

# EVALUATION DER LEHRE

## ALS MASSNAHME ZUR QUALITÄTSSICHERUNG IN DER LOGOPÄDIEAUSBILDUNG

Vorstellung des Reichenauer Modells  
und seine Effektivitätsprüfung

Schlüsselwörter:  
Lehrevaluation,  
Logopädieausbildung,  
Qualitätssicherung,  
Dozentenberatung

### Z u s a m m e n f a s s u n g

Die Qualität der Lehre lässt sich durch Ausbildungseinrichtungen auf verschiedenen Wegen sichern. Der Beitrag stellt in zwei Teilen den Evaluationsansatz der Schule für Logopädie auf der Reichenau bei Konstanz vor – eine Kombination aus qualitativen und quantitativen Methoden.

Der hier vorliegende erste Teil konzentriert sich auf die Einordnung des Evaluationsprojekts unter die Begriffe „Qualitätssicherung“ und „Evaluation“.

Es werden die einzelnen Schritte des Projekts beschrieben, die Ziele sowie die benutzten Instrumente vorgestellt. Eine detaillierte Darstellung der quantitativen Ergebnisse schließt den ersten Teil des Artikels ab.

Der zweite Teil in der nächsten Ausgabe von I.O.G.O.S. interdisziplinär konzentriert sich auf die qualitativen Ergebnisse des Projekts durch Analyse von Einzelgesprächen mit Dozenten<sup>1</sup> und von Konferenzprotokollen.

<sup>1</sup> Aus Anonymisierungsgründen wird im Folgenden ausschließlich die maskuline Sprachform verwendet.

## 1. Warum Qualitätssicherung?

In den 90er-Jahren setzte in verschiedenen Dienstleistungsbereichen eine Entwicklung ein, die Qualität und den Nachweis der Qualität von Dienstleistungen ins Zentrum des Interesses rückte. Vor allem im tertiären Bildungssektor (betriebliche, berufliche und akademische, öffentliche und private Aus- und Weiterbildung) werden an Institutionen erhöhte Ansprüche zur Qualitätssicherung gestellt, neben den Hochschulen sind hier im Besonderen berufsbildende Einrichtungen zu nennen (Keuper u.a., 1995). Zu diesen zählen auch die Lehranstalten für Logopädie: Nicht nur die logopädische Behandlung muss bestimmten Kriterien genügen, sondern auch die Logopädieausbildung stellt eine Leistung dar, die sich der Prüfung ihrer Qualität stellen muss. Der Nachweis der Ausbildungsqualität liegt im Interesse vieler:

- Ausbildungswillige benötigen Informationen, um zwischen verschiedenen Qualifikationswegen und Ausbildungsstätten fundiert wählen zu können,
- Lehranstalten möchten ihre Qualität nach außen darstellen,
- Patienten haben ein Recht auf eine hochwertige Sprachtherapie durch die Absolventen,
- die schon examinierten Kollegen wollen die öffentliche Reputation ihres Berufsstandes durch Qualitätssicherung in der Ausbildung weiter stärken.

Diese verschiedenen Perspektiven lassen sich im gesteigerten Interesse an der Qualität der Ausbildung bündeln. Um diese Qualität der Logopädieausbildung zu dokumentieren und zu verbessern, wurden in den letzten Jahren von Vertretern des Berufsstandes einige Maßnahmen vorgeschlagen (vgl. dbl, 1996a,b). Der nun hier vorgestellte Evaluationsansatz stellt einen Versuch zur Verbesserung der Lehre im Rahmen der Logopädieausbildung

durch Evaluation, Beratung und institutsinterne Qualitätsentwicklungsmaßnahmen dar.

## Evaluation

Unter Evaluation („Bewertung, Beurteilung“) ist die systematische Analyse und empirische Untersuchung von Konzepten, Bedingungen, Prozessen und Wirkungen zielgerichteter Aktivitäten zum Zwecke ihrer Bewertung und Modifikation zu verstehen. Gegenstand von Evaluationen sind Programme, Verfahren, Produkte, Institutionen und Personen. Über adäquate Messung und Beschreibung hinausgehend bewerten Evaluationen und optimieren Prozesse. Formative Ansätze legen hierbei den Schwerpunkt auf die Verbesserung von laufenden Verfahren, summative Ansätze nehmen über die Bereitstellung fundierter Informationen Wirkungsbeurteilungen vor und optimieren so Entscheidungen.

Evaluation zielt auf Qualitätssicherung durch Verbesserung vorhandener Strukturen oder durch Entscheidung über Alternativen:

Evaluation dient als Planungs- und Entscheidungshilfe und hat somit etwas mit der Bewertung von Handlungsalternativen zu tun. Evaluation ist zielorientiert. Sie verfolgt primär den Zweck, praktische Maßnahmen zu überprüfen, zu verbessern oder über sie zu entscheiden (nach Wotawa & Thierau, 1998).

„Qualitätssicherung macht die Qualität der eigenen Arbeit transparent, sie hilft Defizite zu erkennen und zu beseitigen, sie verbessert den interkollegialen Informationsaustausch (...)“ (Bundesärztekammer, 1996, S. 16f.)

Der an der Schule für Logopädie verfolgte Ansatz versucht nun durch Evaluation und durch begleitende Maßnahmen die Ausbildungsqualität am Reichenauer Institut zu dokumentieren und gegebenenfalls zu verbessern. Verbesserung der Lehre impliziert nicht, dass die Lehre zuvor „schlecht“

war. Lehrqualität ist als eine dynamische Größe zu verstehen, die in gewissen Grenzen immer wieder neu durch Aushandlungsprozesse zwischen Dozenten, Schulleitern, Studenten und Abnehmern von Absolventen zu definieren ist.

## 2. Lehrevaluation an der Logopädiehochschule Reichenau

Der nachfolgend beschriebene Evaluationsansatz in der Logopädieausbildung fokussiert einen zentralen Teil der Lehre, nämlich den „theoretischen und praktischen Unterricht“. Dieser umfasst die in der Ausbildungsordnung festgelegten 20 Fächer mit insgesamt mindestens 1740 Unterrichtsstunden (Ständige Konferenz, 1993). Die Inhalte kommen besonders aus dem medizinischen und sozialwissenschaftlichen Bereich wie z.B. „Anatomie“ oder „Hals-Nasen-Ohrenheilkunde“ bzw. „Psychologie“ oder „Pädagogik“. Es gibt aber auch Fächer mit starker praktischer Ausrichtung wie z.B. „Stimmbildung“. Die so genannte „praktische Ausbildung“ mit insgesamt mindestens 2100 Unterrichtsstunden war nicht Gegenstand unserer Evaluation. Die Teilnahme am Projekt war sowohl für Dozenten als auch für Studierende freiwillig, Anonymität wurde zugesichert, Fragebögen sind nicht auf einzelne Studenten zurückführbar, Dozenten erhielten nur Ergebnisse ihrer Veranstaltungen und ein Gesamtergebnis aller Kurse (letzteres das gesamte Institut). Zur Messung und Verbesserung der Lehre wurde eine Mischung aus quantitativen und qualitativen Methoden gewählt (Lamnek, 1995b, S. 247).

### 2.1 Einzelne Schritte einer Lehrevaluation

1. Am Anfang stehen eine Einführung in und Werbung für „Qualitätssicherung durch Evaluation“ bei Lehrenden, Studierenden und Entscheidungsträgern durch Gespräche in Team-Sitzungen, Dozenten- und Studentenkonferenzen im Mittelpunkt.

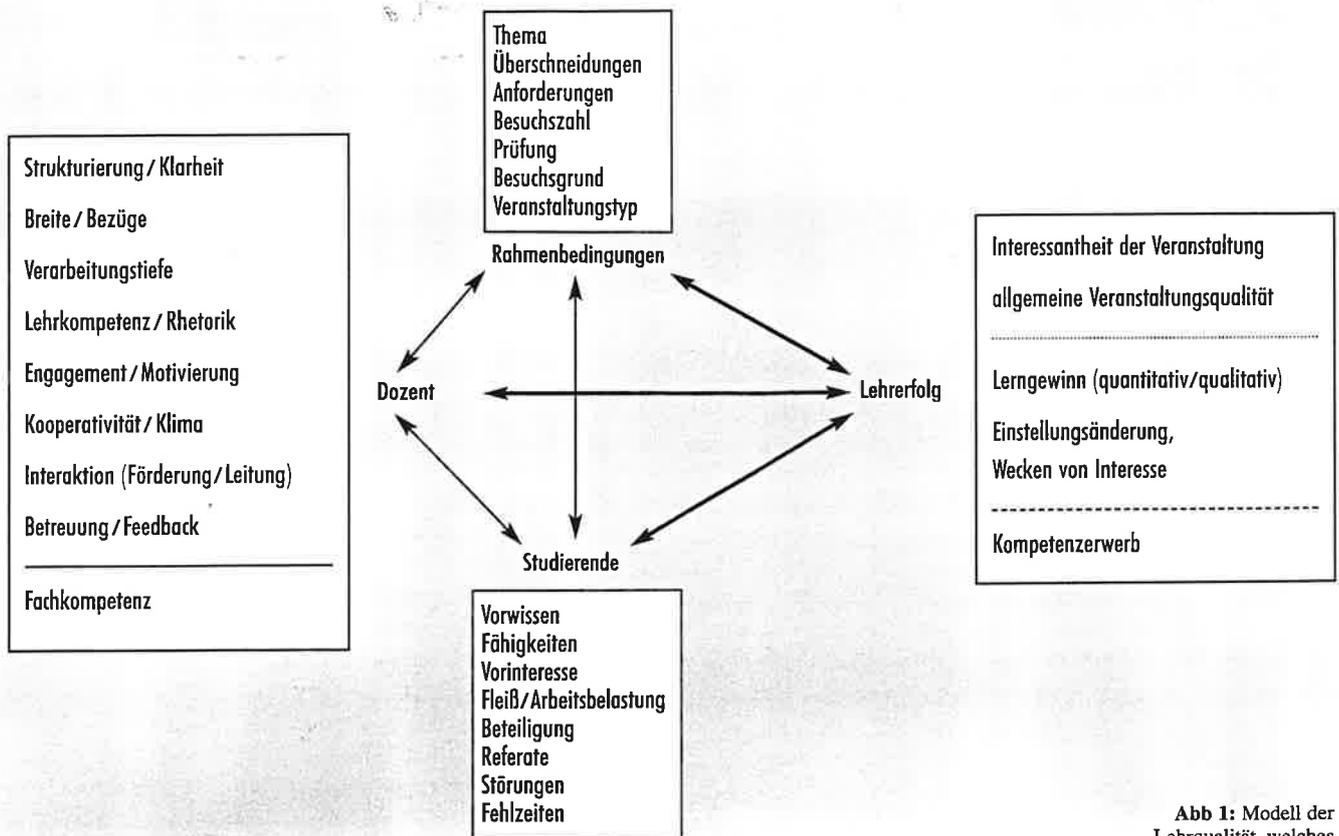


Abb 1: Modell der Lehrqualität, welches der modifizierten Form des HILVE-Fragebogens zu Grunde liegt

2. Auszubildende beurteilen ihre Dozenten im theoretisch-praktischen Unterricht mithilfe eines standardisierten Fragebogens, der aus vorgegebenen Skalen sowie offenen Fragen besteht. Die Auswertung der Fragebögen erfolgt nach wissenschaftlichen Kriterien durch einen Unabhängigen außerhalb der Einrichtung (hier: der Zweitautor dieses Beitrags).

3. Die individuellen Ergebnisse werden dem einzelnen Lehrer in Form eines ca. einstündigen Gesprächs von der leitenden Lehrkraft (Erstautor) übermittelt. Das Gespräch hat außerdem beratende Funktion (s.u.). Der Lehrer wird zur Besprechung seiner Ergebnisse in der Klasse angeregt.

4. Die allgemeinen Ergebnisse oder beispielhaft auch individuelle Ergebnisse werden in anonymisierter Form im Rahmen von Dozenten- oder Studierendenkonferenzen dargestellt und diskutiert.

Die Schritte zwei bis vier werden fortlaufend wiederholt, Evaluation und

Modifikation der Qualität der Lehre stellen einen kontinuierlichen, alternierenden Prozess dar. Nach dem vierten Schritt hat jeder einzelne Lehrer die Möglichkeit, weitere Maßnahmen bis zum erneuten Einsatz des Fragebogens (Wiederholung des 2. Schrittes) zu ergreifen. Vonseiten der Schulleitung und des Schulträgers bestehen hierzu folgende Angebote:

- Unterrichtshospitation bei oder durch andere Dozenten,
- externe Fortbildung, die speziell für Lehrkräfte in der Erwachsenenbildung im Bereich des Gesundheitswesens konzipiert wurde,
- weitere Gespräche nach tieferer Verarbeitung der Ergebnisse. Dabei stehen folgende Fragen im Vordergrund: Wurden Veränderungen in Unterrichtsmethoden vorgenommen und haben diese etwas bewirkt? Zeigte das Ergebnis gefühlsmäßige Auswirkungen? Wurde auf Grund der Ergebnisse eine Diskussion mit der Klasse geführt und wenn ja, welche neuen Erkenntnisse hat sie erbracht?

## 2.2 Ziele

Wichtigstes Ziel der Maßnahme ist die Verbesserung der Lehrqualität in den einzelnen Fächern der theoretisch-praktischen Ausbildung (Kohler & Rindermann, 1998). Die Dozenten erhalten durch die systematische Rückmeldung eine Chance, ihr Lehrverhalten zu reflektieren und gegebenenfalls zu verändern. Das Feedback dient der Konfrontation mit der Fremdwahrnehmung eigenen Handelns. Bei kritischer Rückmeldung wird das Gespräch zur Beratung und dient der Stützung sowie dem Aufbau neuer Lehrtechniken und pädagogischer Handlungsweisen. Es werden gemeinsam Möglichkeiten der Veränderung erörtert und auch konkrete Vorschläge unterbreitet (z.B. zeitweise Kleingruppenarbeit praktizieren statt durchgehendem Vorlesungsstil).

Daneben ist der indirekte Einfluss der Evaluation auf das Gesamtsystem Schule wichtig. Die Kommunikation zwischen Schülern und Lehrern über die Lehre erhält eine sachliche Grundlage. Die Möglichkeit der Einflussnahme der Schülerschaft auf die Lehrqualität soll die Identifikation mit der Schule erhöhen. Das gegenseitige Bewerten von Dozenten und Schülern soll Partnerschaftlichkeit und dialogische Gesprächsformen im Lehr-Lernprozess fördern. Schließlich werden auch Effekte innerhalb der Subsysteme Lehrerschaft bzw. Schülerschaft erwartet. Der Austausch zwischen den Dozenten soll hierdurch intensiviert werden, es wird erhofft, dass sie sich gegenseitig stützen. Einzelne Schüler können ihre Unterrichtsperzeption im Vergleich mit anderen Einschätzungen betrachten und eventuell relativieren. Es soll über die reinen Lerninhalte hinaus auf allen Seiten des Lehr-Lernprozesses Verantwortungsbewusstsein für die Ausbildungsqualität geschaffen werden.

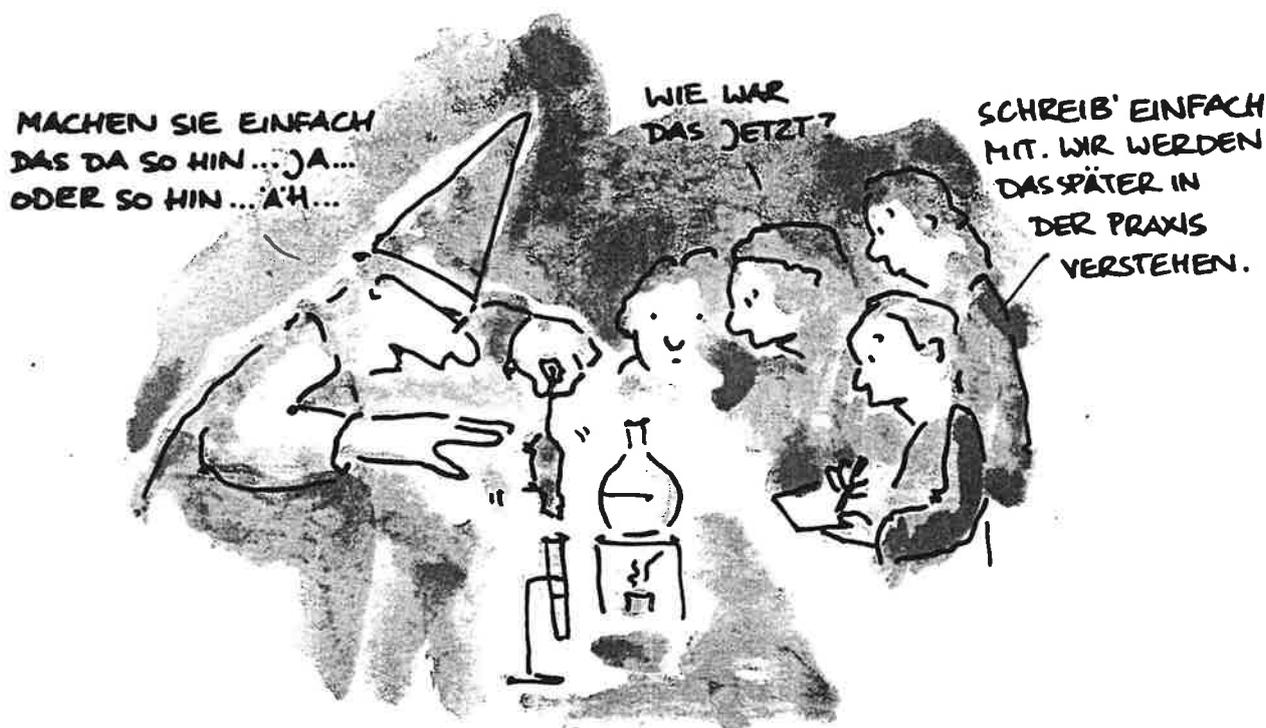
## 2.3 Instrumente

### 2.3.1 Fragebogen

Im eingesetzten Fragebogen und in seinen Skalen spiegeln sich das Modell der Lehrqualität und die Kriterien der Evaluation wider. Die Lehre wird als interaktiver Prozess aufgefasst, zu dessen Gelingen Dozenten, Studierende und Rahmenbedingungen beitragen. Beurteilt werden deshalb das Handeln der Lehrkräfte, Verhalten und Voraussetzungen der Studierenden und äußere Determinanten des Lehrerfolges. Es wird hierfür eine weiterentwickelte Version des Heidelberger Inventars zur Lehrveranstaltungsevaluation „HILVE“ (Rindermann & Amelang, 1994; Rindermann, 1999; Buhl, 1999) verwendet. Dies ist ein Instrument, das in seiner ursprünglichen Version zur Evaluation von Lehrveranstaltungen an der Universität konzipiert und eingesetzt wurde. Zur Anpassung an den Unterricht in einer Schule für Logopädie wurden kleine Veränderungen vorgenommen. Die grundlegenden Dimensionen des HIL-

VE und seine empirisch gefundenen Gütemaße bleiben aber erhalten. Abbildung 1 zeigt das Modell der Lehrqualität, welches der modifizierten Form des HILVE-Fragebogens zu Grunde liegt (vgl. Rindermann, 1998; 1999).

Insgesamt besteht der Fragebogen aus 53 geschlossenen Einzelitems, die zu 19 Dimensionen zusammengefasst werden (Dimensionen s. Anhang). Die Antwortskala variiert zwischen 1 (trifft überhaupt nicht zu) und 7 (trifft genau zu). Zudem besteht für die Studierenden die Möglichkeit zur freien, ungebundenen Bewertung sowie zur Formulierung von Verbesserungsvorschlägen (was ist besonders positiv, was negativ, Anregungen). Die Dozenten erhalten ihr individuelles Ergebnis durch absolute und normierte Mittelwerte mit Streuungsmaßen pro Item bzw. pro Dimension in Form von Zahlenwerten und Grafiken (vgl. hierzu auch Abbildung 1 im nächsten Heft). Normiert wurde mithilfe der in Reichenau gefundenen Ergebnisse. Das heißt, der Dozent erfährt sein Er-



Meister + Schindler  
 Smiltes

gebnis im Vergleich zu anderen Dozenten an der Schule für Logopädie. Die normierten Daten weisen einen Mittelwert von  $M = 100$  und eine Streuung von  $SD = 10$  auf. Ein individueller Wert von  $i = 90$  wäre somit als ein vergleichsweise schwacher,  $i = 110$  hingegen als ein guter Wert aufzufassen. Die Rohwertmittel (bei einer Skala von 1 bis 7) liegen in der Regel zwischen  $i = 5$  und  $i = 6$ , d.h. im Schnitt wird die Lehre am Reichenauer Institut durch die Schüler günstig bewertet.

### 2.3.2 Kommunikation

Das Rückmeldegespräch zwischen leitender Lehrkraft und evaluiertem Dozent sowie die Dozenten- und Studierendenkonferenzen bilden den interaktionalen Rahmen des Evaluationsprozesses. Evaluationsforschung wird als kommunikativer Prozess zwischen Evaluator, Auftraggeber, Betroffenen, Beteiligten und Kunden verstanden (vgl. Beywl, 1988), in diesem Falle zwischen dem Leiter der Schule für Logopädie (gleichzeitig Durchführender der Evaluation und Berater), den Lehrenden und Studierenden und dem

externen Wissenschaftler (Erstellung des Instrumentariums und Auswertung der studentischen Bewertungen).

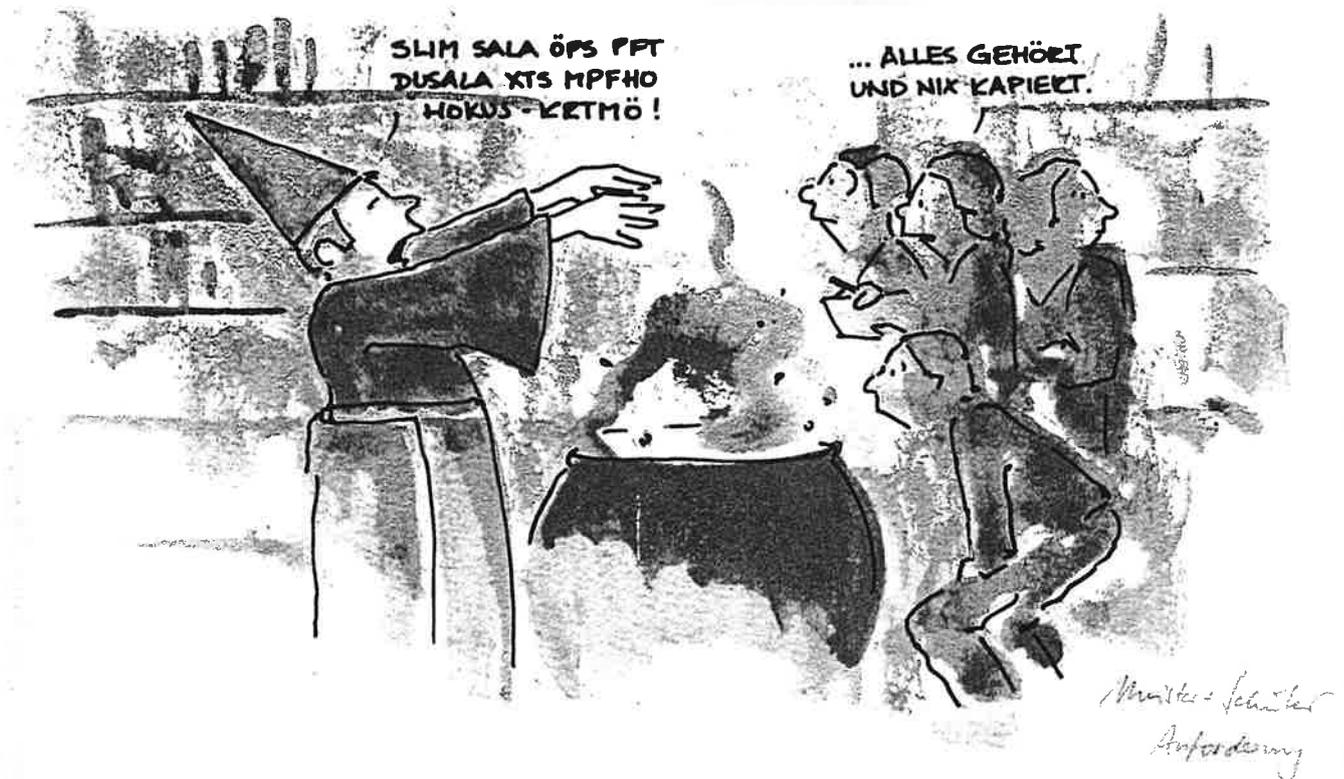
Wichtige Merkmale der Kommunikation innerhalb der Sozialforschung sind nach Lamnek (1995a):

- „Das Einfühlen in die Situation der Betroffenen, das Schaffen eines Vertrauensverhältnisses in der Kommunikationssituation sind wesentliche Voraussetzungen für verlässliche und gültige Befunde.“ (Lamnek, 1995a, S. 58)
- „Vorfremdende Einflüsse durch eine ungewöhnliche, unnatürliche Kommunikationssituation während der Erhebung führen auch zu verfremdeten und unnatürlichen Interpretationen und Deutungen (...). Daher müssen kommunikative Akte eingegangen werden, die Alltagssituationen möglichst ähnlich sind.“ (Lamnek, 1995a, S. 20)
- „Offenheit gegenüber den untersuchten Personen (der Forscher erklärt ihnen den Sinn und Zweck der Untersuchung) und Offenheit in der Erhebungssituation (Flexibilität in der Wahl der Erhebungstechniken

und in der Gestaltung der Erhebungssituation) sind zentrale Merkmale qualitativer Sozialforschung.“ (Lamnek, 1995a, S. 17)

### 3. Quantitative Ergebnisse der Effektivitätsprüfung

Die Wiederholungsmessung ermöglicht, Effekte der Maßnahme auf der Ebene der Lehrveranstaltungsqualität zuverlässig zu beurteilen. Bisher wurden zehn Dozenten zweimal sowie weitere acht Lehrende einmal evaluiert, haben Feed-back zu ihren Ergebnissen erhalten und sind beraten worden. Abbildung 2 zeigt die Veränderung auf den Dimensionen des Fragebogens im Durchschnitt. Die Tendenz ist eindeutig positiv: So ist bei 10 Dozenten in 20 Kursen in Lehrkompetenz eine Verbesserung von  $d = 0,82$  zu beobachten (Engagement  $d = 0,60$ ; Allgemeinurteil  $d = 0,56$ )!<sup>2</sup> Veränderungseffekte ( $d$ ) bei Messwiederholungen auf einer genormten Skala mit einem Mittelwert von  $M = 0$  und einer Standardabweichung von  $SD = 1$  sind bei  $d = 0,2$  schwach, bei  $d = 0,5$  mittelmäßig und bei  $d = 0,8$  stark (Cohen,



# Meßzeitpunktveränderung

10 Dozenten mit 20 Veranstaltungen

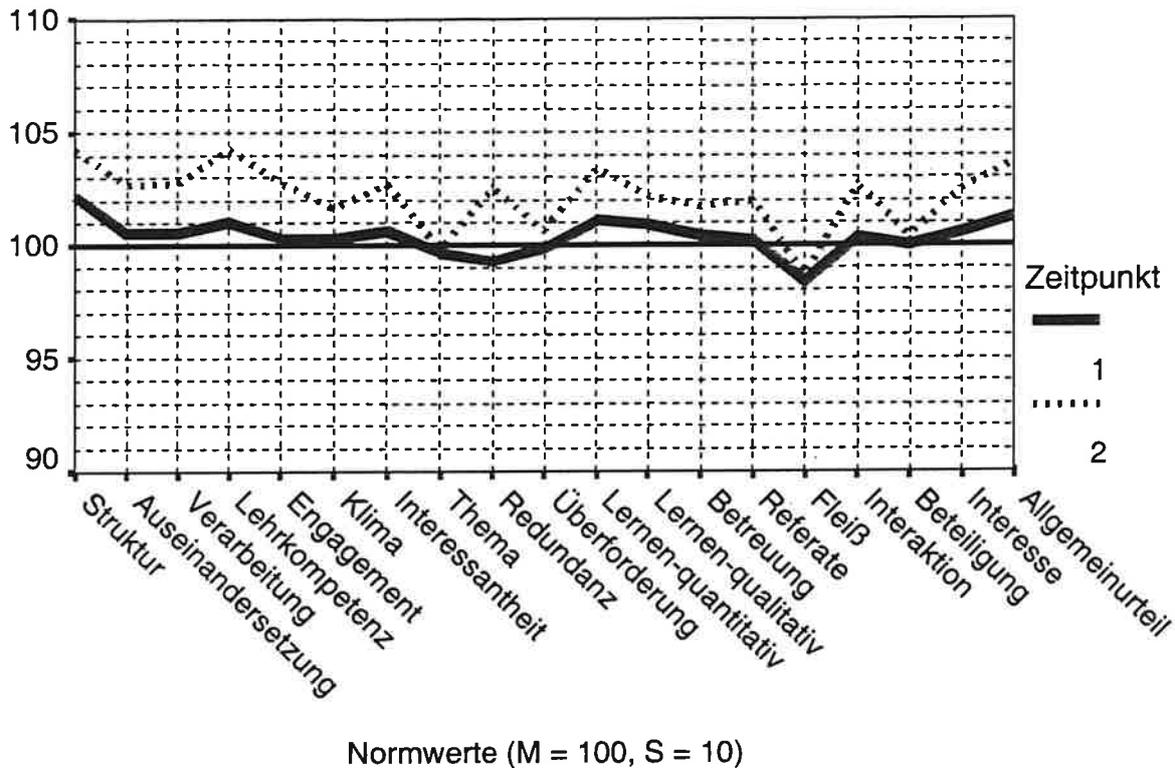


Abb. 2: Messzeitpunktveränderungen bei 10 Dozenten

<sup>2</sup> „d“ stellt ein Effektstärkemaß dar ( $d = \frac{\bar{x}_2 - \bar{x}_1}{s}$ ).

Der Mittelwert der ersten Messung wird von der zweiten abgezogen und die Differenz anschließend durch die Streuung geteilt.

1988). Alternativerklärungen wie allgemein mildere Beurteilung zum zweiten Messzeitpunkt lassen sich ausschließen, da in studentischen Skalen ( $d = 0,01$ ) oder Skalen, die Rahmenbedingungen beschreiben ( $d = 0,19$ ), beides Bereiche, die weniger Grundlage der Beratung waren und bei deren Einschätzung somit auch geringere Verbesserungen erwartbar sind, kaum Veränderungen auftreten. Es lässt sich somit festhalten, dass sich die Qualität des Unterrichts aus Sicht der Studierenden substanziell verbessert hat. Lehrevaluation in Verbindung mit Feed-back und Beratung ist effektiv.

Die Betrachtungsweise über alle Dozenten hinweg hat aber einen Nachteil.

Es gehen auch Messwiederholungswerte von denjenigen Lehrern ein, die schon bei der Erstmessung sehr gute Ergebnisse erzielt hatten. Falls jemand anfangs sehr gut bewertet wurde, gibt es – schon aus teststatistischen Gründen – kaum eine Möglichkeit mehr zur Verbesserung. Die durchschnittliche Verbesserung unterschätzt also die tatsächliche Entwicklung zum Positiven bei den zunächst unterdurchschnittlich Bewerteten. Es ist daher interessant, die kritisch bewerteten Dozenten einzeln zu betrachten. Dies kann hier nur beispielhaft geschehen. Die Tabelle 1 zeigt die Messzeitpunktveränderungen von vier zweimal bewerteten Dozenten. Tabelle 1 und Abbildung 2 enthalten Resultate in unterschiedlicher Darstellungsform, Tabelle 1 zeigt relative

Werte (Veränderungen zwischen erstem und zweitem Messzeitpunkt mit dem Maß  $d$  ausgedrückt), Abbildung 2 normierte Mittelwerte. In der Tabelle 1 kann man pro Dimension die Veränderung durch den Zahlenwert leicht erkennen. Das Pluszeichen steht für Zunahme, das Minuszeichen für Abnahme. Für die meisten Dimensionen gilt, dass eine günstige Veränderung durch das Pluszeichen angezeigt wird. Lediglich die Dimension „Redundanz“ (treten inhaltliche Überschneidungen mit anderen Veranstaltungen auf?) ist anders gepolt. Hier bedeutet das Minuszeichen eine positive Entwicklung. In der Dimension Anforderungen/Überforderung sind extreme Werte (kleiner 90 oder größer 110) als ungünstig zu bewerten, da sie Unter-

**Tabelle 1: Messzeitpunktveränderungen bei vier ausgewählten Dozenten**

Dimensionen	Veränderung d des Dozenten A	Veränderung d des Dozenten B	Veränderung d des Dozenten C	Veränderung d des Dozenten D
Struktur	4,88	6,61	5,91	-2,02
Auseinandersetzung	3,31	7,38	1,94	-4,66
Verarbeitung	5,15	7,43	4,94	-4,93
Lehrkompetenz	4,45	11,04	6,50	-3,09
Engagement	2,30	7,45	7,26	-3,81
Klima	-0,17	3,98	4,04	-1,52
Interessantheit	1,28	6,51	5,56	-7,92
Thema	-2,93	-7,42	-9,68	-3,65
Redundanz	-0,44	-3,22	1,73	3,42
Anforderung	3,13	11,33	4,34	-0,37
Lernen (quant)	7,84	7,54	0,61	-6,80
Lernen (qual)	8,79	4,33	-1,68	-6,29
Betreuung	5,13	1,62	7,62	-4,13
Referate	-	-5,93	-	-
Fleiß	4,96	1,39	-1,58	-10,95
Interaktionsmanagement	2,48	7,87	8,55	-4,01
Beteiligung	0,41	2,57	-3,70	-0,94
Interessenförderung	5,48	2,23	4,71	-7,63
Allgemeinurteil	5,76	6,19	4,36	-4,38
Mittelwert Note 1. Messzeitpunkt	3,28	3,41	3,65	1,16
Mittelwert Note 2. Messzeitpunkt	2,44	2,44	2,86	1,59

**Anmerkungen:** Messzeitpunktveränderungen sind *hier* auf einer Skala mit genormtem Mittelwert  $M = 100$  und Standardabweichung  $SD = 10$  bei einer Veränderung von  $d = 2$  schwach, bei  $d = 5$  mittelmäßig und bei  $d = 8$  stark. + ist Zunahme, - Abnahme. Auf der Dimension Redundanz sind geringe Werte optimal, auf der Dimension Anforderung mittlere. Note auf Notenskala (1-6).

förderung bzw. Überforderung signalisieren.

Wie an der Note zum ersten Messzeitpunkt erkennbar (Skala 1, sehr gut, bis 6, ungenügend), wurden die Dozenten A, B und C im ersten Durchgang kritisch bewertet. Dozent D gehört dagegen zu den sehr gut bewerteten. Die Dozenten A, B und C konnten sich verbessern. Dozent D wurde beim zweiten Mal mit leichten Abstrichen weiterhin äußerst günstig bewertet.

Die Dimension „Lehrkompetenz“ erfasst allgemeine didaktische Fähigkeiten, wie die Verständlichkeit der Sprache im Unterricht, ob Kompliziertes verdeutlicht werden kann und inwiefern die Veranstaltungen gut vorberei-

tet sind. Hier konnte sich Dozent B stark verbessern. Mit „Klima“ wird die Atmosphäre in der Veranstaltung erhoben. Das Klima wird sowohl von Lehrenden als auch von Studierenden beeinflusst. Bei Dozent A war das „Klima“ schon zum ersten Messzeitpunkt schlecht und hat sich zum zweiten Messzeitpunkt noch etwas verschlechtert ( $d = -0,17$ ). „Fleiß“ soll die Mitarbeit der Studierenden erheben: Werden die gestellten Aufgaben wie ‚Literaturlesen‘ oder ‚Vorbereitung der Stunde‘ erledigt? Es ist erkennbar, dass der „Fleiß“ der Studierenden in der Veranstaltung von Dozent D deutlich abgenommen hat. Dies ist aber nicht unbedingt negativ zu bewerten. Vielleicht hat Dozent D die Struktur

seines Unterrichts derart geändert, dass die Schüler in weniger Zeit (also ohne Hausaufgaben) den Stoff vermittelt bekommen. Mit „Interaktionsmanagement“ wird beurteilt, ob der Dozent fähig ist, die Schüler-Interaktion zu moderieren bzw. zu fördern. Dozent B und C erreichen auf dieser Dimension hohe positive Effekte, sie haben sich also stark verbessert.

Das „Allgemeinurteil“ stellt schließlich die generelle Bewertung einer Veranstaltung aus Sicht der Studierenden dar. Ein Item dieser Skala umfasst die Bewertung der Veranstaltung in Form einer Schulnote. Der Vergleich zwischen den Schulnoten (Rohdaten, keine Veränderungsmaße) zwischen erstem und zweitem Messzeitpunkt bringt einen ähnlichen Überblick für die Veränderung wie der Wert für das Allgemeinurteil.

### 3.1 Fazit

Die Lehre sehr gut beurteilter Dozenten kann durch die Maßnahme (Evaluation, Feed-back, individuelle Erläuterung und Beratung) kaum verbessert werden. Bei kritisch beurteilten Dozenten sind jedoch Verbesserungen, vor allem in solchen Skalen, die direkt durch das Dozentenhandeln beeinflussbare Merkmale erheben, unübersehbar. Evaluation mit begleitenden Beratungsmaßnahmen eingebettet in ein unterstützendes und die Qualität der Lehre adäquat honorierendes Umfeld zeigt positive, quantitativ nachweisbare Veränderungen (vgl. Marsh & Roche, 1994). Im Gegensatz zu Hochschulen, in denen die förderlichen Begleitumstände der Lehrevaluation meistens fehlen (Rindermann, 1996; Rindermann, 1997), zeigten sich an einer Schule für Logopädie in privater Trägerschaft inhaltlich bedeutsame Erfolge. Systematische Differenzen zwischen sehr gut bzw. gut beurteilten Lehrern und kritisch bewerteten Dozenten bleiben aber bestehen.

Die gemessenen Effekte zeigen nachweisbare erfreuliche Auswirkungen der durchgeführten Maßnahme. In der nächsten Ausgabe von L.O.G.O.S. in-

## Dimensionen des HILVE-II (Lehrinventar-Logopädie)

**Struktur** beschreibt Aufbau und Organisation der Veranstaltung.

**Auseinandersetzung** fragt nach einer erläuternden Behandlung des Stoffes. Beispiele und ein Theorie-Praxis-Bezug vertiefen das Thema und zeigen die Relevanz des Stoffes.

**Verarbeitung:** Der Stoff wird nicht nur wissenschaftlich vermittelt, sondern Schüler werden zum Mitdenken motiviert, Problemstellungen werden kritisch herausgearbeitet.

**Lehrkompetenz** fragt danach, ob der Dozent in der Lehre überzeugt. Spricht er oder sie verständlich, kann er Kompliziertes verdeutlichen, wird der Stoff regelmäßig zusammengefasst und geht er gut vorbereitet in Veranstaltungen?

**Dozentenengagement** erhebt Motivationsvariablen. Motiviert die Lehrkraft, nimmt sie den Erfolg der Lehre wichtig?

Mit **Klima** wird die Atmosphäre in der Veranstaltung erhoben. Gibt es einen guten Umgang zwischen Dozent und Schülern, ist der Dozent kooperativ? Das Klima wird sowohl vom Dozenten als auch von den Schülern beeinflusst.

**Interessantheit** erhebt im Gegensatz zur Interessantheit des Themas die der Veranstaltung: Ist das Seminar/Vorlesung interessant oder zieht es sich schleppend dahin? Hier soll die Gestaltung der Veranstaltung gemessen werden, nicht das Thema.

Unter **Thema** sind zwei eher heterogene Inhalte zusammengefasst: Haben sich die Schüler schon vor Beginn des Kurses für das Thema interessiert und wird das Thema als relevant betrachtet? Thema stellt weitgehend eine „Biasva-

riable“ dar, d.h. nicht die Veranstaltung oder der Dozent werden bewertet.

Auch **Redundanz** erhebt zwei verschiedene Aspekte: Hatte der Schüler schon ein beträchtliches Vorwissen? Treten inhaltliche Überschneidungen mit anderen Veranstaltungen auf? Niedrige Werte bedeuten günstige Resultate!

**Überforderung** oder **Anforderungen** soll messen, ob die Teilnehmer stark in der Stoffschwierigkeit, der Stoffmenge und der Geschwindigkeit gefordert werden. Im Gegensatz zu anderen Skalen sind hier mittlere Werte (Rohskala um 4, Normskala um 100) sehr gut.

**Lernen-quantitativ** erfasst in der Selbsteinschätzung, ob die Schüler etwas in der Veranstaltung lernen können. In „quantitativ“ wird der Umfang betont.

**Lernen-qualitativ** erfasst qualitatives Lernen: Besteht nach Teilnahme ein grundlegendes Verständnis, wird Sinnvolles und Wichtiges gelernt? Die beiden Lerndimensionen fragen nach der Effektivität der Lehre.

Mit zwei Fragen wird Feed-back auf Schüler-Beiträge und **Betreuung** durch den Dozenten erhoben.

**Referate** beurteilt die Qualität der Schüler-Referate hinsichtlich Didaktik, Inhalte und Lerneffektivität.

**Fleiß** soll die Mitarbeit der Schüler erheben. Werden gestellte Aufgaben wie Literaturlesen oder Vorbereitung der Stunde gemacht? Zusätzlich wird der Arbeitsaufwand pro Woche in Stunden und Minuten direkt angesprochen (anschließend beides zu Minuten summiert).

**Interaktionsmanagement** beurteilt die Förderung und Moderation von Schüler-Interaktionen.

**Beteiligung** erhebt im Gegensatz zu Fleiß die Aktivität innerhalb einer Veranstaltung. Fühlt man sich bei eigener Beteiligung äußernsfähig?

Zwei Fragen beinhalten die **Interessenförderung:** Fördert der Veranstaltungsbesuch das Interesse an der Ausbildung und motiviert der Kurs dazu, sich selbst mit den Inhalten zu beschäftigen?

Mit **Allgemeineinschätzung** wird schließlich die allgemeine Bewertung einer Veranstaltung erhoben. Die Notengebung für die Veranstaltung spricht sie direkt an. Diese Dimension soll einen unspezifischen Eindruck von der Veranstaltung widerspiegeln. Die Note wird vor der Dimensionsbildung in eine umgepolte 7er-Skala transformiert.

Die Fragen nach Bildungsabschluss, Abschlusszeugnisnotenschnitt, Geschlecht, Besuchsgrund, Alter und Ausbildungsjahr dienen der Erhebung biografischer Hintergrundvariablen. Anhand eines Kodewortes lassen sich anonym Fragebögen aus verschiedenen Kursen einem Schüler zuordnen. Mit der prospektiven Frage nach dem Leistungsniveau lassen sich ähnlich wie mit dem Abschlusszeugnisnotenschnitt das Fähigkeitsniveau erheben.

Die Normierung (Mittelwert = 100, Streuung 1= 0; 66% aller Werte liegen zwischen 90 und 110) dient dazu, die eigenen Resultate mit den Werten anderer vergleichen zu können. Es sind so wesentlich einfacher relative Stärken und Schwächen feststellbar.

terdisziplinär werden wir einzelne Gesprächsverläufe in den Beratungen mit den Dozenten und die Diskussion der Evaluation in Institutskonferenzen schildern. Diese Darstellungen erlauben eine Einschätzung der Akzeptanz und der Auswirkungen des Evaluationsansatzes und stellen ein konkretes Beispiel dar, wie mit günstigen oder kritischen Evaluationsergebnissen produktiv umgegangen werden kann.

### L I T E R A T U R

Beywl, W. (1988). *Zur Weiterentwicklung der Evaluationsmethodologie: Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation*. Frankfurt: Peter Lang.

Buhl, Th. (1999). *HILVE II. Pilot-Studie mit Messwiederholungsdesign*. Jena: Friedrich-Schiller-Universität, Forschungsbericht aus dem Psychologischen Institut.

Bundesärztekammer (1996). *Curriculum Qualitätssicherung – Ärztliches Qualitätsmanagement*. Köln: Selbstverlag.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

dbl (1996a). Zertifikat „Lehrlogopädin“ nach den Richtlinien des DBL. *Forum Logopädie*, 3, 38 u. Einhefter VI.

dbl (1996b). Mindestanforderungen zur Gründung

und Qualitätssicherung von Lehranstalten für Logopädie. *Forum Logopädie*, 3, 35 u. Einhefter III-V.

Keuper, R., Mesenholl, H.J., Oppek, E. & Pfitzinger, E. (1995). *Qualität in der Beruflichen Bildung*. Stuttgart: Schriftenreihe des Landesarbeitskreises für berufliche Bildung, Landesgewerbeamt Baden Württemberg.

Kohler, J. & Rindermann, H. (1998). *Beratung von Dozenten zur Verbesserung der Lehrqualität - Erste Ergebnisse einer Effektivitätsprüfung*. Dresden: Poster auf dem Kongress der Deutschen Gesellschaft für Psychologie.

Lamnek, S. (1995a). *Qualitative Sozialforschung. Band 1 Methodologie*. Weinheim: Beltz Psychologie Verlags Union.

Lamnek, S. (1995b). *Qualitative Sozialforschung. Band 2 Methoden und Techniken*. Weinheim: Beltz Psychologie Verlags Union.

Marsh, H.W. & Roche, L.A. (1994). *The use of students' evaluations of university teaching to improve teaching effectiveness*. Sydney: Australian Government Publishing Service.

Rindermann, H. (1996). Verbesserung der Lehre durch Veranstaltungsevaluation? In E. Witruk & G. Friedrich (Hrsg.), *Pädagogische Psychologie im Streit um ein neues Selbstverständnis* (S. 147-156). Landau: Empirische Pädagogik.

Rindermann, H. (1997). Die studentische Beurteilung von Lehrveranstaltungen: Forschungsstand und Implikationen für den Einsatz von Lehrveranstaltungen. In R. S. Jäger, R. H. Lehmann & G. Trost (Hrsg.), *Tests und Trends (Jahrbuch der Pädagogischen Diagnostik, 11)* (S. 12-53). Weinheim: Beltz.

Rindermann, H. (1998). Das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität: Entwicklung, Begründung und Überprüfung. *Beiträge zur Hochschulforschung*, 3, 189-224.

Rindermann, H. (1999). Bedingungs- und Effektivitätsvariablen in der Lehrerevaluationsforschung: Konzeption und Prüfung des Münchner multifaktoriellen Modells der Lehrveranstaltungsqualität. *Unterrichtswissenschaft*, 27(4), 357-380.

Rindermann, H. & Amelang, M. (1994). *Das Heidelberger Inventar zur Lehrveranstaltungs-Evaluation HILVE*. Heidelberg: Asanger.

Ständige Konferenz der Logopädenlehranstalten und Fachtagung der Lehrlogopäden (1993). *Curriculum für die Ausbildung des Logopäden*. Berlin: Gross.

Wottawa, H. & Thierau, H. (1998). *Lehrbuch Evaluation*. Stuttgart: Hans Huber

**Autoren:**  
Jürgen Kohler  
Logopäde, Dipl.-Psychologe  
Schule für Logopädie  
Schloßstr. 1  
87479 Reichenau

**Dr. Heiner Rindermann**  
Dipl.-Psychologe  
Institut für Psychologie, Otto-von-Guericke-Universität Magdeburg  
Postfach 4120  
39016 Magdeburg  
Email:  
heiner.rindermann@gse-w.uni-magdeburg.de