

STUDIENREFORM UND DIE STRUKTURIERUNG DER ERGEBNISEVALUATION — GRUNDSÄTZLICHE ÜBERLEGUNGEN ZUR PRÜFUNGSREFORM IM ZUSAMMENHANG VON STUDIENREFORM

INGRID STIEGLER, HAAN

Zusammenfassung:

Vorausgesetzt wird, daß Prüfungen mit Auslesefunktion eine starke Rückwirkung auf das Curriculum haben, dieses also in Inhalten, Zielen, den Lehr- und Lernstilen usw. wesentlich bestimmen. Das führt zu der paradoxen Situation, daß Prüfungen als Instrumente der Ergebnisevaluation von Ausbildung(1) in der Regel nur ihre eigene Ergebnisevaluation sind. Curriculare Reformen in der medizinischen Ausbildung könnten scheitern, wenn Prüfungen nicht mitreformiert werden und Prüfungslernziele und die ihnen angemessenen -formen curriculare Ziele nicht repräsentieren. Um als Instrumente der Ergebnisevaluation von Studienreform zu fungieren, sollten sie bestimmte allgemeine Kriterien erfüllen.

Summary:

Due to the strong influence of examinations on the curriculum they determine effectually the contents, aims, teaching- and learning-methods. This leads to the paradox consequence that examinations, originally estimated evaluation of educational results¹, tend to become only their own evaluation. Curriculum reforms in the medical education could fail, if examinations are not reformed in the same way and if examinations, aims and methods do not represent curricula aims. To work as instruments for evaluation of results of educational reforms the examination should fulfill certain general criteria.

I. Voraussetzungen

1971 schreiben Pierre Bourdieu und Jean-Claude Passeron im Zusammenhang soziologischer Studien zum französischen Bildungssystem u.a. folgendes zum Thema Examen: "Es ist nur allzu evident, daß das Bildungswesen, zumindest heute ..., nicht nur in den Vorstellungen, sondern auch im realen Verhalten aller Beteiligten vom Examen beherrscht wird. Man bemerkt häufig die sichtbarsten Wirkungen, die der Primat der Auslesefunktion auf Kosten aller anderen pädagogischen Funktionen und vor allem der Bildungsfunktion hervorbringt. Man hat oft die Furcht vor den totalen, brutalen und zum Teil unberechenbaren Urteilen traditioneller Prüfungen beschrieben, die Funktionsstörungen, die daraus entstehen, daß ein einfaches Kontrollmittel zum alleinigen Prinzip der Themen- und Methodenwahl erhoben wird, sowie der Rythmusstörungen, die aus einer Arbeitsorganisation herrühren, die in ihren regellosesten Formen nur noch dem Terminiindruck des Examens gehorcht. Das Examenssystem ist nicht nur der offenbarste Ausdruck der Werte und impliziten Entscheidungen des Bildungssystems: In dem Maße, wie es die Aufmerksamkeit von Lehrenden und Studierenden bindet, ist es eine wirksame Regulierungsinstanz, die dadurch prägt, daß sie die Arbeit orientiert und sowohl Versagen wie auch Erfolge sanktioniert. Die Ausbildung wird mindestens ebenso stark wie von den Studienprogrammen von dem Gewohnheitsrecht beeinflusst, das in der Rechtsprechung der Examina besteht." (2)

Im Sinne Bourdieus/Passerons haben Examina die wesentliche Funktion, die gute Leistung gegenüber der schlechteren zu differenzieren und Leistungen zu plazieren, Kompetenz von Inkompetenz zu scheiden, also eine leistungsgerechte Auslese vorzunehmen und schließlich die Kompetenz zu zertifizieren und damit Berechtigungen für Studien- oder berufliche Anschlüsse zu verleihen. Paradox ist, daß die Auslese- und Berechtigungsfunktion den wesentlichen Sinn von Bildungsinstitutionen - nämlich Kompetenz auszubilden - dominiert und ihn partiell boykottiert.

Das Bildungssystem selbst rechtfertigt Prüfungen mit einer Verantwortung gegenüber der Gesellschaft, nur geeignete Personen in weiterführende Bildungswege oder berufliche Karrieren zu entlassen. Über Examina behauptet es selbst zudem seine relative Autonomie gegenüber dem Beschäftigungssystem. Vorausgesetzt wird, daß Prüfungen einen hohen prognostischen Wert haben, ihr Qualifikationsnachweis nicht nur bildungssystemimmanent, sondern zugleich für außeruniversitäres berufliches Handeln gilt. Qualifikation ist im schulisch erworbenen Titel bzw. in der Berechtigung gewonnen und muß nicht Tag für Tag erneut unter Beweis gestellt werden.(3)

Prüfungen mit Selektions- und Berechtigungsfunktion sind in modernen demokratischen Gesellschaften aber vor allem aus folgendem Grund unersetzbar: Auf der Basis von Leistungen des einzelnen im Bildungssystem, die durch Examina gemessen werden, wird soziale Differenzierung legitimiert. Nicht direkt vererbtes ökonomisches oder soziales Kapital rechtfertigt die "feinen Unterschiede"(4), sondern das Bestehen oder Nicht-Bestehen von systemhaft organisierten, standardisierten und "objektivierten" Prüfungen steuert die unterschiedlichen Sozialkarrieren scheinbar gerecht. Wie wichtig diese Rechtfertigungsfunktion von Prüfungen für den sozialen Status des einzelnen und die differenzierte Sozialstruktur in einer modernen demokratischen Gesellschaft ist, erahnt man in Zeiten einer krisenhaft erlebten Titelinflationierung und Überfüllung, wie sie augenblicklich in der Medizin gegeben ist. Wenn die Selektivität des Bildungssystems über das Examenssystem oder eine vorhergehende Selbstelimination von potentiellen Studienfachbewerbern versagt, wird die Auslese aus dem Bildungs- in das Beschäftigungssystem verlagert. Dort werden dann zunehmend Kriterien der Auswahl zwischen gleichermaßen Berechtigten relevant, die langfristig in modernen demokratischen Gesellschaften nicht legitimationsfähig sind.

Die Selektions- und die Rechtfertigungsfunktion von Prüfungen für soziale Differenzierung haben ihre Ver-

rechtlichung nach dem Grundsatz der formalen Gleichbehandlung zur Folge. Sie setzt den Gestaltungsmöglichkeiten von Prüfungen durch die einzelne Hochschule bzw. Fakultät enge Grenzen und kann letztlich jede weitgehende Studienreform verhindern, wenn man folgendes bedenkt:

1. Examina schreiben das Curriculum in vielfacher Hinsicht mit: Ziele, Inhalte, Methoden, Lehrstile, Lernumgebung und die Organisation der Ausbildung werden rückwirkend darüber beeinflusst und bestimmt. Sie beurteilen an den Ergebnissen einer Ausbildung insofern das, was sie selbst als deren Inhalt und Form definieren. Sie sind - im Extremfall des bestehenden Prüfungssystems in der medizinischen Ausbildung - ihre eigene Ergebnisevaluation.

2. Examina sozialisieren die Studierenden in vielfältiger Weise. Nicht nur Lernstile und Lernverhalten werden von ihnen kollektiv geprägt und nivelliert. Sie bilden den "heimlichen Lehrplan", in dem man z.B. lernt, daß sich Versagungen, Opfer und Selbst-Disziplin auszahlen oder daß der selbstsichere, mit gekonnter Rhetorik verknüpfte Habitus in mancher mündlichen Prüfung siegt.⁽⁵⁾ Deren diagnostische Relevanz für Führungsfähigkeiten in Chefarztpositionen ist deshalb möglicherweise sehr viel höher, als der prognostische Wert der bestehenden MC-Prüfungen für kompetentes ärztliches Handeln.

Obwohl die Unangemessenheit bestimmter Prüfungsformen und -lernziele für bestimmte Studiengestaltungen und -ziele naheliegt und sie sich als Instrumente der Ergebnisevaluation von reformierter Ausbildung nicht eignen, und obwohl die unerwünschten Sozialisationseffekte der MC-Prüfungen in ihrer gegenwärtigen Gebrauchsform für das spätere ärztliche Handeln bekannt sind, steht die Verrechtlichung einer grundlegenden Prüfungs- und damit wirklich wirksamen Studienreform und deren Ergebnisevaluation in der Medizin im Wege.

II. Vorschläge

Dennoch sind begrenzt Wege der begleitenden Prüfungsreform als notwendiger Bedingung der Möglichkeit und der Qualitätskontrolle einer curricularen Reform denkbar, ohne das Kriterium der formalen Gleichbehandlung zu verletzen.

Dabei hätten die Fakultäten folgendes zu berücksichtigen:

1. Prüfungen haben - neben ihrem primären Gebrauchswert der Verleihung von Berechtigungen für Anschlüsse im Studium und im Beschäftigungssystem - die pädagogisch-didaktische Funktion, Lehrenden und Lernenden eine studienbegleitende Standortbestimmung und eine kontinuierliche Rückmeldung über ihre Leistungen zu geben. Curriculare Veränderungen an einer Fakultät und Entwicklungen des persönlichen Lehr- oder Lernstils können mit sogenannten formativen Prüfungen reflexiv begleitet werden.

Formative Prüfungen, an die keine Berechtigungen für einen Studienfortschritt gebunden sein dürfen, sollten - gerade im Rahmen einer curricularen Reform - kontinuierlich und ohne offizielle Vorbereitungszeit für die Studierenden eingesetzt werden. Nur so evaluieren sie Veränderungen im Curriculum für die Fakultät und den Lernfortschritt des einzelnen im Zusammenhang der

Ausbildung und nicht den Fleiß des Autodidakten. Dies müßte im übrigen auch für berechtigende Prüfungen, die am Ende von bestimmten curricular definierten Lernsequenzen "summativ" eingesetzt werden, gelten - zumindest dann, wenn sie zugleich als Instrumente der Ergebnisevaluation einer Studienreform fungieren sollen.

Ein erster positiver Nebeneffekt formativer Prüfungen - unabhängig vom Profit der Fakultäten für curriculare Qualitätssicherung - ist, daß Studierende lernen, Prüfungen als sinnvoll und hilfreich und Prüfungssituationen als selbstverständlich und nicht als angstbesetzte Ausnahmezustände oder Initiationsrituale zu erleben, deren Sozialisationseffekt auf zukünftige Ärzte nicht zu unterschätzen ist.

Ein weiterer positiver Nebeneffekt häufiger formativer Prüfungen könnte sein, daß Studierende, die immer wieder negative Rückmeldungen durch formative Prüfungen bekommen, sich selbst aus Einsicht eliminieren und nicht durch eine Institution, deren prognostische Validität sie bestreiten, eliminiert werden - oder auch nicht.

Ein dritter positiver Nebeneffekt ist, daß von vielfältig gestalteten formativen Prüfungen ein kreativer Schub auf die Gestaltung einer unerläßlichen, bundeseinheitlichen Abschlußprüfung mit beruflicher Berechtigungsfunktion, aber auch auf Fakultätsprüfungen, von denen ein Studienfortschritt abhängig gemacht wird, ausgehen kann.

2. Auch deshalb sollten formativ eingesetzte Prüfungen folgende Gütekriterien, denen diagnostische Tests genügen müssen, annähernd erfüllen:

- Sie sollten in Bezug auf die Vorbereitung, die Durchführung und die Auswertung objektiv sein. Nur auf der Basis der Objektivität können die Zusatzkriterien der Reliabilität und Validität erfüllt werden.

- Reliabilität heißt, daß eine Prüfung das zuverlässig diagnostiziert, was sie diagnostizieren will.

- Validität heißt, daß eine Prüfung das, was sie zu messen vorgibt, tatsächlich mißt.

3. Gerade im Zusammenhang von Reformcurricula und ihrer Evaluation über formative Prüfungen werden Fragen vor allem der Inhalts- und der prognostischen Validität wichtig. Reformcurricula bauen in der Regel auf neue, komplexe und praxisrelevante Lernziele, die über das Ziel des Erwerbs von Wissen im Sinne der Fähigkeit der Wiedergabe von Fakten, Begriffen, Gesetzen usw. weit hinausgehen. Ihnen geht es im Bereich der Ausbildung kognitiver Fähigkeiten der Studierenden auch um die Dimensionen des Verstehens, Anwendens, Analysierens, Synthetisierens und um die Fähigkeit der Bewertung und Beurteilung des Wissens.⁽⁶⁾ Zum Ziel der Ausbildung verschiedener Dimensionen kognitiver Fähigkeiten tritt das Ziel der Förderung praktischer Fähigkeiten, die selbst wiederum mehrdimensional zu beschreiben und mit kognitiven untrennbar verknüpft sind. Daneben wird der Erwerb von Schlüsselqualifikationen wie Lernfähigkeit, Teamfähigkeit usw. wichtig. An Prüfungslernziele, die sie repräsentierenden Aufgaben und die ihnen entsprechenden Formen wird im Zusammenhang der Ergebnisevaluation von Reformcurricula jedenfalls ein hoher Anspruch gestellt werden müssen. Eine Monokultur von MC-

Prüfungen, die sich auf die Überprüfung der Wiedergabefähigkeit von detailliertem Wissen beschränkt, scheint komplexen und vielfältigen Ausbildungszielen jedenfalls nicht angemessen zu sein.

4. Ein letzter Punkt, der bei der Gestaltung von Prüfungen als Instrumente der Ergebnisevaluation von Reformcurricula zu berücksichtigen ist, ist ihre Praktikabilität, d.h. der zeitliche und organisatorische Aufwand, der für die Vorbereitung, Durchführung und Auswertung zu betreiben ist. Das ist vor allen aus Gründen der Übertragbarkeit in andere Fakultäten und auf einheitliche Prüfungen mit Berechtigungsfunktion wichtig. Reformen finden bisher meist nur in kleinen Feldern statt, haben Modellcharakter, wenige sind beteiligt, aber hochengagiert und bereit, große Kosten in Kauf zu nehmen. Für die Ergebnisevaluation wird teilweise zu viel organisatorischer Aufwand betrieben, der unter anderen Bedingungen nicht mehr möglich ist. Das soll nicht heißen, daß in jedem Fall hohe Investitionen in die inhaltliche Vorbereitung gleich welcher Prüfungsform notwendig sind und deren Qualität bestimmen.

So kann z.B. eine gut vorbereitete und strukturierte mündliche Prüfung, die sich potentieller Fehlerquellen wie des Primacy-, Recency- und Halo-Effekts bewußt ist, für die u.a. ein Katalog von Bewertungskriterien und ein nach Kriterien differenzierter Beobachtungs- und Bewertungsbogen erstellt wird, in der die Fragen vorbereitet, klassifiziert und dokumentiert, sowie Kriterien für Form und Inhalt des Protokolls erstellt werden und die vor allem ihre interaktive Form nicht primär zur Wissensabfrage, sondern zur Überprüfung von Problemlösungsfähigkeit nutzt, ein valides, bedingt reliables und deshalb jedenfalls nur formativ einzusetzendes Instrument der Ergebnisevaluation eines Reformcurriculums und seines möglichen Ziels der Ausbildung und Förderung von Problemlösungsfähigkeit sein.

So ist z.B. die MC-Prüfungsform als solche, d.h. unter anderen als den gegenwärtigen Gebrauchsbedingungen, ein hochreliables und zugleich valides Instrument der Evaluation einer Teildimension kognitiver Fähigkeiten, die zur ärztlichen Kompetenz nun einmal ebenso unerläßlich dazugehört, wie bestimmte praktische Fertigkeiten, Fertigkeiten oder Schlüsselqualifikationen.

III. Schluß

Es gibt eine Reihe alternativer, an deutschen medizinischen Fakultäten bereits erprobter Prüfungsformen, z.B. das Objektiv strukturierte klinische Examen (OSCE)(7), die Dreisprung-Prüfung (Triple-Jump)(8), Modified

Essay Question-Prüfungen(9) oder den in Maastricht entwickelten Progress-Test, die unterschiedliche Fähigkeiten entweder differenziert oder integrativ und mehr oder weniger valide und reliabel messen und deshalb mehr oder weniger für Prüfungen mit Berechtigungsfunktion geeignet sind. Auch neu entwickelte Formen sind denkbar. Wichtig für ihren Einsatz als Evaluationsinstrumente curricularer Reformen ist vor allem folgendes: Prüfungsformen und Prüfungslernziele sollten zusammenpassen und diese wiederum die mit einer Studienreform verbundenen vielfältigen Lernziele repräsentieren. Nur dann läßt es sich mit der Paradoxie, daß Prüfungen als curricularer Gestaltungsfaktor ihre eigene Ergebnisevaluation sind, gut leben.

Literatur

- 1 Friedman, Ch.P. u.a.: Charting the Winds of Change: Evaluation Innovative Curricula. *Academic Medicine* 1/90. 8-14
- 2 Bourdieu, P.; Passeron. J.C.: Die Illusion der Chancengleichheit. Stuttgart 1971. 162
- 3 Bourdieu, P.; Boltanski, L.: Titel und Stelle. Zum Verhältnis von Bildung und Beschäftigung. In: Bourdieu, P. u.a.: Titel und Stelle. Frankfurt a.M. 1981. 89-115
- 4 Bourdieu, P.: Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft. Frankfurt a.M. 1988
- 5 Bourdieu, P.; Passeron. J.C.: Prüfung einer Illusion. In: Die Illusion der Chancengleichheit a.a.O. 161-190
- 6 Bloom, B.S.: Taxonomie von Lernzielen im kognitiven Bereich. Weinheim/Basel 1976
- 7 Falck-Ytter, Y.; Stiegler, I.: Beurteilung klinischer Fähigkeiten mit dem OSCE-Verfahren. *Medizinische Ausbildung* 10 (1993). 48-55
- 8 Collin, M.; Ciliska, D.: Revitalizing problemsolving with Triple Jump. *Canadian Nurse* 3/83. 41-44
- 9 Knox, J.D.E.: What is ... a Modified Essay Question? *Medical Teacher* 1/89. 51-55

Dr. Ingrid Stiegler
Kampstr. 112
D 42781 Haan