

医师资格考试内容效度的质量控制措施

何惧, 何佳, 张赛一, 弓姝

国家医学考试中心, 北京, 100097

摘要: 医师资格考试是依据《中华人民共和国执业医师法》举行的准入性行业资格考试, 在规范医疗卫生行业管理、提高医疗卫生服务水平、保障公众身体健康方面发挥重要的促进作用。为提高该考试的科学性和公信力, 需依据现代心理测量理论, 在考试设计、题库建设、审题组卷、考试组织实施、阅卷评分、数据分析等考试的各个阶段采取一系列质量控制措施, 以提高考试的内容效度, 有效发挥考试的社会学作用。

关键词: 医师资格考试, 内容效度, 质量控制措施

Introduction of Quality Control Measures to Improve the Content Validity of National Medical Licensing Examination

HE Ju, HE Jia, ZHANG Sai-yi, GONG Shu

National Medical Examination Center, Beijing 100097, China

Abstract: National medical licensing examination is a vocational qualification test carried out following *Law of the People's Republic of China on Medical Practitioners*, which has played an important role in regulating management of medical behavior and improving the quality of medical service and people's health. In order to improve the scientific nature and validity of national medical licensing examination, various quality control measures should be taken to improve content validity in the process of designing test blueprint, constructing item bank, writing test items, composing test paper, test administration, scoring and data analysis.

Key words: National medical licensing examination, Content validity, Quality control measures

1 背景

医师资格考试是《中华人民共和国执业医师法》规定的卫生行业准入性考试, 考试成绩是考生获得在医疗、预防、保健机构中执业资格的主要依据, 考试结果对考生、卫生机构乃至整个社会都会产生重大的

影响。由于以考试分数为依据对考生进行执业注册, 因此要求考试结果具有高度的有效性——效度。美国《教育与心理测试标准》解释: “效度指根据指定用途支持分数解释的那些事实和理论的有效程度。因此, 效度也是测验制作和评价中最基本的要素。”^[1] 内容效度指考试内容对于应测内容的代表性程度。医师资格考试属于大规模、高利害的标准参照考试, 由于受到考试时间、场地和考试手段的限制, 考试测量只能是抽样测量, 其内容效度取决于多个因素, 对考试内容效度的审查和验证是一个持续不断的收集证据的过程,

该过程起始于考试设计，与考试大纲的制定、组卷方案的确定和题库的建设同步进行，并贯穿整个考试实施、阅卷评分、数据分析、第三方评价、专题研究等考试的各个阶段，指导分数使用部门利用考试成绩进行正确的决策，有效发挥考试的社会学作用。本文介绍了医师资格考试在考试设计、试题开发、组织实施、阅卷评分各个阶段所采取的多种提高考试内容效度的质量控制措施，探讨如何在考前、考中和考后采取有效技术措施改进考试，不断提高考试的内容效度和社会公信力。

2 执业医师基本要求简介

自2009年起，我国医师资格考试委员会颁发的《医师资格考试大纲》（以下简称《考试大纲》）^[2]在前言部分对临床、口腔和公共卫生类别执业医师和助理医师提出执业基本要求，包括专业、学历及工作经历、基本素质、基础理论、基本知识和基本技能的具体要求，这是对考试测量目标的准确描述。首先，该基本要求在理论上确定了考试的测量目标，这是任何一个考试研究和试题开发必须考虑的首要问题。其次，它确定了考试的内容范围和选择试题背景材料的基本出发点。因此，该基本要求对医师资格考试的考试设计和试题开发工作具有决定性的影响，它也被作为专家进行考试内容效度审查的基本依据。例如，在考试设计阶段由专家反复分析哪些基本要求是可以测量的，哪些目前的考试形式尚无法测量；在试卷审查阶段，通过对试卷结构技术指标的分析，来验证哪些测量目标已经在试卷中得到体现，哪些尚未体现，其代表性如何；在考后试题质量分析阶段，结合当年的实测数据分析哪些测量目标已经实现，哪些目标的测量效果不佳。

3 考试设计阶段采取的措施

3.1 考试大纲和组卷方案的制定、审查和完善

《考试大纲》依据执业医师基本要求制定，明确告诉考生医师资格考试将具体测量什么。同时，还需根据考试大纲制定组卷方案，明确每一个测量目标相应的考察目标和标准，即制定考试能够实现的知识内容的层次和能力的考查要求。组卷方案由一系列表格组成（即《组卷细目表》），反映了试卷的内容、难度、

题型、认知层次分布等关键要素之间的关系。在考试大纲和组卷方案的编制、审查与完善中，质量控制措施如下：

（1）课题研究法。采用课题招标的方式，借助高等医学院校和专业教育测量机构的力量开展专题研究，如国内外医师执照考试比较研究、执业医师的岗位胜任力研究、医学考试分类及其考试设计研究等，为考试大纲和组卷方案的制定提供科学依据。

（2）实地调研法。组建调研小组多次赴各类高等医学院校、各级医院（含口腔医院）、预防保健机构和各级卫生行政部门进行实地调研，充分了解各个层次的办学机构、用人机构和政策执行情况，充分了解医学教育情况、用人单位需求以及对考试的意见和建议。

（3）专家小组讨论法。遴选具有代表性的专家参与讨论，除了依靠试题开发专家委员会委员，还在全国范围内选取具有地域、院校、机构代表性的权威医学专家，在多轮次讨论的基础上编写《考试大纲》，并在心理测量专家的参与下完成《组卷细目表》的编制工作。

（4）问卷调查法。根据每年考试实测情况，分析考试大纲和组卷方案是否准确体现了考试的目标，制定针对医学院校、各级各类医疗卫生机构、卫生行政部门和考生的调查问卷，收集各考试利益相关方的意见和建议，不断进行总结、研讨和改进。

3.2 专家小组审查时需重点考虑的几个问题

3.2.1 测试内容是否和测试目标一致

医师资格考试每一年的试卷都严格按照《组卷细目表》编制，由此可见，测试内容是否符合测试目标的关键点在于《组卷细目表》的科学性。专家小组审查需重点考虑《组卷细目表》能否准确诠释测试目标，能否正确体现所要测量内容的知识和能力要求范围。

3.2.2 题型的测试功能是否与能力考查目标一致

题型的选择服从于考试目的，不同题型测试的功能也不尽相同。鉴于医学实践性强的学科特点，考试的测量目标是考查考生运用医学基础知识和基本技能解决临床、口腔、公共卫生工作实际问题的能力。医师资格考试分为实践技能考试和医学综合笔试两部分，两部分均合格才能通过医师资格考试。实践技能考试的题型包括案例分析题和操作题两种，考查较为复杂

的能力，如应用、分析、综合、探究和动手操作的能力。医学综合笔试全部采用客观题，考查记忆、理解和应用能力。通过多种题型和考试形式的组合，充分发挥各类题型的优点，实现考试目标。

此外，还需要审查《组卷细目表》中的难度比例、认知层次比例、测试长度（题量）、考试时间设定等与考试测量的行为目标是否契合。

4 试题开发阶段采取的措施

4.1 针对试题内容代表性的审查

试题的代表性越好，整份试卷的内容效度就越高。为提高每一道试题对测量目标的代表性，采取的控制措施有：

(1) 制定试题命制规范。编写规范的《试题编写指南》和严格的试题入库标准，作为指导专家命题和审校试题的依据和标准。

(2) 专家培训制度。为帮助专家准确领会命题要求，每年组织专家集中进行命题技术培训，结合实际例题讲解命题原则，反馈往年实测试题参数，现场试命题并集中讨论，提高专家命题技巧。

(3) 试题审查制度。每年数次召开审题会，对题库内的试题进行打磨。组织专家审题时应重点考虑：试题内容是否与大纲规定的知识点相吻合；试题是否超出了大纲规定的评价范围；内容取样是否有较好的代表性。同时，审题过程还需审查试题内容是否科学、正确；文字叙述是否简明扼要；备选答案之间是否有相互包容、相互重叠、相互依赖的内容。专家和工作人员还需事先审查每一道试题是否存在对不同群体（地域、种族、性别、文化、宗教）的公平性问题。

4.2 针对试卷内容代表性的审查

《考试大纲》规定了考试要测量的行为总体，组卷方案则是一份试卷所要考查的知识、能力、技能的具体内容及其比例，是进行有效抽样的基础。试卷根据组卷方案组成，是对上述总体的抽样，抽样的代表性直接表征了试卷的内容效度。因此，对试卷的审查重点是试卷是否符合组卷方案。该环节的质量控制措施有：

(1) 内部审查制度。命题专家按照命题规范编制的试题经过数次审校达标后进入题库系统，然后由专业秘书严格按照《组卷细目表》，从题库中抽取符合要

求的试题组成试卷，并由考试机构的工作人员对试卷进行自审，确保试卷的内容、题型、难度和认知层次构成符合组卷方案。

(2) 专家审卷制度。由于不同的学科专家对于考试测量目标的界定、解释和选择总会存在一定的偏差，即使审题专家团队认为题库内的每一道试题都具有较好的内容代表性，组好的试卷中仍然有可能存在部分达不到组卷方案中原定测量目标的试题。因此，在审卷过程中尽量采用独立于命题过程的专家小组进行试卷审查：一是试卷是否对考察的重要方面进行全面的考察；二是试卷是否包含与所要考察的内容无关的试题，确保整套试卷具有较广的覆盖面，其内容代表性符合组卷方案的要求；三是对试题内容的科学性、试题间是否存在相互提示现象、试卷难度和认知层次的审查以及试卷形式、平行试卷的审查等。审卷过程发现问题时，需对质量不高的试题进行淘汰、更换或修改，以满足组卷要求。

5 考试结束后采取的措施

每年考试结束后，通过对试题、试卷的质量的定量分析，审查试卷是否完成了考察目标，试卷的测试内容是否适合测试对象。

(1) 专家试题复审机制。每年考试结束后，根据实测参数进行试题、试卷质量分析，对参数不佳的试题进行复审，由专家逐题评估是否偏离了原定的考试内容、范围，并将复审结果反馈给试题开发专家，以便改进工作。

(2) 试题试卷质量评价分析制度。为了解测试内容是否适合测试对象，首先，考试结束后将考生依学历、专业、工作单位机构类别、隶属级别、地区分布、年龄等分成不同的群体，分别进行考生构成和成绩差异分析，了解不同层次考生在考试中的表现。其次，按学历、毕业专业分别统计每道题的试题参数，以便逐题分析考试内容是否适合于考试对象。第三，通过试题功能偏差分析找出对不同考生群体存在测量偏差的试题，反馈给专家，以便在命题、审题时尽可能降低考试背景材料对考生表现的影响，不会有部分考生因此受益，克服试题功能偏差和考试不公平。

(3) 第三方评价。考生结束后，定期采取问卷调查、座谈会等方式定量、定性地了解专家、考生对考试内容代表性的评价。调查问卷的发放主要针对未参与命题、审题工作的卫生行业专家和考生，问卷调查

的主要目的是定量了解他们对考试内容与执业医师实际工作契合度、考试内容与执业内容的相关性的评价，了解考生内容在实际工作中的使用频次和重要性，从而科学评价试卷的内容效度。

6 结论与思考

可靠的测量结果来源于高质量的试卷，为提高内容效度，在医师资格考试过程中采取了一系列的质量控制措施，建立、健全了各项工作规程，保证了每一年的试卷具有较好的代表性，符合测量目标，适合测试对象。

但是，考试的内容效度取决于多种因素，在今后的考试设计和试题开发过程中，还有一些难点、重点问题有待突破：（1）科学修订执业医师基本要求和《考试大纲》。按照当前医药卫生体制改革的要求，结合目前面临的医疗卫生专业实践中的突出问题，及时修订、完善以上文件。（2）合理设定试卷结构指标，不断完善《组卷细目表》。《组卷细目表》是保证医师

资格考试设计科学性、合理性、规范性和考试结果准确性的依据，也是提高考试效能，确保考试生命力和公信力的关键。（3）加强考试技术的研究。受限于考试形式、题型、时间、场地、经费等因素，在基本要求中提到的有些能力目前尚无法测量，还需加大考试研究方面的投入，探讨新的题型和考试方式。

医师资格考试的目的和性质要求考试起点公平、过程公正、结果可信有效。要证明考试的有效性，需要不断收集基于考试内容、基于内部结构、基于考生反应过程、基于跟其他变量的关系、基于测验后果的效度证据^[1]，这将是一个长期的过程。

参考文献

- [1] 美国教育研究协会，美国心理学协会，全美教育测量学会．教育与心理测试标准 [M]．燕妮琴，谢小庆，译．沈阳：沈阳出版社，2003.
- [2] 卫生部医师资格考试委员会．医师资格考试大纲 [M]．北京：人民卫生出版社，2013.