführung einer leistungsorientierten Professorenvergütung an der Universität Konstanz*.  
[http://www.che.de/projekte.php?strAction=show&PK\\_Projekt=244](http://www.che.de/projekte.php?strAction=show&PK_Projekt=244)
- Coleman, J. & MacKeachie, W.J. (1981). Effects of instructor/course evaluation on student course selection. *Journal of Educational Psychology*, 73, 224-226.
- Dippelhofer-Stiem, B. & Lind, G. (1987). *Studentisches Lernen im Kulturvergleich. Ergebnisse einer international vergleichenden Längsschnittstudie zur Hochschulsozialisation*. Weinheim: Deutscher Studienverlag.
- Dohmen, G., Hg., (1971). *Forschungstechniken für die Hochschuldidaktik*. München: Beck-Verlag.
- Expertenkommission „Reform des Hochschuldienstrechts“ (2000). *Pressemitteilung 57/2000 am 10.04.2000 des BM für Bildung und Forschung*. <http://www.bmbf.de/press/91.php>
- Feldman, K. & Newcomb, Th.M. (1970). *The impact of college on students. Volume II*. San Francisco: Jossey-Bass.
- Finn, J.D. & C.N. Achilles, C.N. (1999). *Tennessee's Class Size Study: Findings, Implications, Misconceptions*. In: *Educational Evaluation and Policy Analysis*, summer, S. 97-109.
- Hage, N.e. (1996). *Lehrevaluation und studentische Veranstaltungskritik. Projekte, Instrumente und Grfundlagen*. Bonn: BM Bildung, Wissenschaft, Forschung und Technologie.
- Institut für Schulentwicklung PH Schwäbisch Gmünd, Hg. (2004). *Standards, Evaluation und neue Methoden. Reaktionen auf die PISA-Studie*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Kempfert, G. & Rolff, H.-G. (2000). *Pädagogische Qualitätsentwicklung. Ein Arbeitsbuch für Schule und Unterricht*. Beltz: Weinheim.
- Kohn, A. (1999). *Punished by rewards. The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin.
- Kohn, A. (2000). *The case against standardized testing. Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.

- Kromrey, H. (1994). Wie erkennt man gute Lehre? Was studentische vorlesungsbefragungen (nicht) aussagen. *Empirische Pädagogik*, 2, 153-168.
- Krueger, Alan B., 2000; Krueger & Whitmore, 1999, Quelle:  
[http://www.irs.princeton.edu/pubs/working\\_papers.html](http://www.irs.princeton.edu/pubs/working_papers.html)
- Kultusministerkonferenz (1999). Leistungsorientierte Besoldung der Hochschullehrer. Fortschreibung des KMK – Diskussionspapiers vom 28.05.1999.
- Landesinstitut für Schule und Weiterbildung (1996). Evaluation und Schulentwicklung: Ansätze, Beispiele und Perspektiven aus der Fortbildungsmaßnahme Schulentwicklung und Schulaufsicht. 2. Auflage. Soest: Kettler.
- Levin, H.M. (2000). Cost-effectiveness analysis as an evaluation tool. Manuscript. To be published in: *International Handbook of Educational Evaluation* (forthcoming).
- Lind, G., Nielsen, A. & Schmid, U. (1979). Moralische Urteile und Hochschulsozialisation. Materialien & Beiträge. SFB Bildungsforschung.  
<http://www.ub.uni-konstanz.de/kops/volltexte/2000/537/>
- Lind, G. (1980). Toward an evaluation of the impact of university on students' moral development. FORM-Bericht. Universität Konstanz.
- Lind, G. (1984). Inhalt und Struktur des moralischen Urteils. Theoretische, methodologische und empirische Untersuchungen zur Moral- und Demokratiekompetenz bei Studierenden. Dissertation an der Sozialwissenschaftlichen Fakultät der Universität Konstanz. KOPS: <http://www.ub.uni-konstanz.de/kops/volltexte/2000/504/> .
- Lind, G. (1985). Attitude change or cognitive-moral development? How to conceive of socialization at the university. In: G. Lind, H.A. Hartmann & R. Wakenhut. Eds., *Moral development and the social environment. Studies in the philosophy and psychology of moral judgment and education*, pp. 173-192. Chicago: Precedent Publishing Inc.
- Lind, G. (1990). Wirksamkeit von Ansätzen zur Förderung der moralisch-demokratischen Urteilskompetenz nach Lawrence Kohlberg. In: U. Steffens & T. Bargel, Hg., Arbeitskreis "Qualität von Schule". Heft 7: Erziehung und Demokratie in der Schule - Konzepte und Erfahrungen zum Ansatz von Kohlberg und Oser. Wiesbaden: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- Lind, G. (2000a). The importance of role-taking opportunities for self-sustaining moral development. *Journal of Research in Education*, 10, 9-15.
- Lind, G. (2000b). Moral regression in medical students and their learning environment. *Revista Brasileira de Educacao Médica*, 24(3), 24-33.
- Lind, G. (2000c). Von der Praxis zur Theorie. Eine Neubestimmung der Funktion der Praxis für die Lehrerbildung. Vortragsmanuskript.
- Lind, G. (2002). Ist Moral lehrbar? Ergebnisse der modernen moralpsychologischen Forschung. 2te Auflage. Berlin: Logos.
- Lind, G. (2003). Moral ist lehrbar. Ein Handbuch zur moralischen und demokratischen Bildung. München: Oldenbourg.
- Lind, G. (2004a). Jenseits von PISA — Für eine neue Evaluationskultur, S. 1 - 7. In: Institut

- für Schulentwicklung PH Schwäbisch Gmünd, Ed., Standards, Evaluation und neue Methoden. Reaktionen auf die PISA-Studie. Baltmannsweiler: Schneider Verlag Hohengehren.
- Lind, G. (2004b). The meaning and measurement of moral judgment competence -- A dual aspect theory. In: D. Fasko & Willis, eds., Contemporary Philosophical and Psychological Perspectives on Moral Development and Education. Creskill, NJ: Hampton Press (im Druck).
- Mclaughlin, M.W. & Phillips, D.C. (1991). Evaluation and education: At quarter century. Ninetieth Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.
- McKeachie, W.J. (1999). Teaching Tips. Strategies, research and theory for college and university teachers. Boston: Houghton Mifflin.
- Millman, J. & Darling-Hammond, L., eds. (1990). The new handbook of teacher evaluation. Assessing elementary and secondary school teachers. Newbury Park, CA: Corwin Press.
- Murphy, Madonna (2002). Evaluating character education initiatives: Can it be done? Boston: Scarecrow Education.
- OECD (2001). Lernen für das Leben. Erste Ergebnisse der Internationalen Schulleistungsstudie PISA 2000. Paris: OECD. Deutsche Fassung im Auftrag des Bundesministeriums für Bildung und Forschung. Kostenloser Bezug von: <http://www.gew.de> und <http://www.pisa.oecd.org>.
- Pascarella, E.T. & Terenzini, P.T. (1991). How college affects students. San Francisco: Jossey-Bass.
- Pascarella, E.T., Whitt, E.J., Noral, A., Edision, M. & Hagedorn, L.S. (1995). What have we learned from the first year of the National Study of Student Learning? Unveröff. Ms., University of Illinois at Chicago, College of Education.
- Popham, W.J. (1974). Evaluation in education. Current applications. Berkeley, CA: McCutchan Publishing Co.
- Popper, K.R. (1968). The Logic of Scientific Discovery. London Hutchinson (Original 1934).
- Rindermann, H. (2001). Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungen an Hochschulen. Landau: Verlag Empirische Pädagogik.
- Rolff, H.-G. (2001). Schulentwicklung konkret - Steuergruppen, Bestandaufnahmen, Evaluation. Hannover: Kallmeyer.
- Sacks, P. (1999). Standardized minds. The high prize of America's testing culture and what we can do to change it. Cambridge, MA: Perseus Publishing.
- Sanders, J.R. (1994). The program evaluation standards. How to assess evaluations of educational program. 2nd edition. Thousand Oaks, USA: Sage Publications.
- Schmidt, C.M & Fertig, M. (2003). Genaueres Hinsehen lohnt. Die Determinanten des Abschneidens deutscher Schüler an der PISA 2000-Studie. In Forschung & Lehre, 6/2003, 313-315. ... mehr
- Scriven, M. (1999): The Nature of Evaluation Part I: Relation to psychology.

<http://ericae.net/pare/getvn.asp?v=6&n=11>

Scriven, M. (1999): The Nature of Evaluation Part II: Training.

<http://ericae.net/pare/getvn.asp?v=6&n=12>

Shepard, L.A. (2002). The role of assessment in a learning culture. *Educational Researcher*, 29, 4-14. (PDF)

Smith, F. (1986). *Insult to intelligence. The bureaucratic invasion of our classrooms*. New York: Arbor House.

Thonhauser, Josef & Patry, Jean-Luc (1999). *Evaluation im Bildungsbereich: Wissenschaft und Praxis im Dialog*. Innsbruck/Wien: Studien-Verlag.

Weinert, F.E., Hg. (2001). *Leistungsmessung in Schulen*. Weinheim: Beltz. Dort vor allem die Beiträge von Weinert (Kap. 1), Brügelmann (Kontroversen), Reinberg (Kap. 4, Kleine Beurteilungsaufgabe), Arnold (Kap. 8), Helmke & Schrader (Kap. 16), Baumert et al. (Kap. 19).

Witte, J. & Schreiterer, U. (2002). *Perspektiven für eine leistungsorientierte Professorenbesoldung im Lichte internationaler Erfahrungen*. Arbeitspapier zur Dienstrechtsreform II. Güterslog: Centrum für Hochschulentwicklung.

## Anmerkungen

1. Die Deutsche Gesellschaft für Qualitätsmanagement schreibt zum Ziel von TQM: "TQM ist eine auf Mitwirkung aller Mitglieder beruhende Führungsmethode einer Organisation, die Qualität in den Mittelpunkt stellt und durch Zufriedenheit der Kunden auf langfristigen Geschäftserfolg sowie auf Nutzen für die Mitglieder der Organisation und für die Gesellschaft zielt." Siehe auch im Internet: Total Quality Management and ISO 9000 web-ressources, bezogen von <http://www.gslis.utexas.edu/~rpollock/tqm.html> (9.10.2004) und die Deming Website: <http://caes.mit.edu/deming/> (9.10.2003).
2. Bestrafung kann keinen Mangel an pädagogisch-didaktischen Kompetenzen ausgleichen. Im Gegenteil, wer wegen schlechter Lehrleistungen mit Geldentzug und vermehrten Aufgaben bestraft wird, hat noch weniger die Möglichkeit, seine Lehrleistung durch Fortbildung zu steigern, sondern wird versuchen, die Evaluationskriterien ohne großen Aufwand zu erfüllen. Wirkliche Lehrkompetenzen müssen in einem Zeit- und Ressourcenaufwendigen Prozess erworben werden. Die stärksten Anreize dafür bilden eine interessante und abwechslungsreiche Arbeit, soziale Anerkennung und Möglichkeiten für Verantwortungsübernahme (Kohn, 1999; Deming, 1994; Frey, 2003).
3. Zu diesen verschiedenen Formen von Kompetenz vgl. Gardner (1983) und auch den Antrag zum ersten Projektabschnitt sowie die Projektwebseite ([www.uni-konstanz.de/itse](http://www.uni-konstanz.de/itse)) und die Web-Seite des Autors zum Thema Evaluation in Bildungsinstitutionen ([www.uni-konstanz.de/ag-moral/evaluation](http://www.uni-konstanz.de/ag-moral/evaluation)).