



医学教育测量

美国全国医学考试委员会的考试和经验

[美] J·P·哈伯德 著

刘秉勋 梅人朗 宗世平 译



人民卫生出版社



017379
05039

医学教育测量

——美国全国医学考试委员会的考试和经验

〔美〕J.P.哈伯德 著

刘秉勋 梅人朗 宗世平 译

金问涛 校



人 民 卫 生 出 版 社

内 容 摘 要

本书是美国医学教育测量专家 Hubbard 等人编写的有关医学考试的专著, 本书除扼要介绍医学教育测量技术的理论与实践外, 还介绍了医学考试制度发展的历史、美国全国医学考试委员会的组织和职能、对医学生和医科毕业生, 包括美国医科学生、医科毕业生、住院医师和申请到美国接受毕业后临床训练或行医的外国医科毕业生进行资格鉴定考试的程序和方法, 并附有美国医学考试委员会历届医学考试中使用过的样板试题100余道。本书对了解现代医学考试的理论与实践, 了解美国医学考试制度、改革医学卫生院校考试方法、组织在职卫生人员和医学卫生院校毕业生业务考试均有参考价值。

Measuring Medical Education —The Tests and the Experience of the National Board of Medical Examiners

Second Edition

John P. Hubbard

Lea & Febiger 1978

医学教育测量

——美国全国医学考试委员会的考试和经验

〔美〕J. P. 哈伯德 著

刘秉勋 梅人朗 宗世平 译

人民卫生出版社出版

(北京市崇文区天坛西里10号)

天津市国营武清印刷厂印刷

新华书店北京发行所发行

787×1092毫米32开本 8 1/4印张181千字

1985年10月第1版 1985年10月第1版第1次印刷

印数: 00,001—5,000

统一书号: 14048·4990 定价: 2.00 元

〔科技新书目 98—63〕

译者的话

这本书是1981年年底我国毕业后医学教育考察组访问费城时由美国全国医学考试委员会(National Board of Medical Examiners)向我们推荐的。当时我们所得到的原文系1971年第一版的版本,为了配合卫生部1982年部分医学院校应届毕业生业务统考的需要,由上海第一医学院医学教育研究室翻译后供内部参考,向全国医学院校提供有关统考的参考资料。而后又得到本书1978年的第二版原文,由于1983年全国已有44所医学院校参加第二次应届毕业生业务统考,并有相当数量的中专卫校也正在筹备采用客观的多选题试行统考,为此,我们在1971年第一版译本的基础上对1978年第二版重译整理,向全国发行。

本书共十一章和三个附录,着重介绍了:考试和评价制度的发展;考试委员会及其在建立客观考试中的任务;多选题的类型;临床能力的客观评价;评分和分析;可靠性、有效性和标准的确定;向各医学院报告考试成绩;利用标准化考题进行的资格鉴定考试,作为指导毕业生学习和继续教育的评价;研究和发展以及全国医学考试委员会和公众利益(相关的工作范围)。

本书介绍了美国两种考试,第一种考试分成三个部分,第一部分是医学生在医学院的二年级学习结束后进行,主要是传统的六门基础医学,即解剖学、生物化学、微生物学、病理学、药理学和生理学。第二部分是在医学院四年级结束时进行,主要是六门临床学科,即内科学、妇产科学、儿科

学、预防医学与公共卫生、精神病学与外科学。第三部分是实习期结束后为取得医师执照所必须的临床能力的考试，所用的试题都是在展示图表资料的基础上提出来的，包括处理病人问题（PMP）。第二种考试是州医学委员会的联合执照考试（FLEX），系一次完成，分成三天内进行，第一天的考试相当于第一部分的内容，传统的六门基础课，但以跨学科的方法混合编排，应试者无法辨别每一道试题的学科来源。第二天的考试相当于第二部分内容，传统的六门临床学科，也是以跨学科的形式编排的。第三天的试题采用全国委员会第三部分的试题，以处理病人问题为依据的编成程序考试，以加权平均数来计算成绩，把评分的重点放在临床能力的考试上。

为了便于了解试题的性质，进一步了解试题的广度和深度，我们把附录所列举的三部分试题一并译出，以供参考。

由于我们业务水平有限，对国外的考试方法、内容和评分等缺乏感性认识，加上本书内容涉及面又广，错误之处在所难免，望读者指正，以便重印时修正。

译者

序 言

本书第一版自 1971 年出版以来，全国医学考试委员会的方针和考试程序已经发生了很大的变化，因此，有必要出版医学教育测量第二版。不论全国医学考试委员会本身所担负的责任还是同其他机构合作进行的工作，这一委员会在建立、评分、分析和报告各种客观考试时，其业务活动已得到了很大的发展。这些考试的结果对医学生、正在接受培训的医生和在职医生的知识和临床能力将作出有效而可靠的测量。对作为教育系统产品的个人的测量，可以从总体上分析和研究年级学生和各类医生在不同学习阶段的知识能力，从而有可能对教育系统的效果作出客观的评价。因此，对教育产品作出评价的同时，也可以对教育过程作出客观的评价。

我们并不认为通过考试所作出的客观测量就一定是评价医学生或医学教育的最合适或最理想的方法。事实上，由他们的教师对学生个人进行连贯的评价具有许多好处，不论校外团体所主持的考试多么好，教师的评价不可能被正式的、偶尔进行的校外考试所取代。对学院一级来说，对各医学院校的正式认可已得到了满意的建立，作为美国医学院协会和美国医学学会的职能之一，它是通过由这两个机构联合组成的医学教育连络委员会来执行，由医学教育连络委员会派代表定期到各医学院校进行现场调查，了解教学设施状况，并同师生进行直接交谈。

但是，对学生个人医学知识和能力的客观测量，正如本

书以下各章所述及的，它已成为教育系统中一个新的和有价值的领域。由于正确地认识到医学教育系统的原料，即医学学生的多变性，所以，当他们进入医学院后，按照教育进程，对他们的知识和能力进行定期考查，这将对教育过程提供各种评价，尽管这种评价也许是有限的。

这些评价，以及用于进行这类评价的方法，正如本书以下各章所述及的，它是二十五年来把教育测量科学应用于医学科学的经验产物。由于这种经验得到了不断的发展，所以，我们既要考虑到全国医学考试委员会的考试服务，也要考虑到全体考试人员所做的工作，这样，我们总是要努力使具有医学和医学教育背景的人员与具有心理测量学背景的人员之间保持平衡。因此，本书第二版是由具有高度能力、富有经验的全体考试人员，以及附录 1. 所列举的全体高级人员共同作出努力的成果。虽然对参予以下各章所述考试程序的全体人员，对他们所起的作用应当予以充分的肯定，但对参予本书第二版某些章节撰写工作的作者更值得提一下。

全国医学考试委员会主席 Levit, E. J. 博士虽然没有担任本书任何章节的作者，但是，她为本书的编写曾做了许多工作，她那富有创造性的想象力，不倦的精力和在全体人员中的威望，同其他外部机构一起，在使医学考试委员会的服务继续站在美国医学教育评价最前列的创造性发展中，她曾发挥了重要的作用。当多选题考试在毕业生和大学生一级的考试中得到全面推广以后不久，Levit 博士又把临床能力的客观测量引进了全国医学考试委员会的第三部分考试，并成为全国医学考试委员会第三部分考试的一个重要特点，目前，这种测量技术亦为美国各专业委员会所采用。在担任全国医学考试委员会目标和重点委员会秘书期间，她是

题为“医学教育全过程的评价”^{*}报告的设计师,这是一份受到全国关注的报告。如同任何具有强有力报告的例子一样,虽然这一报告的某些方面曾受到了人们的批评,但是,从总的方面来说,这一报告所提出的建议,大家是赞成的。在这一报告中,对连续进行的、持续终身的教育全过程提出了一个比较明确的概念(教育全过程是这一报告所提出的非常大众化的术语),并要求为此建立相应的评价制度。

Chase, R. A. 自 1974 年 11 月至 1977 年 7 月任全国医学考试委员会主席期间,正当为提供高质量卫生服务而出现越来越大的压力的时候,他曾成功地处理了全国医学考试委员会在各项活动中所面临的问题。向公众负责成为全国医学考试委员会的一句格言,它在实施资格鉴定考试中的作用得到了进一步的发展,甚至要求使它成为一项法定的考试,医生是否具有充分的能力,应把这些要求写进联邦的法律(卫生职业教育援助法,1976 年)。Chase 博士在这一重要时期里曾担任全国医学考试委员会主席的职务,有关他的论述已被写进本书的第十一章,即“全国医学考试委员会和公众利益:相关的工作范围”这一总结性的章节里。

全国医学考试委员会副主席兼大学生评价科主任 Smith, D. E. 博士,他是作为一名具有医学教育背景的专业人员参加全国医学考试委员会工作的,他曾任弗吉尼亚大学的病理科教授和主任,随后又在美国医学专业委员会工作过一个时期,他非常熟悉全国医学考试委员会的业务。作为全国医学考试委员会决策者之一,他曾担任全国医学考试委员会常设委员会委员多年。因此,他具有充分的资格来指导

^{*} 医学教育全过程的评价 (Evaluation in the Continuum of Medical Education), 全国医学考试委员会目标和重点委员会 (William D. Mayer 任主席), 全国医学考试委员会, 费城, 1973。

全国医学考试委员会的各种考试方法在第一、第二和第三部分考试，以及由此派生出来的其他各种考试中的发展和应用（见第三、四和八章），更加广泛地说，按照他作为副主席所具有的权威，他可以利用他的丰富经验来对全国医学考试委员会不断发展中的政策和活动施加有力的影响。

大学生科协理主任（Co-Director）Kennedy, W. B. 博士，他曾任宾夕法尼亚大学医学院副院长，到全国医学考试委员会任职以来，他推动学院接受了全国医学考试委员会所主持的考试，并以此来代替由学院拟定的多选题考试。由于他具有在医学院任职的经历，所以，他在全国医学考试委员会所属的各考试委员会中工作，以及在控制标准化考试材料的使用中处于一种非常有利的地位。正如本书第八章所述，许多机构目前均利用全国医学考试委员会所保存的标准化考试材料来设计它们自己的考试。由于他非常熟悉全国医学考试委员会所储存的数量众多的试题集，因此他知道如何用于考试，本书附录里所列的示范性考试——这也是本书第一版最受人们欢迎的特点之一，他就是这项考试的设计者之一。

全国医学考试委员会副主席兼毕业生和继续医学教育科主任 Burg, F. D. 博士，以前曾一度在美国儿科委员会兼职，他作为全国医学考试委员会同毕业生和继续医学教育最积极和最有力的联系者是大家公认的，他与美国医学专业委员会，各专业委员会和各医学专科学会合作，与该科的副主任 Templeton, B. 博士一起，发展了全国医学考试委员会在正如本书第九章所论述的医学教育全过程评价中的作用。

心理测量学科协理主任 Schumacher, C. F. 博士，他曾在美国医学院协会的研究处任职，来全国医学考试委员会

后，使全国医学考试委员会在教育测量科学方面得到了必要的加强。在迅速发展的时期里，他不懈地坚持把心理测量学的原理同使研究成果更富有现实主义的固有观点结合起来，本书有关评分与分析，可靠性、有效性和标准的确定等章（第五章和第六章）就是由他撰写的，这也是他对本版最直接的贡献。但是，为保持全国医学考试委员会所主持的各种考试的精确性，他的同事和副主任 Kelley, P. P. 博士也为这些方面做了大量的工作。

全国医学考试委员会研究和发展科主任 Andrew, B. J. 博士，正如他本人在本书第十章中所论述的，全国医学考试委员会对教育测量的研究和发展已经给予并将继续给予重视。全国医学考试委员会的这一特点开始于对论述题考试和多选题考试的比较研究。为了测量应试者的临床能力，尔后又进一步探索更好的考试方法，同时也对各道试题的合适性进行了研究。在全国医学考试委员会所属的目标与重点委员会建议的推动下，成立了研究和发展科，以便加强对研究工作的领导，进一步寻找更好的测量各级医学教育的方法。

大学生医学评价科副主任 Weinberg, E. 博士，她担任卫生辅助专业评价方案主任，这一方案是全国医学考试委员会承担的职能之一，因为这类卫生专业人员已得到了全国的承认。因此，她对考试的有效性也产生了相应的兴趣，并为第六章有关考试有效性的讨论提供了实质性的贡献。

诚然，这些同道是 1977 年前我赖以使全国医学考试委员会的方针和考试程序达到现代水平的积极成员。但是，他们仅仅是对医学教育和医学考试具有丰富经验的全体人员中的一部分。在本书第二版的编写中，我们也得到全国医学考试委员会管理人员的大力支持。由于同道们的紧密配合，全国

医学考试委员会的成员已经成为国内外所熟悉的专家，为适应医学教育作出正确的和有意义的测量，国内和国际上都十分重视他们的领导才能。在这里，我对本书编写工作作出贡献的每一个人深表谢意。

最后，虽然对一名出版者来说通常也许不他为对本书的出版表示感谢。但是，我对我们在 Lea 和 Febiger 出版公司的编辑、前全国医学考试委员会职员 Pattison. R. 女士对本书所作的熟练的和有判断力的编辑表示感谢。

费城

Hubbard, J. P. 医学博士
(梅人朗)

目 录

第一章 考试和评价制度的发展	1
个人的评价.....	1
带徒培训制度.....	1
州执照颁发委员会.....	3
关于建立全国医学考试委员会的建议.....	6
多选法考试的采用.....	14
教育过程的评价.....	16
教育机构的认可.....	18
全国医学考试委员会评价工作的扩展.....	21
向专科委员会和专科学会提供考试服务.....	23
医助证书考试.....	24
考试服务的发展.....	25
第二章 考试委员会及其在建立客观考试中的任务	29
考试委员会的工作方法.....	29
考试大纲.....	31
综合性的多学科考试.....	33
考试的组织.....	35
第三章 多选题的类型	39
试题的三种基本类型.....	39
最佳回答型(A型题).....	39
配伍型(B型题和C型题).....	45
复合是非型(K型题和X型题).....	47

275599-2.00元-86/6

图表资料	49
审定试题的标准	52
第四章 临床能力的客观评价	55
临床能力的定义	55
病人处理问题 (PMP) 的程序考试	59
线性病人处理问题 (PMP) 的命题	63
确定问题和正确的病因	64
书写鉴别诊断	64
确定模拟的类型	64
确定评价的顺序	65
明确情况	65
书写题干	65
书写问题	66
对暗示的审查	66
病人处理问题的评分	66
第三部分考试不及格	69
第五章 评分与分析	72
评分	72
试题分析	74
试题难度(P值)	76
试题的区别性 (r_{bis})	77
答案有效化	80
成绩换算	81
全国医学考试委员会信任证书的严肃性	84
第六章 可靠性、有效性和标准的确定	88
可靠性	88
有效性	93

内容有效性·····	94
结构有效性·····	95
标准相关有效性·····	97
确定等级标准·····	99
第七章 向各医学院报告考试成绩·····	107
申请者和非申请者·····	107
过分相信全国医学考试委员会·····	108
第一部分考试的一份报告·····	109
第八章 利用标准化试题进行的资格鉴定考试·····	114
联合执照考试 (FLEX)·····	114
联合执照考试产生的背景·····	114
考试·····	116
联合执照考试和全国医学考试委员会考试	
双重考试制度·····	118
加拿大医学委员会考试·····	119
外国医科毕业生教育委员会 (ECFMG) 考试·····	120
签证资格鉴定考试 (VQE)·····	123
第九章 作为指导毕业生学习和继续教育的评价·····	126
对毕业生学习的评价·····	126
作为指导学习的住院医师训练期考试·····	128
继续医学教育的自我评价·····	131
发展考试对指导学习潜在作用的方法·····	133
学习提纲·····	134
关键词反馈·····	135
病人处理问题 (PMP) 的反馈·····	136
自学的自我评定和评价·····	137
继续医学教育——一个发展中的领域·····	139

第十章 研究和發展	142
历史的回顾.....	144
近期的研究工作.....	148
今后方向和趋势.....	154
第十一章 全国医学考试委员会和公众利益	
——相关的工作范围.....	161
有效性的策略.....	163
教育的全过程.....	166
对卫生、教育和福利部要求的响应.....	167
未来.....	169
高度科学的标准对最低程度知识的教育	
——“唯理智论对相关论”.....	170
专业上自我约束的立法动议.....	171
医疗执照的再注册.....	171
专科证书的再注册.....	172
影响公共(政府)卫生政策的立法动议.....	173
为专业能力评价承担教育的任务.....	173
资料、方法和证书国际交流的创始.....	174
小结.....	174
附录1. 全国医学考试委员会的高级工作人员	176
附录2. 全国医学考试委员会考试中所用的例题	178
附录3. 试题答案	248

第一章 考试和评价制度的发展

为了保证教育系统为医学同行和公众培养出合格的人才，美国医学界已陆续开展了三方面的重要工作，即(1)对作为教育成果的个人的评价、(2)对教育过程的评价和(3)对教育机构的评价。第一项工作是同由专业团体或立法机关批准个人行医的认可过程（发给医疗执照）密切相关的。第二项工作涉及对教育机构在完成教育目标过程中各种效果的评价。第三项工作是审查各教育机构是否符合设置标准，这些机构包括医学院校、教学医院、以及当前提出的同继续教育有关的团体和组织。

个人的评价

带徒培训制度

在殖民统治时代及其后相当长的时期里，美国的医生主要通过带徒培训来培养，他们也许以后到国外再接受某种形式的医学教育，受训者所学的知识和技能，主要在于向患者提供某种信心。如果同现代正规的医学教育和通过立法程序来管理行医的制度相比，我们当然不能认为带徒培训是一种完善的教育制度。但是，鉴于当前有大量的学生正在教育系统接受培训，而且为了加强学生同文科医科教师的密切联系而出现的各种困难，我们仍然不能否定带徒培训制度的价值。

Genevieve Miller 曾为我们列举了一个对学生个人进行评价并认为有能力在殖民统治下的美国行医的生动例

子。她引证了弗吉尼亚的一名医生，这一医生曾于 1775 年为他的徒弟写了如下的证书，这一证书是这样写的：

这些评语将向一切有关的人们报告：弗吉尼亚殖民地威廉县 Cary Henry Hempton 先生在我的教导下已经完成了带徒培训。在为期四年中，他已经勤奋地学完了解剖学、外科学、药物学和产科学；他对切西滕氏 (Cheseldens) 解剖学、海斯特氏 (Heisters) 外科学，考伦氏 (Cullens) 药物学、斯曼利氏 (Smeiles) 产科学已具有良好的基础；他已经用拉丁语学完了先师悉登汉姆 (Sydenham) 和希波克拉底 (Hippocrates) 的著作，以及同本专业有关的许多其他著作。在培训期间，他已在病人床边和其他地方受过实践训练。因此，我对他所掌握的知识表示满意，并推荐他向一切需要他提供服务的人们提供医疗照顾。

医学博士 Andrew Robertson (签字)

1775 年 8 月 1 日

当时，这一证书就是青年医生 Hampton 的一分医疗开业“执照”。当然，这一证件决不是随手可得的，只有导师有权向他的学生提供这样的合法证书，正如各类工匠一样，只有他的师傅才能在中世纪时代行会职业训练中向他的学徒传授技艺。在马萨诸塞州的一份 1734 年的学徒契约中规定，在为期五年又四个月的培训期中，导师必须向他的学生提供膳宿，同时，对学生的任务也作了如下的规定：

在受训期间，学生应当忠实地为他的导师服务，要严守秘密，任何情况下都要乐意地完成合理和合法的指示，决不有意做有损导师声誉的事情……，学生不得经常到酒馆肆意酗酒，玩纸牌或聚赌，也不准参加任何非法的娱