

ZUR GESCHICHTE DER ÄRZTLICHEN PRÜFUNGEN

H. E. Renschler, Bonn ⁽¹⁾

Zusammenfassung:

Am Ende des 17. Jahrhunderts wurde in den deutschen Staaten für die Zulassung zur ärztlichen Berufstätigkeit nach Abschluß der universitären Ausbildung zum Arzt das Bestehen einer vor der Landesgesundheitsbehörde abzulegenden Prüfung eingeführt. Damit sollte die Bevölkerung vor schlecht ausgebildeten Ärzten geschützt werden. Seit dem Einführen der Krankenversicherung im Deutschen Reich im Jahre 1883 hat jeder Bürger Anspruch auf eine qualifizierte ärztliche Betreuung. In der Prüfung mußte daher am Ende des vergangenen Jahrhunderts jeder Kandidat 7 Patienten über bis zu 14 Tage an insgesamt 27 Prüfungstagen behandeln. Wie die dargestellte Übersicht zeigt, wurden mit den späteren Prüfungsordnungen die Anforderungen in den Prüfungen stark reduziert. Da mündlich-praktische Prüfungen die testtheoretischen Forderungen nach Objektivität und Zuverlässigkeit nicht erfüllten, wurden sie unter Vernachlässigung der Gültigkeit stark eingeschränkt. Nach der Approbationsordnung von 1970 wird nur noch eine zeitlich begrenzte Prüfung an 1 - 3 Fällen und ohne Abfassen eines schriftlichen Berichtes gefordert. Nach 1956 wurde die Methodik der Prüfungen am Patienten besonders für die Facharztprüfungen im Ausland soweit verbessert, daß mündliche Prüfungen auch unter hohen Ansprüchen an Objektivität und Zuverlässigkeit abgehalten werden können. Die dazu erforderlichen Bedingungen werden besprochen. Diese sollten bei der Wiedereinführung von mündlichen Prüfungen eingehalten werden.

Ein Bericht über die Geschichte der ärztlichen Prüfungen entspricht der besten Tradition der deutschen Universität.

Als im Jahre 1900 die Hochschulpädagogik als neue Disziplin gefordert wurde und der damals berühmte Berliner Astronom W. Foerster die Anwaltschaft dafür übernommen hatte, hat Paulsen mit seiner ganzen Autorität widersprochen. Er hielt für den für ihn im Mittelpunkt des Universitätsstudiums stehenden Vorlesungsbetrieb eine reine Meisterlehre für ausreichend (Paulsen 1912 a). Zugestimmt hat er dagegen einer Erforschung der Geschichte des Hochschulunterrichtes und einer kritischen Darlegung der gegenwärtigen Zustände und Erscheinungen im

A History of Medical Licensing Examinations

Abstract:

At the end of the 17th century state examinations were introduced in the German states. Promotion by a medical faculty of a university was no longer sufficient for the practice of medicine. An examination held by the health care authority of the state had to be passed. This was to protect the population from unqualified health care professionals. Every citizen in Germany is entitled to qualified medical care since 1883. The state examination had to guarantee the competence of all licensed physicians. The examination, therefore, had to be a valid measurement of performing medical care. At the end of the last century in the clinical examinations candidates had to treat 7 patients for up to 14 days, seeing each patient at the least twice a day. The total number of days on which candidates had to see patients was 27. Written case reports, which were mandatory before 1901 were no longer required after 1970. All conditions were consequently reduced. With the regulations of 1970 candidates could be examined on 1 to 3 short cases only. Otherwise the previously oral and practical examinations were replaced by non valid multiple choice examinations, as oral examinations had low reliability and no objectivity. Since 1956 the methodology of practical oral examinations was improved mainly for the certification of specialists. Requirement for reliability and objectivity can now be met. Corresponding conditions should be applied when oral examinations are introduced on a larger scale in 1988.

Hochschulwesen aller Länder. Ich werde mich dementsprechend auf die geschichtliche Entwicklung im deutschen Kulturraum konzentrieren, aber auch kritische Vergleiche mit internationalem Ausblick anschließen. Ich lasse offen, ob wie bei Paulsen deutsche Bildung gleich Menschheitsbildung sein soll (Paulsen 1912 b).

Die erste für uns noch heute wichtige Entscheidung fiel 1685. Sie unterscheidet die deutschen Prüfungen von denen fast aller anderen Staaten der Welt. Der Große Kurfürst richtete aus "Landesväterlicher Fürsorge um die Wohlfahrt Unserer Untertanen" in Übereinstimmung mit den meisten deutschen Staaten dieser Zeit mit dem Medizinedikt für Brandenburg eine Landes-

(1) Kurzreferat gehalten am 2. Juni 1988 bei der 17. Arbeitssitzung der Gesellschaft für Medizinische Ausbildung in Münster

gesundheitsbehörde ein (Fischer 1933). Danach reichten die Zeugnisse der Universitäten allein nicht mehr zur Ausübung des ärztlichen Berufes aus, es war eine zusätzliche Approbation des Collegium Medicum erforderlich. Daraus entwickelten sich die über lange Zeit zentral in der Landeshauptstadt abgehaltenen Staatsexamen.

Das Ziel der Staatsprüfung war seit 1685 politisch bestimmt und nicht wissenschaftlich. Die Auseinandersetzung über die Unvereinbarkeit dieser beiden unterschiedlichen Ziele hält in den Vereinigten Staaten bis heute an.

In Deutschland trägt die Verantwortung für die ärztliche Prüfung ausschließlich das Landesprüfungsamt. Die Universitäten haben die Rolle von kritischen, aber interessierten-Zuschauern.

Die Staatsprüfung sollte die Bevölkerung vor Schaden durch unfähige Heilpersonen, auch durch Ärzte, die ja früher nicht alle Kranken betreuen konnten, schützen. Seitdem die gesundheitliche Versorgung den Rang eines Grundrechtes hat, soll die Staatsprüfung sicherstellen, daß die jetzt allen Patienten zu leistende ärztliche Versorgung eine Mindestqualität hat. Im Gegensatz zur früheren Maxime "NIL NOCERE" soll in modernen Staaten mit Sozialgesetzgebung, was Deutschland seit 1883 ist, die Kompetenz der zukünftigen Ärzte durch das Staatsexamen gesichert werden.

Es lassen sich somit 2 Kriterien für die ärztliche Prüfung definieren:

1. Der positive Nachweis einer umfassenden ärztlichen Kompetenz
2. Die prospektive Aussage des Ergebnisses:

Aus dem Verhalten in der Prüfung soll auf die zukünftige Leistung bei der Betreuung von Patienten geschlossen werden. Dies ist um so besser möglich, je mehr die Prüfungssituation der Berufspraxis entspricht, hat allerdings zur Voraussetzung, daß dieses Verhalten auch schon eingeübt worden ist. "Zuschauen", was das Merkmal der Tätigkeit im Praktischen Jahr geworden ist, reicht also nicht.

Der Norddeutsche Bund und das Deutsche Reich

In den verschiedenen deutschen Staaten unterschieden sich die Staatsprüfungen nur gering. Voraussetzung für die Zulassung war bis 1869 die Promotion durch eine medizinische Fakultät, also der erfolgreiche Abschluß der wissenschaftlichen Ausbildung (Kletke 1874). Häufig, wie in Bayern und Württemberg, war eine weitere Voraussetzung ein nach dem Universitätsstudium und vor der Zulassung zum letzten Teil des Staatsexamens abzuleistendes Praktisches Jahr.

Ich werde mit der genaueren Besprechung in der Mitte des vergangenen Jahrhunderts

einsetzen. Die Prüfungsordnung von 1852 war insofern wichtig, als damit alle Ärzte auch in Chirurgie geprüft werden mußten. Diese Vereinigung der beiden medizinischen Hauptdisziplinen war 1795 bereits in Frankreich allgemein eingeführt worden (Ecole de Santé de Montpellier 1795). Die deutsche Medizin war somit in dieser Beziehung am Ende der 1. Hälfte des vergangenen Jahrhunderts fast 60 Jahre im Rückstand gegenüber der führenden Nation Frankreich. Dies galt aber auch für die neuen Methoden der Klinik wie Perkussion und Auskultation, die erst nach 1840 Eingang fanden (Wunderlich 1859). Die Steilheit des deutschen Aufstieges zur Weltspitze, die nach 1860 erreicht wurde, wird dadurch noch deutlicher.

Welchen Anteil hatten daran die Ausbildung und die Prüfungen?

Ich werde jetzt die Staatsprüfung von 1869 eingehender besprechen. Sie wurde nach der Reichsgründung ab 1872 in allen deutschen Bundesstaaten übernommen. Voraussetzung für die Zulassung zur ärztlichen Prüfung war das Reifezeugnis und das Abgangszeugnis von der Universität, das auch noch politischen Inhalt haben konnte, sowie der Nachweis über die Teilnahme als Praktikant an der chirurgischen, medizinischen und geburtshilflichen Klinik. Das 1861 abgeänderte "Tentamen physicum", jetzt als "naturwissenschaftliche Prüfung" bezeichnet, mußte in der medizinischen Fakultät bestanden sein. In der Gewerbeordnung vom 21. Juli 1869 war bereits festgelegt worden, daß die Approbation nicht von der vorherigen Doktorpromotion abhängig gemacht werden darf. Das Staatsexamen wurde nicht mehr, wie bisher, in der Hauptstadt, sondern in den Universitätsstädten abgenommen. Zum Prüfer konnten wissenschaftlich gebildete Fachmänner aus allen Zweigen der Heilkunde berufen werden, die klinischen Prüfungen fanden in größeren Krankenhäusern oder Universitätskliniken statt.

Bemerkenswert ist, daß außer in den klinischen Fächern Chirurgie, Ophthalmologie, Innere Medizin sowie Geburtshilfe und Gynäkologie nochmals, aber jetzt vor der staatlichen Prüfungskommission, in Anatomie und Physiologie geprüft wurde. Ein weiteres Fach war die pathologische Anatomie. In die alle Stoffgebiete umfassende Schlußprüfung war Staatsarzneikunde oder Hygiene einzu beziehen. In Anatomie war ein Nervenpräparat anzufertigen, in Pathologie die Sektion einer Leiche oder mindestens eines Leichenteiles zu machen.

Die klinischen Prüfungen verdienen besondere Beachtung und können zum Vergleich mit der jetzigen Wiedereinführung herangezogen werden. Die Einzelheiten sind in der Tabelle 1 enthalten. In den drei großen Kliniken wurden den Kandidaten jeweils von 2 Prüfern je ein Patient zugeteilt. Diese mußten in Gegenwart des Examinators untersucht werden. Diagnose, Prognose und Heilplan mußten sofort angegeben

werden und in Klausur unter Aufsicht und ohne Zeitbegrenzung schriftlich abgefaßt werden. In den folgenden 7 Tagen mußten die betreuten und von den Kandidaten selbst behandelten Patienten zweimal täglich von den Kandidaten besucht und die Verläufe geschrieben werden. Die beiden Prüfer hatten an je 3 Morgenvisiten der Kandidaten teilzunehmen und dabei auch an anderen Patienten zu prüfen. In der Chirurgie mußten kleinere Operationen am Patienten, eine große an der Leiche ausgeführt werden, in Geburtshilfe die bei normalem Geburtsverlauf üblichen Hilfen. In Geburtshilfe sowie über Frakturen und Luxationen wurde am Phantom geprüft. In den theoretischen Prüfungsteilen wurden die Aufgaben durch Los bestimmt. Die gesamte Prüfung hatte 5 Abschnitte, von denen mindestens 4 mit der Note "mittelmäßig" bestanden sein mußten.

2 Patienten behandelt und darüber schriftlich berichtet werden. Ein anonym englischer Autor, der Wien 1817 besucht hatte, kam zu dem Schluß, daß jeder der etwa 60 Studenten in Wien im letzten Studienabschnitt wenigstens 10 - 12 Patienten unter der Aufsicht der Professoren selbst betreut hat (Anonymus 1819).

Vor der nächsten umfassenden Änderung der Prüfung im Jahre 1901 wurde 1883 die Mindeststudiedauer von 8 auf 9 Semester erhöht und zusätzliche Prüfungen in Zoologie und Botanik in die ärztliche Vorprüfung sowie Hygiene in die ärztliche Prüfung eingeführt.

Mit der ärztlichen Prüfung von 1901 wurde die Trennung zwischen Vorklinik und Klinik für uns endgültig festgelegt. Die ärztliche Vorprüfung war jetzt vor einer von der vorgesetzten Zentralbehörde berufenen Prüfungskommission abzulegen.

Tabelle 1

Übersicht über die ärztliche Prüfung

	1869	1901	1965	1970
Vorklinische Fächer	2	0	0	0
Klinisch-theoretische Fächer	2	2	4	*
Nichtklinische Prüfer	4	2	4-6	0
Klinische Fächer	3	5	8	3
Klinische Prüfer	7	8	12-13	3-5
Anzahl der Fälle	7	7	10	1-3
Prüfungstage mit Fall	27	23	14	1
Beobachtungstage der Fälle	51	29	3	0
Vorgeschrriebene Visiten	28	26	5	0
Davon mit Examinator	12	14	1	0
Schriftliche Berichte	7	7	10	0
Schriftliche Prüfungen	1	1	1	3

* Gesamtzahl der Stoffgebiete: 48

Eine Besonderheit, die nur das deutsche Prüfungssystem aufweist, ist in der Prüfungsordnung festgelegt: Die maximale Anzahl der Kandidaten pro Termin ist mit 3 bzw. 4 angegeben. In der österreichischen Prüfungsordnung war 1810 gegensätzlich bestimmt worden, daß niemals 2 Kandidaten oder Kandidatinnen zugleich geprüft werden durften (Fakultäts-Kommission 1847). Dies entspricht meines Wissens bis heute der allgemeinen Regelung außerhalb Deutschlands. Es wäre eine vor allem im internationalen Vergleich interessante Fragestellung für eine Untersuchung, wie sich diese Bedingung auswirkt.

In Wien mußten ab 1810 bei der Meldung zur Prüfung Berichte über 2 von den Kandidaten im 5. Ausbildungsjahr unter Aufsicht behandelte Patienten eingereicht werden. Diese mußten von allen Professoren der Fakultät schriftlich genehmigt werden; wurde diese Genehmigung von mehr als 2 Professoren verweigert, mußten erneut

Nach vollständig bestandener ärztlicher Prüfung war das Praktische Jahr abzuleisten (Joachim und Korn 1914). Im Staatsexamen war von den vorklinischen Fächern Anatomie und Physiologie nur noch der Prüfungsteil topographische Anatomie in der Chirurgie enthalten. Klinisch-theoretische Fächer waren pathologische Anatomie und Pathologie sowie Hygiene als 2 eigene Abschnitte. Die Anzahl der klinischen Prüfungsabschnitte wurde auf 5 Fächer erhöht, Augenheilkunde und Irrenheilkunde waren selbständig geworden. Kinderkrankheiten, Hals-Nasen-Ohren-Krankheiten, Haut- und venerische Erkrankungen sollten in anderen Fächern mitgeprüft werden. Die Kandidaten - Frauen wurden damals noch abgelehnt - mußten über 2 Semester an der medizinischen und chirurgischen Klinik sowie an der Klinik der Geburtshilfe über 2 Semester und an weiteren 6 Kliniken über ein Semester als Praktikant teilgenommen haben. Bei einer Gesamtstudiedauer von jetzt 10 Semestern mußten 4 nach Ablegen

der ärztlichen Vorprüfung zurückgelegt sein.

Bei gleicher Anzahl der Patienten, die in Anwesenheit des Prüfers oder, wie in der Geburtshilfe, in Anwesenheit eines beauftragten Assistenten untersucht werden mußten, wurde die Anzahl der Beobachtungstage pro Fall von 7 auf 4 reduziert. Bei diesem Parameter ergab sich somit die relativ stärkste Reduktion der Prüfungsleistungen auf 57 % (Tab. 1). Die Anzahl der schriftlich zu erstellenden Fallberichte blieb gleich. Die Zahl der in der Prüfungsordnung vorgeschriebenen Visiten durch die Kandidaten reduzierte sich nur gering, die der Visiten mit Anwesenheit des Prüfers erhöhte sich formal, da die in der Augenheilkunde jetzt im Wortlaut der Ordnung aufgeführt wurde.

Andere Leistungen als ein extemporierter Vortrag waren weiterhin die vollständige Sektion mindestens einer Körperhöhle und das Herstellen eines histologischen Präparates in der Pathologie. In der Chirurgie mußten 2 Operationen, darunter eine Arterienunterbindung, an der Leiche durchgeführt werden. An Phantomen waren Leistungen bei Frakturen und Luxationen sowie in Geburtshilfe zu erbringen. Als schriftlicher Prüfungsteil blieb die Lösung einer pharmakologischen Aufgabe in der medizinischen Klinik.

Die letzte Bestallungsordnung

Die letzte Fassung der Bestallungsordnung vom 31. 05. 1965 kennt ein Teil von Ihnen noch aus der Sicht des Kandidaten, unsere in Kürze abtretende Generation noch als Prüfer. Es waren darin neben 4 nichtklinischen Abschnitten 8 klinische Prüfungsfächer zu absolvieren (Rachold 1967). Die Forderung, daß die Patienten in Gegenwart des Prüfers zu untersuchen sind, wurde noch in 5 Fächern gestellt, der Professor der Frauenheilkunde konnte sich durch einen Assistenten bei der Untersuchung der Gebärenden vertreten lassen, was bei der Dringlichkeit und der Tageszeit dieser Untersuchung verständlich ist. 4 Visiten waren bei der Wöchnerin in den der Entbindung folgenden 48 Stunden vorgeschrieben. In der Kinderheilkunde mußte das am 1. Tag untersuchte kranke Kind am 2. Tag unter Aufsicht des Prüfers behandelt werden. Es war also nicht möglich, Arzt zu werden, ohne in der Lage zu sein, ein Kind halten zu können. Pädiater beanstanden dies an der Approbationsordnung.

Die schriftlich zu erledigende Aufgabe des Rezeptierens war von der Inneren Medizin in die jetzt getrennte Pharmakologie verlegt worden.

Nach der Prüfungsordnung von 1869 wurden insgesamt 51 Beobachtungstage der Fälle gefordert, während denen 28 Visiten, 12 davon mit Anwesenheit des Examinators vorgeschrieben waren. Diese Dimensionen entfielen mit der Approbationsordnung von 1970 völlig. Die Vorschrift, Gebärende über 2 Tage zu besuchen und das in

der Kinderklinik untersuchte Kind nach einem Tag zu behandeln, was bei meinem Staatsexamen verlangt worden ist, ist mir aus dem Gedächtnis entschwunden. Nur durch das Nachlesen der Bestallungsordnung konnte ich dies nachvollziehen. Die Forderung, daß der Prüfer bei der Untersuchung des Patienten durch den Kandidaten anwesend sein muß, wurde nach meiner Erinnerung nicht im strengen Wortlaut durchgeführt, bei mir genügte dafür das Wiederholen eines Teiles der Untersuchung.

Vergleichende Übersicht über die Entwicklung des deutschen Staatsexamens

Der in der Tabelle 1 dargestellte systematische Abbau der geforderten Leistungen über ein Jahrhundert ist offensichtlich und beeindruckend. Dies gilt vor allem für die Anforderung, daß die Prüfungssituation möglichst der Berufssituation entsprechen soll, die Prüfungen also valide sind. Die Gültigkeit gilt als das wichtigste Merkmal von Prüfungen (Veloski et al 1987). Über die Auswirkung dieser Diskrepanz gibt es bei uns weder theoretische Grundlagen noch empirische Untersuchungen (Renschler 1987 a). Wie ist nun der Zeitraum zwischen 1860 und 1910 zu beurteilen, in dem die deutsche Medizin führend in der Welt war?

Entsprechend der Ideologie des deutschen Universitätsstudiums wird in der Ausbildung kein direkter Zwang auf die Studenten ausgeübt. Jaspers spricht dies klar aus: "Die Lehrer geben keine Anweisung und keine persönliche Führung", da die Studenten die Reife der Selbstverantwortung tragen (Jaspers u. Rossmann 1961). Die Universität nimmt in Kauf, daß Studenten "aus Mangel an Leitung und Vorschrift möglichst gar nichts lernen" (Seite 76). Da es zu dieser Zeit keine Gegenstandskataloge und keine Lernziele gab, mußten die Studenten ihre Zielvorstellung und ihr Lernen von den Prüfungsanforderungen ableiten. Eine Verstärkung der Wirkung der Prüfungsbedingungen auf das freie Lernverhalten wäre zu erwarten gewesen, wenn durch die Anwendung der damals sehr strengen Bestehensbedingungen für die Prüfung eine Auslese getroffen worden wäre. Dies hätte auch der politischen Forderung der Sicherstellung der Kompetenz der Ärzte, die zur Berufsausübung zugelassen werden sollen, entsprochen. Für Deutschland bestand diese Forderung ab 1883 mit der Einführung der Sozialgesetze.

Üblicherweise setzen etwa 10 Jahre nach Einführung einer Änderung der Studienbedingungen Diskussionen ein, die nach weiteren 20 Jahren zu einer erneuten Änderung führen, mit der dann oft der alte Zustand wieder hergestellt wird (Sprengel 1911). Im Rahmen dieser Diskussionen befragte Quincke seine damaligen Kollegen und erhielt das Ergebnis, "daß jeder in

die Staatsprüfung eingetretene Kandidat schließlich die Zulassung zur Praxis erreichte" (Quincke 1890). Mit Recht erregte dies seine Zeitgenossen, die die Richtigkeit seiner Aussagen in Frage stellten. So konnte denn auch in Freiburg gefunden werden, daß zwischen 1874 und 1883 bei etwa 500 Kandidaten 3 Fälle waren, die auch beim dritten Versuch und somit endgültig das Staatsexamen nicht bestanden hatten (v. Kries, 1893). In München fand von Ziemssen 4 entsprechende Fälle. Es ist also davon auszugehen, daß mit wenigen Ausnahmen vor 1900 von den deutschen Medizinstudenten ärztliches Verhalten vor dem Staatsexamen nicht ausreichend eingeübt wurde und so auch die abstrakt dargebotene Theorie nicht praxisrelevant wurde (Guttstadt 1892). Dies war der Grund für die Wiedereinführung des praktischen Jahres, die aber schon nach 10 Jahren als mißglückt angesehen wurde (Renschler 1987 b).

Die Neigung zur milden Beurteilung der Kandidaten blieb bis zum Ende der Bestallungsordnung bestehen. Von 639 Kandidaten der Jahre 1968 bis 1972 einer westdeutschen Universität haben 91 % das Staatsexamen mit der Note "gut" oder "sehr gut" (43 %) bestanden, 0,5 % hatten endgültig nicht bestanden. 3,4 % hatten beim Erstellen der Statistik die Prüfung noch nicht abgeschlossen. Nach der Mitteilung unseres Landesprüfungsamtes an die Fakultäten erhalten jetzt (1984/85) noch 60 - 70 % "gut" oder besser und nur noch etwa 20 % "sehr gut". Die Durchfallquote liegt weiter bei etwa 1 %.

In einer Untersuchung des in Autobiographien festgehaltenen Studierverhaltens deutscher Koryphäen zwischen 1872 und 1901 konnte Irrgang zeigen, daß 14 der von ihm untersuchten 21 Wissenschaftler während ihres Studiums Patienten in eigener Verantwortung unter Aufsicht und Bewertung durch Dozenten ärztlich betreut hatten (Irrgang 1988). Dies geschah vor allem bei der freiwilligen Famulatur, im Regelunterricht geschah es nur dreimal in einer Poliklinik bzw. in einer Universitätsfrauenklinik.

Das letzte Kriterium wäre die Überprüfung der prospektiven Bedeutung des Staatsexamens, also eine Beurteilung der Qualität der ärztlichen Berufstätigkeit mit einem Bezug zur Prüfungsleistung. Aus Deutschland sind mir derartige Untersuchungen nicht bekannt. Für die ärztliche Fortbildung ist diese Forderung im letzten Jahr gestellt worden und wird zur Zeit von allen Landesärztekammern diskutiert. Unsere eigenen Ansätze im Jahre 1982 blieben stecken, obwohl die KV ihre finanzielle Unterstützung damals zugesagt hatte.

Einstellung der mündlich-praktischen Prüfungen

Als letztes Thema, das bereits historisch bearbeitet werden kann, möchte ich die

Hintergründe für die Änderungen zwischen 1965 und 1970, also zwischen der Bestallungsordnung und der Approbationsordnung beleuchten.

Mit der Einführung objektiver Prüfverfahren, bei denen die Prüfungsleistungen frei von der Variabilität menschlichen Urteilens bewertet werden, wurden an alle Prüfungen strenge Maßstäbe angelegt. So wurden auch die mündlichen Prüfungen und die praktischen Prüfungen am Krankenbett psychometrisch neu bewertet. Es zeigte sich, daß sie den Forderungen nach Objektivität und Zuverlässigkeit nicht entsprachen. Wurde also dieselbe Prüfungsleistung von verschiedenen Prüfern oder von einem Prüfer zu verschiedenen Zeiten benotet, ergaben sich starke Unterschiede. Selbst bei Korrelationskoeffizienten um 0,80 wichen die Noten, die von verschiedenen Prüfern gegeben wurden, sehr stark voneinander ab. Bei den Versuchen von Wilson und Mitarbeitern in Glasgow, die 28 aufgezeichnete Prüfungen von 14 Prüfern bewerten ließen, fielen nur 13 Studenten bei keinem Prüfer durch, hatten also als Mindestnote noch ausreichend erhalten, dafür aber mindestens einmal die Note "sehr gut" (Wilson u. Mitarb. 1969).

Einige der Autoren mit kritischen Arbeiten sind in die Zeittafel eingetragen (Tab. 2). Schon fast

Tabelle 2
Mündliche Prüfungen

RELIABILITÄT			
Unzureichend		Ausreichend	
Jahr	Aut. or	Jahr	Aut. or
1929	Barnes & Pressey		
1934	Triebler		
1935	Hartog & Rhodes GB		
		1958	Bull
1958	Goldstein	1962	Carter
1963	Hubbard (NBME)		
		1965	RCGP (London) *
1969	Wilson u. Mit. GB	1969	Foster *
		1970	Levine & McGuire *
		1971	Kelley *
1975	Kosteni	1975	Kittle *
		1975	Meskauskas *
		1980	Maatsch *

* = Facharztprüfung Zitiert nach Neufeld & Norman 1985

gleichzeitig wurden aber erfolgreiche Versuche unternommen, die Zuverlässigkeit der mündlichen Prüfungen zu verbessern. Darauf möchte ich ganz zum Schluß eingehen.

1963 entschloß sich in den USA daher das National Board of Medical Examiners (NBME), die alte mündliche Prüfung zu Gunsten komplexerer Testverfahren, die unter dem Begriff PMP (Patient Management Problem) bekannt sind, aufzugeben. Dies war aus 4 Gründen, die aber für Deutschland nicht gelten, möglich.

1. Vor der Zulassung zur Staatsprüfung muß in den U.S.A. der erfolgreiche Abschluß des Gesamtstudiums durch die Fakultät bestätigt werden.
2. Während des Studiums führt der amerikanische Student mit gesetzlicher Billigung durch Staatsgesetze mit selbständigen Entscheidungen ärztliche Maßnahmen in großer Anzahl durch. Jede Fakultät verfügt über eine hinreichend große Anzahl von Dozenten, die die klinische Ausbildung der Studenten überwachen.
3. Zuständig für die Erteilung der ärztlichen Lizenz ist der einzelne Bundesstaat, der für seine Entscheidungen in freier Verfügung mehrere Prüfverfahren, nicht nur die des National Board verwenden kann.
4. Teil III des National Board wird erst nach einem erfolgreichen Abschluß des 1. post-universitären Jahres gültig. Während dieser Zeit hat der junge amerikanische Arzt eine unerhörte Arbeitsleistung zu vollbringen. Es wird allerdings jetzt diskutiert, daß die Wochenarbeitszeit im Krankenhaus von 110 auf 80 Stunden im Durchschnitt reduziert werden soll. Für dieses erste Jahr, das nur im Rahmen eines anerkannten Weiterbildungsprogrammes absolviert werden darf, stehen zur Zeit 30 % mehr Stellen als für die Absolventen benötigt zur Verfügung.

Es ist mir nicht verständlich, warum die Aufgabenform der Patient Management Problem (PMP) nicht in den schriftlichen Prüfungen nach Paragraph 14 der ÄAppO benutzt werden. Es wäre dies nach dem Wortlaut "Der Prüfling ... hat dabei anzugeben, welche der mit den Fragen vorgelegten Antworten er für zutreffend hält" möglich. Noack hatte am 08.02.1968 in einem Referat vor der Kleinen Kommission zur Änderung der Bestallungsordnung den Wert dieser Form der objektiven Prüfung dargelegt (Noack 1968). Es kann angenommen werden, daß mit diesem Instrument Fertigkeiten gemessen werden, die mit den üblichen Fragen mit Auswahlantworten nicht erfaßt werden und die besonders das Problemlösen betreffen (McGuire 1984). In der Zwischenzeit wurden mehrere Bewertungssysteme entwickelt und geprüft (Neufeld u. Norman 1985). Auch eine erste deutsche Untersuchung konnte zeigen, daß das Ergebnis einer solchen Prüfung vom verwendeten Bewertungssystem abhängt

(Fuchs, in Vorbereitung). Vor einer Einführung des PMP als Prüfungsinstrument müßten diese in den Fakultäten erprobt werden und in einer wissenschaftlichen Untersuchung unter unseren Bedingungen das beste Auswertungssystem ermittelt werden. Bei dieser Gelegenheit sollte gleichzeitig die bisher benützte, aber überholte Auswertung auf der Grundlage der klassischen Testtheorie verlassen werden.

Nach einer 1977 eingeleiteten Vorbereitungsphase wurde nach 1979 die Auswertung der Prüfungen durch das NBME auf das nach 1950 entwickelte Auswertungsverfahren nach dem RASCH-Modell umgestellt (Kelley & Schumacher 1984). Seine Anwendung setzt eine Spezialausbildung in Statistik voraus, wie sie nur ausnahmsweise bei Klinikern zu finden sein dürfte.

Bei der Kritik blieb die Beurteilung der Gültigkeit, die bei der alten klinischen Prüfung gegeben war, weitgehend außer Betracht, da sie der direkten Meßbarkeit nur schwer zugänglich ist. Analysen der Inhalte der mündlichen Prüfungen zeigten aber, daß nicht, wie angenommen, vorwiegend klinisches Denken geprüft wurde. Zu 70 % wurde Wissen abgefragt, Problemlösen hatte einen Umfang von nur 13 % (McGuire 1966).

Die ersten Ergebnisse der psychometrischen Bewertung der mündlichen Prüfungen stammen aus dem Jahr 1929. In Deutschland begann die pädagogische Diagnostik erst Mitte der Sechziger Jahre, ein erstes Buch darüber erschien 1978 (Birkel 1978).

Fast gleichzeitig mit der Formulierung der Kritik an der mündlichen Prüfung wurden Verbesserungen des Prüfungsverfahrens entwickelt, die den Anforderungen nach Objektivität und Zuverlässigkeit entsprachen (Tab. 2).

Die zu fordernde Anzahl von Prüfungsfällen und der gesamte übrige Aufwand haben die Durchführung erschwert. Erst nachdem eine Verbesserung der Reliabilität erreicht war, kam die Überlegenheit der Validität mündlicher Prüfungen wieder zur Geltung. Mit verbesserten, standardisierten mündlichen Prüfungen konnte gezeigt werden, daß damit andere Fertigkeiten gemessen werden, als durch die Fragen mit Auswahlantworten (Wakeford 1986 b). Die Einführung der standardisierten mündlich-praktischen Prüfung erfolgt daher in Nordamerika und England vor allem im Bereich der Facharztprüfungen, wo die jährliche Anzahl der Kandidaten nur wenige Tausend beträgt (Tab. 2 und Wakeford 1986 a).

Um mündliche Prüfungen, besonders solche am Krankenbett, mit der jetzt für erforderlich gehaltenen und auch möglichen Zuverlässigkeit durchführen zu können, ist ein großer Aufwand erforderlich (Moulds u. Mitarb. 1978). Die dadurch entstehenden Kosten wirkten sich hemmend auf eine breite Einführung aus. Ein in Schottland entwickeltes Format, das objektive standardisierte klinische Examen OSCE, ermöglicht jedoch einen

Einsatz bei großen Zahlen von Kandidaten (Harden u. Mitarb. 1975). Diese Prüfungsform kann auch in Entwicklungsländern wie Ägypten eingesetzt werden und müßte auch bei unseren geringen Mitteln diskutiert werden. Wichtig ist, daß außer der testtheoretischen Grundlage Theorien für die zu prüfenden Fertigkeiten (McGuire 1984) und ihre unterschiedlichen Bewertungen (Maatsch u. Mitarb. 1986) verwendet werden.

Als **Summe** dieser jüngsten Entwicklung ergibt sich, daß unter bestimmten Bedingungen praktisch mündliche Prüfungen mit der erforderlichen Reliabilität abgehalten werden können. Dazu gehören die in Tabelle 3 aufgeführten Merkmale standardisierter Prüfungen.

Tabelle 3

Merkmale standardisierter Prüfungen

1. Zielbestimmung und Planung der Prüfung
2. Schulung und Einübung der Prüfer
3. Evaluation der Prüfer
4. Endgültige Auswahl der Prüfer
5. Objektivierung der Prüfung
6. Gemeinsame Vorbereitung der Fälle
7. Strukturierung der Aufgaben
Orientierung: Modell ärztlichen Denkens und Handelns
8. Ausreichende Anzahl von Fällen (mehr als 10)
9. Wissenschaftliche Begleituntersuchung
Testtheoretische Begründung
10. Standardisierung der Patienten.

Prüfung mit einer Diskussion in der Prüfungskommission verbunden.

Bis zu 100 Kandidaten mußten von einer Kommission geprüft werden, erst wenn es mehr waren, durfte eine zweite Kommission eingerichtet werden. Damit erfolgte eine Einübung in das Prüfen automatisch. In der Literatur finden sich viele Hinweise - aber auch Klagen - über den großen Aufwand, den die Prüfungen verursachten und über die große Belastung auch der Patienten.

Die große Zahl der Examensfälle war nach dem heutigen Erkenntnisstand eine wichtige Voraussetzung für die Reliabilität der Prüfung. Die ursprüngliche Zahl von 7 Fällen konnte bei dem damaligen Umfang der klinischen Fächer gerade noch genügen. Die Steigerung auf 10 Examensfälle im gesamten Staatsexamen nach der Bestallungsordnung erfüllt zwar die heutige Forderung, wird aber nicht den unterschiedlichen Bedingungen der Berufsausübung, zu der die Prüfung berechtigen soll, gerecht. Es müßte dazu eine Strukturierung unter Festlegung zumindest auf Fächergruppen und auf unterschiedliche Anforderungen und Arbeitsbedingungen erfolgen. Es müßten in Teilprüfungen mit je 10 Fällen etwa die folgenden Dimensionen enthalten sein:

- Ambulante Betreuung
- Stationäre Betreuung
- Operative Behandlung
- Medikamentöse Behandlung
- Physikalische Behandlung
- Akute (Not-)Fälle
- Chronische Krankheiten
- Medizinische Indikationen
- Soziale Indikationen.

Nach internationalem Standard müßten diese Bedingungen bei der Wiedereinführung mündlicher Prüfungen eingehalten werden. Wir können abwarten, welche Kriterien die Landesprüfungsämter erfüllen werden.

Besprechung der ärztlichen Prüfungen im Deutschen Reich

Es sollen in einer kurzen Übersicht die Erfüllung der heutigen Anforderungen an mündliche Prüfungen durch die um die Jahrhundertwende im Deutschen Reich durchgeführten Prüfungen dargestellt werden.

Mit der Notwendigkeit, Patienten über insgesamt 51 Tage zu betreuen, war die Prüfung valide.

Die Prüfungsaufgaben für die mündliche Prüfung waren für die in der Prüfungsordnung seit 1869 vorgeschriebene Auswahl durch Los schriftlich im Voraus zu formulieren. Damit war eine Festlegung auf die Ziele jeder einzelnen

Jede Teilprüfung kann nur eine begrenzte Anzahl der sich teilweise widersprechenden Bedingungen erfüllen. Die vorläufige Mindestzahl von etwa 10 Prüfungsfällen gilt nur für eine definierte, in der Prüfung realisierte Situation und muß noch durch weitere Untersuchungen festgelegt werden (McGuire 1984). Für die Sicherstellung der vollen ärztlichen Kompetenz ist die Prüfung unter allen im Beruf zu erwartenden Bedingungen Voraussetzung. Die gesamte Prüfung müßte nach einer vorläufigen Schätzung somit an mindestens 40 Fällen erfolgen.

Stand in Nordamerika

Die Medizinschulen in den U.S.A. führen keine Abschlußprüfung durch, sondern verlegen die Bewertung der klinischen Fertigkeiten in die Clerkships. Diese müssen in den von jeder Fakultät verschieden festgelegten Bedingungen wie Zahl und Form vor der Promotion zum M.D. (Medical Doctor) erfolgreich abgeleistet worden sein. Gefordert werden 5 - 10 Pflichtfächer und 5 - 10 Wahlfächer. Jedes Clerkship wird in freier Form, jedoch durch eine Vorgabe der Klassifikation, strukturiert bewertet. Das National

Board hat eine andere Richtung eingeschlagen und wird sich mit den Computerprüfungen auf die Bewertung des Denkens beschränken. Der Beginn dieser Entwicklung am Ende der 60er Jahre verliert sich schon in der Geschichte, die Einführung ist sichere Zukunft. Ihr Prinzip stellt eine Weiterentwicklung der PMP dar. Demonstrationsfälle, die im November 1987 in Washington vorgestellt wurden, werden an uns später verschickt. In Kanada werden computerisierte Prüfungen seit 1974 für Facharztprüfungen eingesetzt.

Literaturverzeichnis

- Anonymus (1819) Sketches of the medical school of Vienna. Quart J Foreign Med Surg 1:84-192
- Birkel P (1978) Mündliche Prüfungen : Zur Objektivität und Validität der Leistungsbeurteilung. Kamp, Bochum
- Ecole de Santé de Montpellier (1795) Programmes des cours d'enseignement dans l'école de santé de Montpellier. Imprimerie des Sciences et Arts, Paris
- Fakultäts-Kommission (1847) Die älteren Statuten der Wiener medizinischen Fakultät, nebst einer systematischen Zusammenstellung der auf diese bezüglichen gesetzlichen Bestimmungen : im Auftrage der medizinischen Fakultät. Medizinische Fakultät Wien, Wien
- Fischer A (1933) Geschichte des deutschen Gesundheitswesens Band 1: Vom Gesundheitswesen der alten Deutschen zur Zeit ihres Anschlusses an die Weltkultur bis zum Preussischen Medizinaledikt. Herbig, Berlin
- Fuchs U (1988) Untersuchungen zur Qualitätsverbesserung der medizinischen Ausbildung in Gynäkologie und Geburtshilfe durch Einsatz von Fallsimulationen. In Vorbereitung. Tübingen
- Guttstadt A (1892) Über die praktische Ausbildung der Aerzte in den Kliniken. Springer, Berlin
- Harden RM, Stevenson M, Downie WW et al (1975) Assessment of Clinical Competence Using Objective Structured Examination. Br Med J 1:447-451
- Irrgang D (1988) Aspekte der Ausbildung des Mediziners im deutschsprachigen Kulturraum zwischen 1872 und 1901 anhand von Selbstzeugnissen deutscher Ärzte. Dissertation. Medizinische Fakultät, Bonn
- Jaspers K, Rossmann K (1961) Die Idee der Universität. Springer, Berlin, Göttingen, Heidelberg
- Joachim H, Korn A (1914) Grundriss des Deutschen Aerzterechts für Studierende, Aerzte und Verwaltungsbeamte. Fischer, Jena
- Kelley PR, Schumacher CF (1984) The Rasch Model : Its use by the National Board of Medical Examiners. Evaluation & the Health Profession 7:443 - 454
- Kletke GM (1874) Die Medicinal-Gesetzgebung des Preussischen Staates. Grosse, Berlin
- Kries J von, (1893) Zur Organisation der ärztlichen Prüfungen. Mohr, Freiburg/Br, Leipzig
- Maatsch JL, Huang R (1986) An Evaluation of the Construct Validity of Four Alternative Theories of Clinical Competence. Annu Conf Res Med Educ 25:69-74
- McGuire C (1966) The Oral Examination as a Measure of Professional Competence. J Med Educ 41:267-274
- McGuire C (1984) Medical Problem Solving : A Critique of the Literature. Annu Conf Res Med Educ 23:3-13
- Moulds AJ, Hayes TAB, Young KHM (1978) The MRCGP Examination : A comprehensive guide to preparation and passing. MTP Press, Lancaster, UK
- Neufeld VR, Norman GR (1985) Assessing Clinical Competence. Springer, New York
- Noack H (1968) Gründe gegen die mündliche Prüfung. Medizinstudent 9:34-5
- Paulsen F (1912a) Deutsche Bildung - Menschheitsbildung. Aus: Das Deutschtum im Auslande, 22, 1903. In: Spranger E (Hrsg.) Gesammelte pädagogische Abhandlungen von Friedrich Paulsen. Cotta, Stuttgart und Berlin, pp 349-358
- Paulsen F (1912b) Hochschulpädagogik? Hamburgische Wochenschrift "Der Lotse", 1. Jahrg., Heft 10, 1900. In: Spranger Eduard (ed) Gesammelte Pädagogische Abhandlungen von Friedrich Paulsen. Cotta, Stuttgart und Berlin, pp 256-261
- Quincke H (1890) Kritische Bemerkungen zur ärztlichen Prüfungsordnung. Deutsch med Wochenschr 16:530-531
- Rachold R (1967) Bestallungsordnung für Ärzte : 10. Aufl. Deutscher Ärzte-Verlag, Köln, Berlin
- Renschler HE (1987a) Die Praxisphase im Medizinstudium. Springer, Berlin, Heidelberg
- Renschler HE (1987b) Definition der Fallmethode aus ihrer geschichtlichen Entwicklung in den Medizinschulen Europas. Schweiz Rundschau Med (Praxis) 76:981-996
- Sprengel (1911) Zur Frage des "praktischen Jahres" der Mediziner. Deutsch med Wochenschr 37:599-601
- Veloski JJ, Hojat M, Gonella JS (1987) The Validity of Part III of the National Board Examination. Annu Conf Res Med Educ 26:54-59
- Wakeford R (1986a) Development of the Oral Examination as Part of Specialist Certification Examinations - An International Perspective. Annu Conf Res Med Educ 25:341-342
- Wakeford RE, Norman GR, Belton A (1986b) "Something Old, Something New": the Certification Examination of the United Kingdom Royal College of General

Practitioners. Annu Conf Res Med Educ
25:211-216

Wilson GM, Lever R, Harden RM, Robertson JIS,
MacRitchie J (1969) Examination of Clinical
Examiners. Lancet 1:37-40

Wunderlich CA (1859) Geschichte der Medicin.
Ebner & Seubert, Stuttgart

Prof. Dr. H. E. Renschler
Institut für Didaktik der Medizin
Sigmund-Freud-Str. 25
D-5300 Bonn 1