

Methoden der Unterrichtsbewertung

F. Eitel, München

Im Rahmen der deutschen Reformdiskussion um die Qualität der Lehre hat das große Zählen und Messen begonnen (Hartmer 1993). Es stellt sich somit nicht die Frage, *ob* evaluiert werden soll, sondern *was* gemessen werden und *wie* bzw. *womit* gemessen werden soll, um die Lehrqualität hinreichend zuverlässig und gültig zu erfassen. Während diese Fragestellungen im angloamerikanischen Sprachraum intensiv bearbeitet werden (Walberg et al 1990), ergibt eine Erhebung an allen 38 deutschen Medizinfakultäten, daß an keiner Fakultät eine systematische, d.h. alle Lehrveranstaltungen methodisch adäquat erfassende Evaluation der Lehrqualität stattfindet. Auf die Umfrage haben 15 Fakultäten geantwortet: 13 Fakultäten führen eine Unterrichtsbewertung in mehr oder weniger methodischer Form durch, 2 Fakultäten evaluieren nicht. In den 13 evaluierenden Fakultäten werden von 10 Institutionen, Instituten oder Kliniken Erhebungen mittels Fragebogen durchgeführt. Die übrigen 5 Evaluationsaktivitäten sind weniger formalisiert als die genannten Fragebo-

genaktionen. Es handelt sich demnach um Einzelaktivitäten methodisch unterschiedlicher Reife.

Diese Erhebung zeigt, daß eine der Grundvoraussetzungen für eine Unterrichtsbewertung, nämlich die Datensammlung auf empirischer Grundlage (mittels Fragebogen, Interviews oder Beobachtung), nicht in systematischer Weise in den Fakultäten gegeben ist.

Ziel des vorliegenden Artikels ist es deshalb, Informationen zur Verfügung zu stellen, die den Leser und die Leserin in die Lage versetzen, die Evaluation des eigenen Unterrichts zu *planen*. Der vorliegende Artikel kann auch dazu dienen, die entsprechenden Organe der Fakultäten darüber in Kenntnis zu setzen, was veranlaßt werden könnte, um die Qualität der Lehre zu sichern. Eine in die Einzelheiten gehende Darstellung der Methoden (statistische Prozeßkontrolle [SPC] oder Fehlermöglichkeiten und -einflußanalyse [FMEA]) würden den Rahmen dieser Übersicht hier sprengen und ist für eine nachfol-

gende Veröffentlichung vorgesehen. Hier sollen zunächst 2 Fragen beantwortet werden: Was soll wie evaluiert werden?

Was soll evaluiert werden?

Bevor man sich um den Beobachtungsansatz bzw. Versuchsplan kümmert, ist es angezeigt, sich über die zu untersuchenden Beziehungen, Relationen, Ursache-Wirkungszusammenhänge klar zu werden. Man sollte über eine Vorstellung, genauer: ein Modell, dessen verfügen, was evaluiert werden soll. Man kann erprobtermaßen den Evaluationsansatz in sogenannten Ursache-Wirkungsdiagrammen darstellen (Abb. 1). Ein Beispiel für ein Modell des Unterrichtsablaufes findet sich in einer früheren Veröffentlichung (siehe Eitel et al. 1992, Seite 240).

Die nachfolgenden Erörterungen beziehen sich vorwiegend auf die Qualität der Unterrichtsdurchführung (Prozeßevaluation), nicht auf die Messung der Lernerfolge (Ergebnisevaluation), letzteres ist ein eigenes trauriges Kapitel der Geschichte deutscher Lehr-Evaluation. Da die Qualität der Unterrichtsdurchführung eine der wesentlichen Determinanten für den Unterrichtserfolg ist, beschäftigen wir uns im folgenden mit der **Prozeßevaluation**, mit anderen Worten: mit der Messung der Unterrichtsdurchführung auf statistischer Grundlage.

Es stellt sich zunächst die Frage, was Prozeßqualität ist. Als erstes gilt es, sich darüber klar zu werden, wie die "Qualität" des zu beobachtenden Unterrichtsprozesses definiert wird, was mit anderen Worten in der Unterrichtsdurchführung einen "Wert" darstellt, der gesichert oder gesteigert werden soll. Was also macht *gute* Lehre aus? Die Beantwortung dieser Frage setzt voraus, daß in dem zu untersuchenden Unterricht eine Beziehung besteht zwischen **strukturellen Variablen der Qualität** (z.B. der Lehrbefähigung, oder der Einstellung der Lehrer zur Lehre), und **Prozeßvariablen** (z.B. dem aktuellen Verhalten der Lehrer im Unterricht bzw. dem Verhalten der Lerner). Weiterhin ist zu überprüfen, ob eine Beziehung besteht zwischen den genannten Prozeßvariablen und **Ergebnisvariablen**, d.h. dem Lernerfolg der Schüler (vergl. Abb. 1).

Wenn also die genannten Strukturvariablen die Prozeßvariablen beeinflussen und diese wiederum die Ergebnisvariablen, dann kann man davon ausgehen, daß die Messung der zentralen Variablen, nämlich der Prozeßvariablen, ein hinreichend genaues Bild der Unterrichtsqualität ergibt. Die Ergebnisse der Messung der Unterrichtsdurchführung können dann als Begründung für Entscheidungen zur Organisationsänderung benutzt werden. Damit erfolgt die Verbesserung der Durchführungsqualität auf empirischer

Grundlage und nicht aufgrund von Einstellungen, Vermutungen, Vorurteilen und ungeprüften Traditionen wie weithin beobachtbar.

Ursache - Wirkungs - Diagramm (Qualitative Wirkungsanalyse)

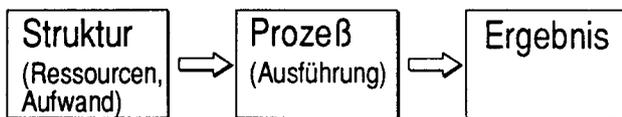


Abb. 1: Ursache-Wirkungs-Diagramm für Systeme der Gesundheitsversorgung. Dieses Schema kann zugleich dazu dienen, die verschiedenen Bereiche der Gesundheitsversorgung zu beschreiben. Die **Struktur** besteht in der materiellen und personellen Ausstattung, der Organisationsstruktur des Gesundheitswesens, dem finanziellen Aufwand sowie der Medizinischen Wissenschaft und der Aus-, Weiter- und Fortbildung des Personals (Qualifikation). Der **Prozeß** besteht in der Ausführung der Gesundheitsversorgung und das **Ergebnis** besteht in dem erreichten gesundheitlichen Status bzw. der Lebensqualität. Die Lehre selbst kann auch nach diesem Schema beschrieben werden. Die Struktur der Lehre besteht im Unterrichtsprogramm (Curriculum, Lehrplan), der Ausstattung in personeller und materieller Hinsicht und dem Aufwand; der Prozeß beinhaltet die Unterrichtsdurchführung, insbesondere das Lehr- und Lernverhalten der Beteiligten; und das Ergebnis schließlich besteht im Lernerfolg (Kompetenz und Performanz).

Metaanalysen im Schrifttum ergeben, daß als Indikatoren für didaktische gute Unterrichtsdurchführung gelten:

- die einfühlsame Zuwendung des Dozenten¹ gegenüber den ihm anvertrauten Studenten¹ (Empathie),

- ein strukturierter, d.h. geordneter Unterrichtsablauf und dem Stand der Wissenschaft entsprechender Unterrichtsinhalt (Didaktik im engeren Sinne),
- ein aktives Engagement der Schüler für die Bewältigung der gestellten Aufgaben (Lernmotivation)
- und eine größtmögliche Rückmeldung (Feletti et al. 1982, Galton 1987) der erreichten Kompetenz bzw. Performanz an die Lerner (Feedback).
- Weiterhin ist die aktive Teilnahme der Lerner (Stritter et al. 1975, Jason 1962) am Unterrichtsgeschehen von entscheidender Bedeutung für die Qualität des Unterrichts.

Die genannten Zielkriterien haben wir in Fragen (Items) gefaßt, die in einen Fragebogen zur Prozeßevaluation aufgenommen wurden. Weitere Items zur Feststellung der Akzeptanz und des subjektiv eingeschätzten Lernerfolges sind darüberhinaus im Fragebogen enthalten (Abb. 2).

¹ Alle maskulinen Bezeichnungen betreffen in gleicher Weise Männer und Frauen

Fragebogen zum Praktikum der Chirurgie

<p>1. Lernziel der heutigen Kursstunde (was sollen die Studierenden am Ende der Kursstunde tun können? - Dozenten fragen!) Wie weit bin ich <u>zu Beginn</u> vom obigen Lernziel entfernt?</p> <p>In der heutigen Unterrichtsveranstaltung</p> <p>2. besaß ich bereits Vorwissen</p> <p>3. bestand für mich die Möglichkeit selbst tätig zu werden</p> <p>4. war der Kursaufbau strukturiert/chaotisch</p> <p>5. ging der/die Dozent/in auf die Teilnehmer/innen ein</p> <p>6. hat sich der/die Dozent/in bemüht</p> <p>7. habe ich verstanden worum es ging</p> <p>8. bin ich gefordert worden</p> <p>9. waren Rückfragen möglich</p> <p>10. hat der/die Dozent/in mir gesagt was ich kann</p> <p>11. war ich konzentriert/abgelenkt</p> <p>Nach der heutigen Unterrichtsveranstaltung</p> <p>12. möchte ich mich weiter mit dem Thema beschäftigen</p> <p>13. erteile ich der Veranstaltungsqualität insgesamt die Note(1 ... 6)</p> <p>14. Wie weit bin ich <u>am Ende</u> der Veranstaltung vom obigen Lernziel entfernt?</p> <p>15. Die heutige Kursstunde hat mir sehr gut gefallen/gar nicht gefallen</p>

Abb. 2: Fragen (Items) zur Evaluation der Durchführungsqualität der Veranstaltung "Praktikum der Chirurgie". Zielkriterien sind: subjektiv eingeschätztes Lernziel, Unterrichtsorganisation, didaktische Kriterien wie aktive Teilnahme am Unterricht oder Rückmeldung des Lernerfolges durch Dozenten, die Lernmotivation (z.B. Weiterbeschäftigung mit dem Thema), subjektive Einschätzung der Lehrqualität und Akzeptanz seitens der Studenten. Die Fragen werden auf einer sechsstufigen Skala, die hier nicht abgebildet ist, von den Studenten eingeschätzt. Auch die Möglichkeit, eigene Kommentare abzugeben, ist in einer freien Frage gewährleistet.

Es ist klar, daß die genannten Zielgrößen (repräsentiert durch Items) *nur eine Seite* der Medaille "gute Lehre" darstellen. Sie sind hier herausgehoben worden, da sie wesentliche Determinanten der Qualität sind, und weil hier in der momentanen Situation vieles im Argen liegt. Natürlich müßte die Unterrichtsevaluation auch andere Zielgrößen erfassen (fachlich medizinisch-wissenschaftliche Gesichtspunkte, Anforderungen des Berufsfeldes oder andere qualitative Aspekte wie Patientenzufriedenheit oder Fähigkeit der Lernenden zur adäquaten Informationsverarbeitung als Ausbildungsziele und damit als Evaluationsgegenstände).

Wie soll evaluiert werden?

In der Lehrevaluation ist ein experimentelles Design aus praktischen Gründen kaum durchführbar. Deshalb ist es üblich, nach einem *quasi-experimentellen Versuchsplan* vorzugehen (vergl. Cook et al. 1979), d.h. ohne Randomisation, ohne Blindbedingungen und gelegentlich auch ohne Vergleichsgruppe, stattdessen mit Vergleich in sich (z.B. Panel-Design oder Zeitreihenanalyse; bezüglich Einzelheiten vergl. Rossi et al. 1988, Eitel et al. 1993a und 1994).

Es werden im allgemeinen Kohortenstudien durchgeführt, wobei möglichst alle Unterrichtsteilnehmer erfaßt werden sollen (Totalerfas-

sung). Es handelt sich also meist nicht um den Vergleich zweier unverbundener Stichproben, sondern vielmehr um den Vergleich von Meßdaten, die vor, während und nach dem *Wirksamwerden* einer Einflußgröße (Intervention) - oder noch kontinuierlicher mit mehr Meßzeitpunkten (Panel-Design oder Zeitreihenanalyse, siehe Rossi et al. 1988) erhoben werden. Beide Versuchspläne werden unter dem Begriff *Time-series-design* zusammengefaßt. Dieses Design ist dazu geeignet, die *Wirkung* von Maßnahmen der Unterrichtsorganisation zu überprüfen, aber auch um beispielsweise ein Lernergebnis vor, während und nach der Instruktion zu messen. Letzteres ist aber nicht Gegenstand des vorliegenden Artikels. Die *Wirkungsanalyse* mit dem Panel-Design bzw. der Zeitreihenanalyse ist relativ kostenaufwendig. Um bei möglicherweise geringen Ressourcen oder geringer Evaluationserfahrung einen empirischen Datensatz zu gewinnen, der gewissermaßen als historische Kontrolle für zukünftige, weitere Untersuchungen dienen kann, genügt es am Anfang einer Evaluationsserie im allgemeinen, im sogenannten Post-Test-Design am Ende einer Maßnahme die relevanten Zielkriterien (siehe Abb. 2) zu untersuchen.

Letztgenanntes Design alleine ist aber für Wirkungsanalysen ungeeignet, wenn es nicht als

erster Schritt für ein *Time-series-design* dient. Wendet man allerdings Maßnahmen an, deren Wirkung empirisch belegt ist, dann würde eine gleichzeitige Wirkungsanalyse redundante Ergebnisse liefern. Sie kann dann aus Ökonomiegründen unterbleiben und man überprüft im Post-Test-Design lediglich, ob die in den jeweiligen Kontext eingeführte Maßnahme das beabsichtigte Ziel erreicht hat. Die Qualitätserfassung besteht hier also in der sogenannten *Zielerreichungsanalyse*. (Beobachtungsstudie mit Vergleich der gemessenen Ist-Werte mit einem Sollwert). Bei der Zielerreichungsanalyse muß ein Soll-Wert festgelegt oder aus empirischen Daten abgeleitet werden, welcher dann mit den empirisch erhobenen Ist-Werten verglichen wird.

Die *Datensammlung* bei den Qualitätserfassungen in der Lehre erfolgt im allgemeinen durch informelle oder mehr oder weniger strukturierte Interviews, durch direkte oder indirekte Beobachtungen und durch Erhebungen mit Fragebögen. Fragebögen zur Prozeßevaluation (siehe Abb. 2) können zu Beginn des Semesters jeweils für jede Veranstaltung an jeden Studenten ausgegeben werden. Der Student füllt den Fragebogen für jede von ihm besuchte Veranstaltung aus. Der Fragebogen dient gleichzeitig zur Bestätigung der Teilnahme an der jeweiligen Unterrichtsveranstaltung. Am Ende vom Seme-

ster werden die Fragebögen wieder eingesammelt.

Wir verwenden eine 6-stufige Skala zur Beurteilung der Fragen in Anlehnung an die allseits bekannte Schulnotenskala. Statistische Gründe sprechen für die Anwendung dieser Skala.

Die von jedem Studenten zu der von ihm besuchten Veranstaltungen für jedes Item abgegebenen Noten i werden auf einem PC mit dem Tabellenkalkulationsprogramm EXEL[®] datenbankmäßig in einer Rohwerttabelle (vergl. Abb. 7) erfaßt.

Datenreduktion: Die Einschätzungen der n -Studenten werden für jeden der p -Veranstaltungsorte mit dem genannten Programm durch Mittelwertbildung m mit Standardabweichung SD zur Beurteilung der Streuung bzw. mit Median zur Beurteilung der Verteilungsform so reduziert, daß für jedes Item i eine Durchschnittsnote pro Lernplatz bzw. Veranstaltungsort o zur Verfügung steht. Die Datenaufbereitung erfolgt für nominal und ordinal skalierte Variable in Histogrammen (Abb. 3), oder für ordinal skalierte Variable in Pareto-Histogrammen (Abb. 4). Speziell für die Wirkungsanalyse bei intervallskalierten Variablen bzw. bei Nutzenanalysen bieten sich Streudiagramme an (Abb. 5). Für die Zielerreichungsanalyse sind Kontrolldiagramme geeignet (Abb. 6). Diese Darstellungsweisen stam-

men aus der statistischen Prozeßkontrolle, wie sie in der Industrie üblich ist.

Die *Datenanalyse* erstreckt sich auf die inferenzstatistische Analyse der parametrisch (siehe vorherigen Absatz) aufbereiteten Daten. Der Zielerreichungsgrad für jeden Veranstaltungsort bzw. -zeitpunkt wird bestimmt, indem der errechnete Mittelwert m vom Sollwert s subtrahiert und die erhaltene Differenz durch den Sollwert s dividiert wird. Ist die Effektivität der untersuchten Maßnahme gewährleistet, so stellt der Zielerreichungsgrad den Nutzen der Unterrichtsmaßnahme dar. Es ist aber in jedem Fall Voraussetzung, eine Wirkungsanalyse oder eine Meta-Analyse (siehe Rossi et al. 1988, Fricke et al. 1985) im Schrifttum zur Feststellung des Wirkungsgrades einer Unterrichtsmaßnahme vorgenommen zu haben. Falls eine Wirkungsanalyse (vergl. Abb. 3-5) sowie eine Zielerreichungsanalyse (vergl. Abb. 6) vorgenommen wurde - (dies wird im allgemeinen bei innovativen Unterrichtsprogrammen, d.h. bei durchgreifenden Organisationsänderungen, zu tun sein), - können der *Wirkungsgrad* w und der *Zielerreichungsgrad* z in einem Streudiagramm korreliert werden (vergl. Abb. 5). Dann stellt der Korrelationskoeffizient $r_{w,z}$ das Maß für den Nutzen dar, d.h. je größer r , desto größer der Nutzen.

Die Erfassung des Nutzens ist aber für die Prozeßevaluation einer Unterrichtsmaßnahme alleine nicht hinreichend, da der Nutzen ja auch von dem Aufwand bzw. den entstehenden Kosten abhängt.

Histogramm
(Panel-Design, Zeitreihenanalyse)

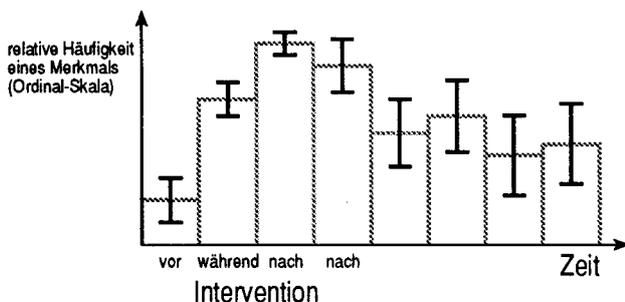


Abb. 3: Histogramm zur Darstellung der Wirkung einer Intervention (Maßnahme der curricularen Organisation, z.B. Einführung des problemorientierten Lernens, oder Vermittlung fachlicher Fertigkeiten, z.B. Nähen am Schaumgummimodell, vergl. Eitel 1992). Auf der Ordinate sind die aus den Zielkriterien abgeleiteten Prozeßgrößen oder Ergebnisindikatoren abzutragen. Die Wirkung wird dann aus dem Vergleich in sich (Pretest-Posttest-Design) z.B. der Werte vor und nach Instruktion ermittelt, sei es mittels nonparametrischer Signifikanzprüfung bei ordinal skalierten Variablen oder nicht normalverteilten Daten; oder Korrelationskoeffizient bei intervallskalierten Variablen.

Pareto-Diagramm
(Quantitative Wirkungsanalyse)

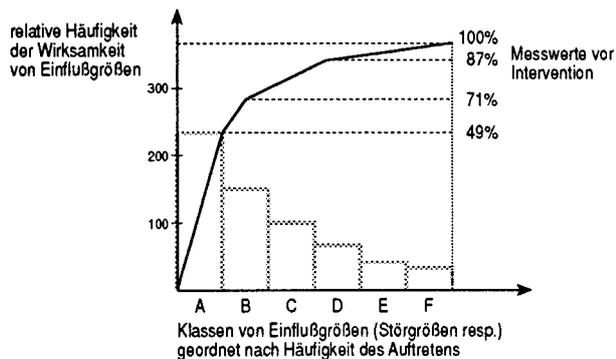


Abb. 4a: Auf der Abzisse werden die Einflußgrößen A-F, die bei einer qualitativen Wirkungsanalyse festgestellt wurden (siehe Abb. 1), nach der Häufigkeit ihres Auftre-

tens geordnet abgetragen. Auf der Ordinate wird die z.B. in Strichlisten (vergl. Abb. 7) festgehaltene absolute oder relative Häufigkeit des Auftretens bestimmt, die Höhe der jeweiligen Säule entspricht der beobachteten Merkmals-häufigkeit. Rechts wird die kumulative Häufigkeit der Merkmale aufgetragen.

Pareto-Diagramm
(Quantitative Wirkungsanalyse)

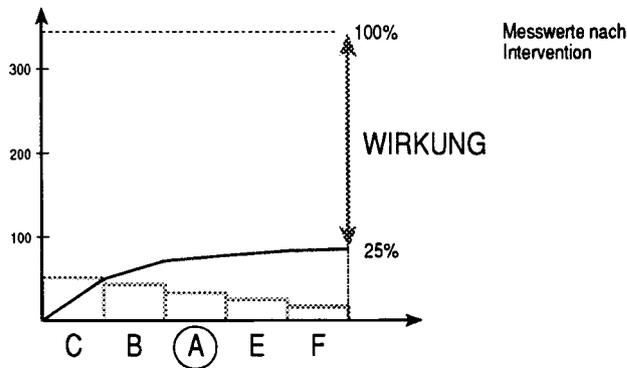


Abb. 4b: Die Intervention führte zur Reduktion der Merkmalshäufigkeit, die zuvor häufigste Störgröße hat jetzt einen mittleren Platz in der Rangfolge. Die Differenz zwischen dem vor Intervention gemessenen Ausgangswert (Pretest) und dem nach Intervention gemessenen Wert (Posttest) stellt die Wirkung dar, hier beträgt sie 75 %.

Streudiagramm
(Quantitative Nutzenanalyse)

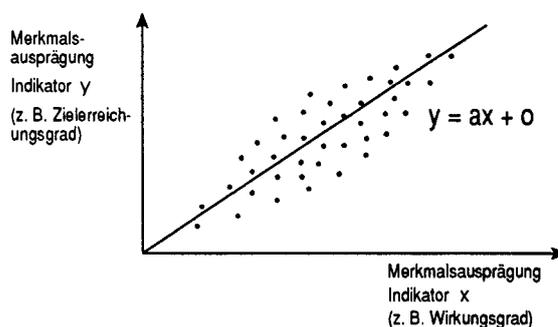


Abb. 5: Streudiagramme bilden die lineare Korrelation zweier kontinuierlicher Variablen ab. Die Steigung a ergibt die Stärke des Zusammenhanges: $a_r > 0,8$ = starker Zusammenhang (Kausalität), $a_r > 0,2$ = schwacher Zusammenhang. r = Pearson Product Moment Correlation Coefficient (Voraussetzung der Anwendung: Normalverteilung).

Kontrolldiagramm (Zielerreichungsanalyse)

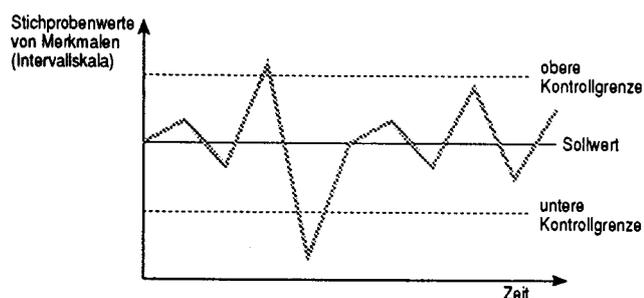


Abb. 6: Zielerreichung liegt vor, wenn die zufällig um den Sollwert schwankenden Meßwerte die obere oder untere Kontrollgrenze nicht überschreiten, die natürliche Variabilität also in Grenzen bleibt. Der Verlauf der Kurve in der 3. und 4. Zacke zeigt an, daß eine nähere Überprüfung des Prozesses erforderlich ist. Auch systematische Fehler können aus einer konstanten Abweichung von Sollwert bzw. konstant hohen Varianzen der Meßwerte erschlossen werden. Bei diesem Diagramm erfolgt ein Vergleich von Meßzeitpunkten; Evaluation besteht immer im Vergleich mindestens zweier Datensätze.

Matrizes (Tabellen)

z. B. Auftretenshäufigkeit von Merkmalen				
z. B. Meßzeitpunkte				

Abb. 7: Beispiel für die Möglichkeit der Gestaltung einer Rohwerttabelle. Sie stellt die Matrix für eine Auszählung von Merkmalen dar (z.B. Strichliste).

Es gilt also, eine *Kosten-Nutzenanalyse* durchzuführen. Hierzu werden die Kosten für die jeweilige Unterrichtsmaßnahme erfaßt und mit dem ermittelten Grad des Nutzens (Zielerreichungsgrad oder Nutzen-Korrelationskoeffizient) wiederum in Beziehung gesetzt bzw. mittels eines Streudiagrammes auf die Größe des Korrelationskoeffizienten r (Pearson-Product-Moment-

Correlation-Coefficient) hin untersucht. Voraussetzung hierfür ist, daß die erfaßten Variablen intervallskaliert, kontinuierlich und normalverteilt sind; ist dies nicht der Fall, muß mit den entsprechenden statistischen Methoden (z.B. Chi-Quadrat-Test oder r_s) vorgegangen werden. Der dadurch ermittelte Korrelationskoeffizient ergibt die *Effizienz q* einer Unterrichtsmaßnahme.

Wir definieren Prozeßevaluation als die *Bestimmung der Effizienz* einer Unterrichtsmaßnahme, was letztendlich auf eine Kostennutzenanalyse hinausläuft. Die Effizienz kann in einem Funktionsdiagramm dargestellt werden (Abb. 8). Hierbei wird die Korrelation zwischen Aufwand und Nutzen im allgemeinen nicht linear sein.

Die Qualität der Unterrichtsdurchführung kann, - um es auf eine Formel zu bringen -, mit dem Parameter der Effizienz q abgebildet werden.

Relevanz der Evaluation

Das Funktionsdiagramm der Kosten-Nutzen-Analyse (Abb. 8) kann zur Beantwortung einer wichtigen Frage herangezogen werden: Wann soll man qualitätsverbessernde Maßnahmen (continuous quality improvement, Qualitätsmanagement) aufgeben zugunsten einer Innovation im Sinne der grundlegenden Erneuerung? Qualitätsverbesserungen der Lehre wären zum Beispiel organisatorische Änderungen mit dem Ziel

der gleichmäßigeren Verteilung des Lehrdeputates ("Lehrdienstplan", siehe Eitel et al. 1993b, Seite 252); eine Innovation wäre zum Beispiel eine Studienreform wie sie der Entwurf zur 8. Novelle der Ärztlichen Approbationsordnung vorsieht. Es ist klar, daß letzteres wesentlich aufwendiger ist als eine kontinuierliche Qualitätsverbesserung.

Soll also weiter nur verbessert oder muß innoviert werden? Es ist vernünftig, dann zu innovieren, wenn bei der laufenden Qualitätsverbesserung der Aufwand immer größer und die damit einhergehende Zunahme des Nutzens immer kleiner wird, wenn also mit anderen Worten der break-even-point (Kostendeckungspunkt) erreicht wird (siehe gestrichelte Linie in Abb. 8). Denn bei Überschreitung des break-even-point "schreibt man rote Zahlen".

Die Evaluation ist - wie dargestellt - in der Lage, festzustellen, ob in der Lehre "rote oder schwarze Zahlen" geschrieben werden. Evaluation ist zwar derzeit in Deutschland nicht so weit entwickelt, daß dies möglich wäre. Im Rahmen der Strukturveränderungen im Gesundheitswesen wird aber die Ökonomie auch der Aus-, Weiter- und Fortbildung Diskussionsgegenstand werden. Dann wird Evaluation als Methode der Effizienzbestimmung eine gefragte Fertigkeit werden.

Funktionsdiagramm (Aufwand - Nutzen - Analyse)

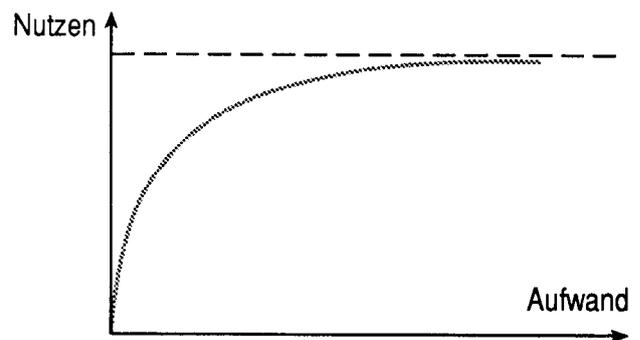


Abb. 8: Bestimmung der Effizienz einer Intervention durch Aufwand(Kosten)-Nutzen-Analyse. Hier wird also der Ressourcenverbrauch eines Prozesses mit dem Prozeßergebnis (siehe Abb. 1) verglichen. Evaluation ist also Effizienzmessung. Bei angemessen zunehmendem Nutzwert (aufsteigender Ast der Kurve) kann man sich mit einer Strukturverbesserung (Qualitätssicherung, Erhaltung des Status quo) begnügen, was billiger ist als eine Innovation.

Literatur

1. Cook TD, Campbell Dt (1979) Quasi-experimentation- Design and analysis issues for field settings
Chigago: Rand Mc-Nally
2. Eitel F (1992)
Zur problemorientierten studentischen Ausbildung und deren Evaluation
In: L. Schweiberer, J.R. Izbicki (Hrsg.)
Akademische Chirurgie, S.235-250
Springer, Berlin
3. Eitel F, Prenzel M, Schweiberer L, Lyon HC (1993a)
Quality assurance of education in surgery
I. Approach to improving its quality
Theor Surg 8:194-202
4. Eitel F, Kanz KG, Seibold R, Sklarek J, Feuchtgruber G, Steiner B, Neumann A, Schweiberer L, Holzbach R, Prenzel M (1993b)
Verbesserung des Studentenunterrichts - Sicherung der Strukturqualität Medizinischer Versorgung

In: D. Habeck, U. Schagen, G. Wagner (Hrsg.) Reform der Ärzteausbildung - Neue Wege in den Fakultäten
Blackwell Berlin, S.243-266
5. Eitel F, Lackner CH (1994)
Qualitätsmanagement als Kostenvermeidungsstrategie
ALERT 1:58-62

6. Eitel F, Prenzel M, Schweiberer L, Lyon HC (1994)
Quality assurance of education in surgery
II. Evaluation approach assessed by meta-evaluation
Theor Surg 9:1-9
7. Feletti GJ, Doyle E, Petrovic A, Sanson-Fisher R
(1982) Medical students' evaluation of tutors in a
group-learning curriculum
Med Educ 16:319-325
8. Fricke R, Treinies G (1963)
Einführung in die Metaanalyse
Huber, Bern 1925
9. Galton M (1987)
Structured Observation
In: Dunkin MJ (ed.)
The International Encyclopedia of Teaching and Teacher Education, S.142-147
Pergamon Press, Oxford - New York
10. Hartmer M (1993)
Das große Zählen und Messen
Süddeutsche Zeitung (11.10.93) 235:47
11. Jason H (1962)
A Study of Medical Teaching Practices
J Med Educ 37:1258-1284
12. Rossi PH, Freeman HE, Hofmann G (1988)
Programm-Evaluation
Einführung in die Methoden angewandter Sozialforschung
Enke, Stuttgart
13. Stritter FT, Hain JD, Grimes AD (1975)
Clinical Teaching Reexamined
J Med Educ 50:876-882
14. Walberg HJ, Haertel GD (1990)
The International Encyclopedia of Educational Evaluation
Pergamon Press, Oxford - New York

Prof. Dr. med. Florian Eitel
Chirurgische Klinik und Chirurgische
Poliklinik, Klinikum Innenstadt
Ludwig-Maximilians-Universität
Nußbaumstraße 20
80336 München