

教学论文



蔡红星，中共党员，教授、副主任法医师，徐州医科大学副校长。担任教育部高等学校法医学类专业教学指导委员会委员，首届教育部高等医药院校现代教育技术与计算机教学指导委员会委员，中国健康医学教育联盟副理事长，江苏省司法鉴定协会副理事长，江苏省高等学校医药教育研究会常务理事，徐州司法鉴定协会副理事长兼秘书长。主要研究方向为损伤、疾病及死亡关系的法医学鉴定。在全国卫生职业院校中首创建了司法鉴定技术专业，并成为该专业的学科带头人。近年主编教材1部，主持省级以上课题5项，获江苏省“六大人才高峰”第七批次项目资助，2017年获省高等教学成果奖二等奖1项、徐州市科技进步奖一等奖1项，发表SCI论文8篇，中文核心期刊论文21篇。

强化终结性评价内涵，全面检验医学培养目标达成度

乔伟丽¹，孙红¹，蔡红星^{2,*}

徐州医科大学¹基础医学院生理学教研室；²教务处，徐州 221004

摘要：终结性评价在教育评价体系中具有“一锤定音”的决定性作用，是衡量教育目标是否达成的标尺。社会的发展不断对人才培养提出更高的要求。目前医学院校采用的“一纸定终身”的终结性评价已经无法满足当前医学教育培养目标的需求，强化终结性评价内涵建设势在必行。通过丰富评价方式、优化试卷内容、规范评价标准、提升教师评价技能、增加学生作为评价主体等多项策略，可以较好地完善终结性评价的标尺功能，全面检验医学培养目标达成度。

关键词：高等医学教育；终结性评价；培养目标

中图分类号：G640；G642

Optimization of medical educational objectives by improving summative assessment system

QIAO Wei-Li¹, SUN Hong¹, CAI Hong-Xing^{2,*}

¹Department of Physiology, Basic Medical College; ²Academic Affairs Office, Xuzhou Medical University, Xuzhou 221004, China

*Corresponding author: xzmccaih@163.com

Abstract: Summative assessment plays a decisive role in the educational assessment system, which is a yardstick to measure the cultivating goal of higher education. The rapid progress of modern society has put forward higher standard for higher medical education. Traditional summative assessment system with single dimension that focuses on evaluating the student's learning outcome via a standardized examination cannot meet the higher requirements for undergraduate medical education. We have improved the summative assessment system by optimizing the assessment content, criteria and method, as well as teachers' assessment skills and students' evaluation. The reform greatly increases the teaching quality and learning effect in our university.

Key words: higher medical education; summative evaluation; educational objective

*通讯作者：蔡红星，Tel: +86-516-83262017, E-mail: xzmccaih@163.com

习近平总书记在2018年全国教育大会上指出,要深化教育体制改革,健全立德树人机制,扭转不科学的教育评价导向,从根本上解决教育评价指挥棒问题。社会的发展不断对人才提出更高的要求。今天,医学人才不仅要具备丰富的医学知识和娴熟的专业技能,而且要具备高尚的职业道德、良好的表达沟通能力和学习创新能力^[1]。因此,为了适应现代医学培养目标的需要,高等医学教育在课程整合、教育方法和技术手段等方面不断进行改革^[2]。而教学和教学改革是否成功,需要用教育评价体系来衡量其是否达成教育培养的目标。本文重点剖析了当前我国医学教育评价现状,认为医学院校目前采用的“一纸定终身”的终结性评价已经无法满足当前医学教育培养目标的需求^[3],提出强化终结性评价内涵建设的需求,使之能真实、全面评价教学和学习效果,从而实现全面检验医学培养目标达成度的目的。

1 教育评价——终结性评价和形成性评价

教育评价是根据一定的教育价值观或教育目标,运用可行的科学手段,对教育活动、教育过程和教育结果进行价值判断的过程。根据评价的功能和用途不同,教育评价可划分为终结性评价和形成性评价。

终结性评价是在某一相对完整的教育阶段结束后,对整个教育目标实现的程度做出的评价,也称为结果性评价。它以预先设定的教育目标为基准,考察学生达成目标的程度^[4]。在现行的教育评价体系中,终结性评价占有决定性地位。医学生的期中、期末考试、三阶段考试以及毕业考试、执业医师资格考试都是属于终结性评价。也就是说,课程能否通过、能否升级、能否拿奖学金、能否毕业、能否读研、是否能做医生都由终结性评价来决定,所以有“分,分,学生的命根!”一说。

形成性评价是指在某项教学活动的过程中,为使活动效果更好而不断修正其本身轨道所进行的评价,也称为过程性评价^[5]。它的主要目的是为了明确教学活动中存在的问题和改进的方向,及时修改或调整教学计划,以期获得更加理想的效果^[6,7]。形成性评价是在终结性评价的时间段内进行的评价,起反馈调节教学活动的作用。

终结性评价和形成性评价的关系是你中有我、我中有你。如果说教学是一场长跑,终点是我们的

培养目标,到达终点的时间(结果)就是终结性评价,如何从起点到终点(过程)就是形成性评价。学生是运动员,教师是教练员和裁判员。到达终点的时间由过程中运动员的速度、技巧等决定,也就是过程决定结果,结果是过程的体现。教练员在跑步过程中的指导和帮助,能够及时修正运动员的不足,使结果更好;双方从结果中分析得失可以指导下次的过程。跑步结果不好可以重来,而跑步过程中运动员学到的方法和技能却可以受用终生,从这个角度来说,教育中的形成性评价意义更大。但现实生活中,人们都不希望重来,都希望有一个好的终结性评价结果。因此,教学过程中如何协调终结性评价和形成性评价的关系、如何做出客观公正的终结性评价使之符合培养目标的要求就成为至关重要的问题。

2 当前我国医学教育评价现状及分析

中国教育部临床医学专业认证工作委员会于2016年公布的《中国本科医学教育标准——临床医学专业(2016版)》^[8]规定:中国临床医学本科专业毕业生应在科学和学术、临床能力、健康与社会、职业素养四个领域达到基本要求。学生学业成绩考核要围绕培养目标制定和实施,确保考核覆盖培养目标的各个方面,确保学生能够实现预期的教育结果并有利于促进学生的学习。

根据上述标准考量,笔者认为我国医学院校的终结性评价存在以下不足。

2.1 评价方式单一,导致评价出现重视结果而忽视过程的现象

许多学校一直存在一张试卷成绩等于终结性评价的“一纸定终身”现象,缺乏对学习过程的评价指标。一门课程一般是一次性期末考试,整个本科过程中还有三阶段考试和毕业考试,都是闭卷笔试。试卷中知识点的考察偏重知识的记忆,缺乏对知识点的理解、分析及综合应用知识能力的考核。另外,一次性考试不利于调动学生平时学习的积极性。

2.2 评价内容不全面,重视定量评价而缺乏定性评价

医学是一门实践性很强的科学,一张试卷形式的终结性评价导致在理论考试中不易体现的技能、态度、职业素质方面的培养结果往往被忽视。评价缺乏基于学生的实验、实习的表现、观察、讨论等方面的定性评论^[9]。尤其是在临床实习阶段,带教教师缺乏对实习生的职业作风、对患者的态度、工

作责任心等行为的评价，甚至于学生的出科考核也只有试卷成绩，严重影响了临床实习的质量。

2.3 评价标准不够规范

目前我国许多医学院校正在尝试把形成性评价的信息纳入终结性评价，但工作刚刚起步，对于如何评价缺乏统一的标准，导致结果因人而异，差异很大。

2.4 评价缺乏专业特色

目前医学院校专业分化，每个专业都有各自的培养目标，例如临床医学专业和公共卫生管理专业就业方向差异巨大，同样的生理学课程，两个专业的教学和考核都要有所区别。但医学院校教师大多关注自身学科的发展，对本学科在各专业整体培养目标中的定位认识不足，尤其是一门课程同时对多个专业学生授课时情况更为严重。教师往往只能根据课时的多少增减教学内容和考核内容。

2.5 评价主体单一

传统的终结性评价由教师做出，学生作为评价对象完全是处于被动的一方，学生只能接受评价而无权参与评价，这在很大程度上忽视了学生的主观能动性。

显然，目前医学院校实行的终结性评价与上述国家考核标准之间还是存在较大的差距，不能满足医学培养目标的要求，其结果必然使得学生普遍重视知识学习而轻视能力提升，严重影响了毕业生质量^[10, 11]。近年来，大批学者与教师开始关注医学生的终结性评价，就什么是有效的终结性评价以及如何强化终结性评价内涵进行尝试和探索。

3 有效的终结性评价应具备的指标

终结性评价的首要目的是对学生在某门课程或某个学习阶段所取得的学习成果进行全面的结论性评定，其结果主要是为教务管理者、教师以及学生提供教学效果、学习成绩信息，为未来的教育管理、教学以及学习调整提供依据。教学目标的实现要通过终结性评价来高度综合、概括、分析和总结。相较于教学过程中的个体，课程目标达成的整体情况和普遍性受到更多的关注。因此，为了真实、客观、全面地反映教学、学习状况，有效的终结性评价应该保证具有有效性、可靠性、冲击性和可行性^[12]。

3.1 有效性

有效性，即效度，是指测试或评价在多大程度上考核了教学大纲所需考核的内容，评价的内容、

方式有多大程度体现教学目标。任何一种评价的好坏，最重要的就是看有效性是否实现^[13]。尽管考核或评价方式不同，标准也各有差异，但标准的制定都依据一定的科学规律。一般来说，保险的做法是参照教学大纲和课程标准，将考核内容与教学大纲所列内容进行比照，从中筛选出有代表性的知识点、核心技能和能力等列入考核内容。考核内容与教学大纲所列内容一致性越高，则有效性越高。对于医学教育来说，评价工具所采用的成绩测试应该是与教学大纲紧密相扣的标准参照测试，学生的成绩取决于与预先设定的标准是否相符。学生所获得的成绩或评价只针对他自身学习表现与课程标准的差距，而不是该学生与其他学生的差距。也就是说，可以所有同学都达标或所有同学都取得最高成绩。这种评价方式能增加学生的学习动力，减少源自同学的人为压力。

3.2 可靠性

可靠性，即信度，指的是采取同样的方法对同一对象重复进行测量时，其所得结果一致的程度，也就是评价工具和结果的可靠性、稳定性和客观性。可靠性可以从两方面考量，一是不同的教师评价同一个测试，结果应该大致相同，二是同一个教师在不同的时间评价同一个测试，结果也应大致相同。实际工作中可以通过规范评分标准或评价量表、统一评分程序等，尽量提高可靠性。

3.3 冲击性

评价不仅仅是衡量表现情况，而且还要对教学和学生的学习动机等起到促进作用。评价结果通常对课程的开设和教学有着强烈的冲击。因此，尽量减少负面影响是非常重要的。

3.4 可行性

进行终结性评价所需要的资源，包括评价内容、评价标准、评价方式、教师的时间和专业知识、费用以及学生的学习时间等，都应该具备。

4 强化终结性评价内涵的策略

医学教育的目的是培养符合新形势要求的合格的医生或医务工作者，而终结性评价旨在判断学生是否阶段性或最终达到此目标。这种以结果为导向的考核理念也被称为基于能力的评价，评价所涉及的评价内容、评价标准、评价方法和实践等均需符合预期目标。综合国内多所医学院校和我校的实践探索，我们提出强化终结性评价内涵的具体策略，

以达到通过教育评价来提高教学质量、促进教育目标达成的目的。

4.1 增加考核形式, 兼顾结果与过程、定量与定性

课程的考核形式中包括定量考核(如随机测试、阶段考试、期末考试、毕业考试)和定性考核[如基于问题的学习(problem-based learning, PBL)评价、基于案例的学习(case-based learning, CBL)评价、实践操作考核、撰写论文]^[14]。不同的考核方式达成不同的目的。

医学生临床期还需要有特有的考核方式, 如临床操作观察和考核、微型临床演练、客观结构化临床考试、临床虚拟模拟考核等。我校设立了临床多站考试中心, 包括数字化临床技能实验室(培训中心)、试题库和网络平台。利用该考试中心作为出科考试和客观结构化考试的考场, 丰富了临床期终结性考核的形式, 多种先进的考核方法立体综合地考察了学生对于知识和技能的掌握情况^[15]。

在实际应用中, 上述考核方式可以根据专业和课程特点灵活掌握, 选择合适的组合。如在理论性强的医学基础课程可以适当上调定量考核的比例, 而在应用性较强的课程中可以适当增加定性考核的比例。

定量和定性评价结果按照一定的比例组成学生的最终成绩。以我校生理学课程为例, 学生最终的生理学课程成绩由平时成绩、期中成绩和期末成绩三部分构成。平时成绩包括考勤、课堂表现、随堂测验、单元测试、实验操作、实验报告、CBL评价, 各占一定比例(表1)。上述比例根据不同专业特点有所调整。这样, 作为终结性评价的成绩中兼顾了学习过程中的表现, 使得终结性评价的结构更加合理, 结果也更加客观、全面。

多种评价方式的使用, 无疑会占用教师和学生大量的时间和精力。因此, 考核方式和次数要适当,

一是避免让学生记忆过量的信息或占用学生大量的时间, 对学习产生负面影响, 二是不要过量增加教师负担。如果教师和学生能够共同商定课程考核标准和考核方式组合, 对学生的积极性和主动性都有很大帮助。

当前互联网的发达和大范围线上课程的应用, 为多种考核方式的实施提供了极大的便利^[16]。在实施中, 我们利用学习通和在线课程进行考勤、课堂提问、随堂测验和单元测试, 教师设置好要求, 系统自动对学生的各方面表现进行分析和打分。一方面教师可以随时、全面掌握全体学生的学习情况, 也可以对单个学生个体进行分析^[17], 从而及时调整教学进度和教学策略; 另一方面, 网络的自动分析和统计功能为教师节省了大量时间和精力, 这对学生数量大但教师数量相对较少的高校尤为重要。

4.2 优化试卷内容

虽然很多学校都在优化终结性评价的构成, 但试卷考试的占比一直很大。优化考试内容要注意两点: 一是突出对学生理解、分析及综合应用能力的考核; 二是密切联系学生专业特点^[18]。

在知识与能力并重的思想指导下, 考核从强调记忆和描述转为专业能力和综合素质的考查, 命题方式也相应改变。例如, 对于知识点“下丘脑的视上核和室旁核分泌抗利尿激素”, 以前的命题方式是“下丘脑的视上核和室旁核能够分泌下列哪种激素?”, 注重记忆; 如今改为“患者, 女, 27岁, 3天前骑电动车下班途中遭遇车祸, 致头部受伤, 被路人送医院急救, 病情稳定后转入重症监护病房。护士发现病人口唇干燥、尿量增多, 2小时尿量1550 mL。体温37.2℃, 血压82/63 mmHg, 脉搏95次/分, 呼吸18次/分。该患者的临床表现最可能与下列哪种激素有关? ”。同样的知识点, 命题方式的改变导致评价内容从单纯的知识考核上升为

表1. 徐州医科大学2019年临床医学专业生理学课程考核方法

成绩构成	成绩比例(%)	考核方法
考勤/课堂表现	5	手机签到/随堂观察/教师打分
随堂测验	5	教师根据课程内容组织测验, 5~10次
实验操作/实验设计/实验报告	15	实验课随堂观察, 教师评价打分和学生自评打分
在线课程	20	利用省在线开放课程, 完成在线课程规定的各项任务(视频、章节作业、单元测验、在线讨论等)
CBL	5	教师评价/学生自评/学生互评
期中考试	10	卷面考试, 时长1小时
期末考试	40	卷面考试, 时长2小时

CBL: case-based learning.

知识与能力并重，同时还兼顾课程思政。

考试内容也要与学生专业特点相联系。针对不同专业，例如临床、影像、麻醉、护理及药学等，同一门课程的教学内容和考核内容应有所区别，各有侧重，掌握好考核的难易度，以符合本专业人才培养目标的要求^[14, 19]。目前多所高校实行的课程负责人制度很好地弥补了这方面的不足。

4.3 规范化评价标准

规范的评价标准能够保证评价的可靠性。有标准答案的考核，如选择、名词解释、填空、问答等客观性试题，评价标准很容易规范。但对于没有标准答案的考核，如非标准化考核和形成性评价信息，规范可行的评价标准的制定至关重要^[20]。评价标准的设立，必须考虑到考核目的。由于非标准化考核更加注重学生对知识的运用，在考核模式设计时，要根据学科特点合理规划，试题注重发散和创造。在最终成绩评定时，考核标准的重点应是学生是否能结合课程基础知识进行发散性思考，提出有理有据的答案；是否对知识融会贯通，是否有清晰的解题思路；是否能够举一反三、推陈出新。考核试题相应的参考答案则应为开放式和非标准化答案。对于过程性和能力评价的信息，常用的做法是制定评价量表、统一评分程序，把收集的信息进行分级和水平判定，之后把不同等级量化为数字成绩，按比例记入总分^[21, 22]。当然，我们没有必要、也不可能把所有评价信息都进行量化，只有那些具有普遍意义并能够进行分级量化处理的信息，如课堂出勤、课堂展示、随堂测验、课后作业、实验操作及报告、课程论文、答辩等，才进行量化。

4.4 强化教师的评价能力

终结性评价的主体是教师，作为裁判员的教师的评价技能和水平很大程度上影响了评价的有效性和可靠性。基于这个原因，有效的终结性评价对教师提出了更高的要求。首先，教师应该接受相应的岗前和在职培训^[23]，理解评价目的、掌握评价技能、有比较完善的评价方案，足够了解所教授课程在专业人才培养目标中的定位，在考虑全方位学习目标的基础上做出判断，减少评价误差。其次，评价程序应该是非常透明的，教师的判断要以事实为依据。

4.5 增加学生作为评价主体

目前，我们正在尝试把学生本人和同学纳入评价主体，让学生也尝试担当裁判员的角色。一是让学生参与评教。学生对教师的教學态度、教学内容、

教学方法、教学效果等方面进行评价，并在分析评价结果与学生意见的基础上向教师进行信息反馈^[24]。二是组织学生以个人或小组为单位参与出题和阅卷，已经进行了两年。学生通过认真阅读教材和教学大纲、上课听讲、上网查资料等方式发现重点和难点，创造性地把各个知识点转化成考题。为了不过度增加学生负担，每个学生只负责一小部分内容。小型的测试全部选用学生出的题目，学生分组阅卷评分。优秀的题目纳入期末考试题库。出题和阅卷的过程就是自主性学习和研究性学习的过程，对学生的创造性思维和批判性思维很有益处。

综上所述，终结性评价内涵的强化，对教学质量、教学实施情况、学生一个学段整体学习状况给出鉴定性评价，从而为学校、教师和学生群体在未来教学中更新教学内容，改进教学方法，提高学习质量等下一步教学决策提供重要依据。其评价结果相对稳定、客观、可信度高，具有评定学生知识技能掌握程度、预测学生在后继学习过程中成功的可能性等作用，兼顾了过程和结果，成为一种客观、全方位、多元化的综合性评价方式，让教师和学生从评价中真正有收获、有成长，为全面达成医学培养目标提供了保障。

参考文献

- 1 陈佳龙. 新时期医学生培养目标的定位和教育模式的思考. 西北医学教育 2014; (6): 1065-1066, 1076.
- 2 吕立夏, 姜文霞, 邵志华, 李玉娟, 曹金凤, 蔡巧玲, 杨文卓, 徐磊, 徐国彤, 郑加麟, 章小清. 临床医学专业基础医学课程体系的建设和实践. 中华医学教育杂志 2018; (5): 665-667, 675.
- 3 赵炜明, 张晓杰, 云长海, 孙宝志. 医学教育全球标准中的学生考核新理念探析——以美国名校为例. 医学与哲学 2020; 41(8): 56-59.
- 4 Kibble JD. Best practices in summative assessment. Adv Physiol Educ 2017; 41(1): 110-119.
- 5 Prashanti E, Ramnarayan K. Ten maxims of formative assessment. Adv Physiol Educ 2019; 43(2): 99-102.
- 6 Kulasegaram K, Rangachari PK. Beyond “formative”: assessments to enrich student learning. Adv Physiol Educ 2018; 42(1): 5-14.
- 7 Evans DJ, Zeun P, Stanier RA. Motivating student learning using a formative assessment journey. J Anat 2014; 224(3): 296-303.
- 8 教育部临床医学专业认证工作委员会秘书处. 中国本科医学教育标准——临床医学专业(2016版). 北京: 北京大学医学出版社, 2017.

- 9 赵莲, 沈守荣, 邓芳, 彭小青, 王晓艳. 基于多元智能理论的八年制医学教育评价体系研究. *中国继续医学教育* 2017; 9(21): 43–45.
- 10 张付芝, 刘莹. 形成性考核在高等医学教育中的应用研究. *卫生职业教育* 2013; 31(24): 5–7.
- 11 刘莹, 崔涛, 刘磊. 形成性考核对提高学生病理学学习质量的作用研究. *继续医学教育* 2010; 24(5): 29–31.
- 12 Cilliers FJ, Schuwirth LW, Herman N, Adendorff HJ, van der Vleuten CP. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract* 2012; 17(1): 39–53.
- 13 Chur-Hansen A, Vernon-Roberts J. The evaluation of undergraduate students' written English language skills. *Med Educ* 2000; 34(8): 642–647.
- 14 Cramer N, Asmar A, Gorman L, Gros B, Harris D, Howard T, Hussain M, Salazar S, Kibble JD. Application of a utility analysis to evaluate a novel assessment tool for clinically oriented physiology and pharmacology. *Adv Physiol Educ* 2016; 40(3): 304–312.
- 15 李艳杰, 王志荣, 黄一虹, 徐晤, 李洪春, 韩斐斐, 孟冰. 由“多站式考核”设想五年制医学生临床技能培训体系的构建. *卫生职业教育* 2013; 31(6): 102–103.
- 16 刘勤勇, 刘晓燕. 教育信息化2.0时代医学教学评价变革. *解放军医院管理杂志* 2019; 26(12): 1127–1130.
- 17 潘晨, 刘莹, 蔡红星. 医学教育认证背景下学生评价改革初探. *西北医学教育* 2016; 24(1): 30–31, 54.
- 18 朱永香, 肖赞英, 王倩, 于玮, 南瑛. 生理学考试改革的探索. *中华医学教育探索杂志* 2010; 9(6): 763–764.
- 19 陈建芳, 刘友, 任可, 梅建军, 刘晓梅, 刘锦龙, 张涛, 于桂霞. 以临床执业医师考试为导向的医学免疫学教学改革初探. *基础医学教育* 2020; 22(3): 173–175.
- 20 易祖木, 李春雨, 赵呈智, 邓子昂, 张雨薇. 非标准化答案考核在医学教育中的探索和应用. *中国医学教育技术* 2019; 33(5): 551–553.
- 21 唐玉红, 周华, 朱敏佳, 陈华华, 陈丽, 岳利民, 何亚萍. 生理学非标准化考核与考试的探索与实践. *四川生理科学杂志* 2017; 39(3): 163–165.
- 22 夏欣欣, 孙连庆, 胡珊, 赵娇, 韩萍萍, 孙焯. 形成性评价与终结性评价相结合的成绩评定模式在留学生中医学教学中的应用. *中国中医药现代远程教育* 2019; 17(22): 32–35.
- 23 潘晨, 张付芝, 刘莹, 蔡红星. 中美高校教师发展中心发展及现况分析. *教育教学论坛* 2016; (17): 30–31.
- 24 张静, 刘莹, 孙晓楠, 蔡红星. 大数据背景下理想化学生评教系统的构建. *卫生职业教育* 2015; (20): 5–6.