

# 多维、质化、技术辅助的教学质量评价体系的建构

## ——以广东外语外贸大学为例

袁长青 林慧华 蔡文娟

(广东外语外贸大学, 广东 广州 510420)

**摘要:** 本文从大学评教对象单一、单纯量化, 区分度、信度、效度的质疑出发, 通过文献、调查问卷和实践案例, 提出构建以改进教学为目的的多维、质化、技术辅助的教学质量评价体系: 通过多维度, 增加质性、建设性评教, 奖励教学分享, 拓宽反馈途径等方式, 提高大学评教的区分度、信度、效度, 通过技术辅助提升其可操作性和效率。

**关键词:** 教学质量; 学生评教; 多维度; 质量评价; 质量体系

### 一、问题的提出

如何建构一套客观、准确、有效的教师教学质量评价体系, 既是高校建立科学评价体系的关键环节, 又是高校质量文化建设的重要内容。

#### (一) 评教的目的是什么

斯塔弗尔比姆说: “评价最重要的目的不是为了证明 (prove), 而是为了改进 (improve)。” 同理, 学生评教最重要的意义在于改进教学, 提高教学质量。当下国内高校将评教结果普遍运用于教师个体的奖罚, 是一种对结果的误用或滥用。

#### (二) 学生评教的效度和信度如何

作为评教主体部分的学生评教是一种“舶来物”, 从1927年至今在美国已被连续使用、大量研究、激烈争论了90多年, 有学者认为, 单一的学生评教制度下, 教师容易给学生高分以换取好评, 严格的优秀教师不容易得高分, 而对低评教分的担忧则可能扼杀教学创新, 导致教学要求简单化、教学方式娱乐化、教学内容碎片化。另有学者认为, 美国自推行 Weinberg 学生评教以来就一直存在“分数膨胀”现象。也有学者研究得出, 学生的实际学习情况与其评教分数是弱相关的, 很多学生打分时并没有意识到他们到底学了多少; 学生评教更多地受到对分数期望的影响, 而不是教师的教学有效性; 学生经常认为课堂氛围轻松愉快和低期望值的教师是好的, 所以教师在压力之下不得不给学生高分并减少作业量与难度; 学生经常抗拒要求更多投入、挑战更大的新教学法, 不迫使学生更努力的教师, 反而经常能得到更高的评教分。

从20世纪60年代起, 受教师评教影响, 美国高等教育学生分数开始显著膨胀, 课程任务量和难度同时下降。哈佛本科生学院院长哈雷曾公开告诫教师, 分数并非只有A、B两个等级, 不能只给学生打高分。

“分数膨胀、评教随意”现象在国内高校也普遍存在, 清华校长顾秉林就曾呼吁管理层重视教学管理机制建设, 因为清华自开展学生评教以来, 学生平均分从原来的70多变成了90多分。

要求严格、打分严谨的老师评教时往往得低分, 而很多课上内容不多、段子不少的教员反而屡得高分。“严师遇冷, 水课被捧”, 北大、清华、复旦学生2014年联合调查披露, 即使是名校学生, 也追捧“课水、人好、分高、好过”的教师和课程。据《中国青年报》(2010)报道, 某大学全校教师学生评教分大多在97分以上, 80分以下的只有一人, 大学生给老师打分成“一分钟游戏”。

#### (三) 同行评教是否可靠

众所周知, 国内高校普遍采用的“督导评定+同行评议+学生评教”相结合的多元评价体系中, 前两者往往因为涉及一线教师过多、落地费时费力, 难以有效全面铺开, 且同样也存在分数膨胀、干扰因素多的现象。譬如, 听课覆盖面有限, 同行之间存在竞争关系等。

#### (四) 如何合理使用评教结果

关于学生评教结果的公开性, 国内高校较少在学生中公开评教结果, 有学者表示: “结果不公开, 等于没评教。”虽然有个别大学(如山东大学)网上公开评教结果供学生选课参考, 但从我国高校整体情况来看, 相对于国外的普遍公开, 不公开评教结果仍是国内高校学生评教制度的普遍两岸现象之一。官方不公开, 民间盛流言, 个别“师兄师姐选课宝典”或者对老师的偏见已成为选课、选老师的指南。

### 二、课题组的调研结果

本课题组选取了广东外语外贸大学五个学院一、二、三年级的学生作为受访对象, 有效学生问卷206份。并对个别教师展开访谈调查。调研结果显示以下问题。

#### (一) 存在分数虚高、区分度低的情况

从该校教学管理部门在校园网公布的连续7个学期的数据显示, 每学期评教全校平均分都在93分以上(满分95分), 为了验证这是否意味着学生对老师普遍很满意, 课题组在调查问卷中提问: 根据本校评教系统, 选“优秀”(95分), “良好”(85分), “中”(75分), “及格”(65分), “差”(55分), 统分时去掉最高和最低的5%, 剩下90%的平均分为最后得分。去年一学年本校教师平均分分别是93.03和93.11(满分95分), 你认为这个结果表明什么? 结果只有3.17%的学生选择“本校学生对本校老师的教学情况非常满意”, 73.81%的学生选择“大多数学生评价教师教学随意, 不重视”。

当问卷要求学生填空, 给本校至今教过他们的所有大学老师打分(分数范围95-55分)时, 平均分只有81.31分, 其中最好的老师平均得分是92.42, 最差的老师67.72。这三个分数让学生完全靠“直觉”打分, 没有给任何具体的参考标准, 只是把“选

择题”换成“填空题”，但受访师生普遍认为，调研填空得出的数据比官方 ABCD 选择题统计的数据更符合常识和正态分布规律，后者分数虚高，区分度太低。

(二) 学生评教打分很少参考细则和标准，随意性大

课题组从问卷中了解到，42.06% 的学生一学期要评 6-9 位老师，49.21% 的学生一学期要评 10-14 位老师，30.95% 的学生一学期花在评教上的时间少于 5 分钟，38.1% 用了 5-10 分钟，也就是说，约 70% 的学生给每位老师评教的时间平均不到 1 分钟。很多学生访谈时坦言，一分钟可以评完所有老师。学校提供了详尽的细则和标准，但大多数学生不愿意花时间，且完全不按标准评。37.3% 的学生为了“尽快完成，基本上每个老师都选‘优秀’，有时老师的名字都没看清楚”，50.79% 的学生表示“多多少少会受自己的主观意愿影响，很少按照各项评估细则选择”，只有 11.11% 的学生选择“严格按照评估细则，根据每个教师的实际情况逐项、仔细、客观打分”。这些数据又从另一个侧面印证了常年平均分 93 分以上的评教分数和实际教学满意度并不完全一致。

(三) 学生对评教的意义不清楚，或感觉不到评教的作用

为什么要评教？调查数据显示，对评教的意义和作用不清楚的学生比例高达 57.14%。评教的主要作用应该是什么？77.78% 的学生认为应该是“学生给老师反馈、建议，让老师教得更好、学生学得更好”。这与管理部门的意愿是基本一致的。但是，42.86% 认为自己所参与的评教实际作用“很小”，49.21% 甚至说“基本没用，只是走过场”。对于学生而言，评教的意义没有得到强化，作用也无从体现，这就很好地解释了为什么学生评教时会比较随意，不愿意花时间。

(四) 评教的方式和结果对促进教师重视、改进教学的作用有限

评教最终以及最重要的目的在于促进教学改进，但如果将目前简单的量化结果孤立地传达给教师，不仅没有意义，还有误导性。访谈时 A 老师举了个例子，B 老师是本校的“科研牛人”。有一次公开课课后同事们建议在语言教学中多与学生互动交流，课堂不要偏学术，因为非专业本科生和英语专业研究生差别是很大的。但是，B 老师觉得自己上课有特色，且评教有 90 分，学生应该适应也满意他的授课方式，所以没必要改变。而据 B 老师的一名学生反映，上课时常常“睡倒一大片”，还有学生跑到隔壁班老师那“蹭课”。然而，很多老师都像 B 老师一样，对自己过高的分数有错误的解读，或者因为评教分数已经很高，没有改进教学的迫切性。

课题组问卷调查还发现，约 62% 的学生评教从来不写或一般不写反馈，只有不到 5% 的学生会给大部分老师写反馈。学生在什么情况下会更愿意给老师具体的评语和反馈建议？81.75% 表示，感觉自己的意见会受到重视时（比如老师接受反馈，并尝试改进）。

C 老师还讲了另外一种常见情况。自己工作 10 年只有 1 次评教成绩进了系部前 30%，获得了“校优秀教学奖”，但他却觉得自己教得最差的一次。当时他负责的是英语水平较低的班级，所以给了学生很多轻松有趣的学习内容，除了一个小组展示几乎

没有其他作业，期末考前还划了备考复习重点。

可见，如果把虚高、缺少区分度、甚至有误导性的学生评教成绩作为评优评奖评职称的重要依据，不仅难以令人信服，更不可能促进教师重视、改进教学。而学生看不到评教的意义和作用，就会更加敷衍应付，评教分数就会恶性循环，越来越没有有效度和信度，与评教的初衷背道而驰。所以，学生评教改革迫在眉睫。

### 三、思考与展望

无论从国内还是国际上，如何建立科学有效的教学质量评价标准与体系，都是一个难题，以下几个问题仍值得作进一步的思考和探索。

(一) 增加评价维度，提高评价结果的区分度

增加评价的维度。高校管理者和研究者普遍认为不应该放弃学生评教，而应该在不断改善的同时增加评价的维度。于剑等（2018）提出，为提高量化学生评教结果的信度，可以通过数据清洗来剔除非正常数据，消除评价主体潜在的非理性行为，过滤无关、无效数据等，比如学生作弊弊存在诚信问题，缺课太多失去评教资格等。赖亚尔斯建议学生评教在评价总分中的占比不超过 50%，因为学生权重越大，教师就越会投其所好。Berka（2006）认为评教有学生访谈、课堂录像、专家评教、校友评价、教学档案袋等 13 个证据来源，其中不少是质性评教。

优化评价指标体系。深化高校教师考核评价制度改革，逐步实现分类指导与分层次评价相结合，根据不同岗位教师的职责特点，教师分类管理和分类评价办法，分类分层次分学科设置评价内容和评价方式，提高评价分数区分度。

(二) 加大课程评价的力度

建立以课程为中心、以学生为基础，持续改进的理念。课程是教学改进的核心，学生是教学行为最直接的利益相关者，推进课程评估，不是评价老师，而是评价这门课程，评学生学得怎么样。具体以图 1 所示。

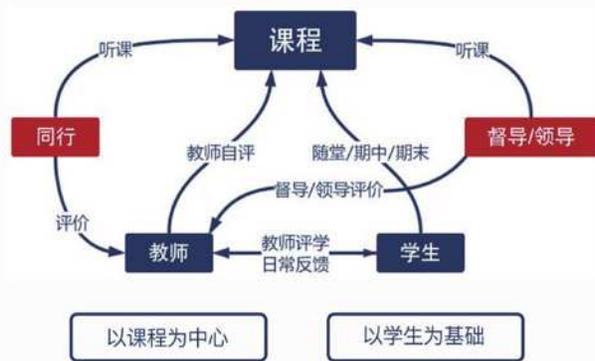


图 1 课程评价

(三) 奖励质化和建设性评教

单纯量化，是评教直接利益相关的教师、学生和管理者无法获得有价值信息的根源。质性评教反对把复杂的教育评估简化为数字，主张评价应全面反映教育现象，为改进教育实践提供可靠依据。比如可以借用网络平台，打开了师生交流的通道，让学生可以随时发表对教学的看法，可以多次、分阶段评教，避免学生

意见无处诉说形成积怨,也有助于教师养成倾听学生意见、交流改进的习惯。

#### (四) 奖励教学分享,完善同行评教

大力嘉奖、推动教师之间听课互评,以及线上线下教学展示、分享和帮扶互助,特别是要使之成为申报“优秀教学奖”的必要条件,把同行评议落到实处,使之更有效果和效率。

#### (五) 拓宽教学反馈的服务渠道

有学者建议,学校教学管理部门为教师提供发展性、个性化教学咨询服务项目。比如“期中学生反馈教学咨询服务”,20世纪70年代从华盛顿大学开始,如今已成为西方许多高校教学发展中心的常规服务项目,包括“前期会谈、课堂观摩(同时收集学生反馈)和后期会谈”,其宗旨是帮助教师发现问题、寻求解决办法,从而提高教学水平改善教学效果。这种做法有助于教师在课程结束前及时调整,更积极主动地改进教学(教师可以询问最密切相关的信息甚至征求批评,不必担心负面影响,因为反馈信息只给老师,避免了教师间的对比和来自领导及行政部门的压力)。同时,学生有机会帮助教师设计自己的教育过程,对学习效果存在积极影响。上海交大教学发展中心2012-2014年为50位教师和1800余名学生提供了该服务,受调查教师和学生的认可度分别是97%和99%。可见,这样的发展性评教方式是行之有效的,既能让学生感受到自己的意见是有价值的,也让老师感受到自己的改变是有意义的,能够为教与学带来积极的变化。

#### 参考文献:

- [1] 胡乐乐.美国高校学校评教研究90年:历史、争议与反思[J].高等教育研究,2019,38(9):91-109.
- [2] Marshall W.Meyer 著.姜文波译.绩效测量反思:超越平衡记分卡[M].北京:机械工业出版社,2005:43-60.
- [3] 雷宇.简单评价搞乱师生关系,大学生给老师打分成一分钟游戏[N].中国青年报,2010-02-04.
- [4] 于剑,韩雁,梁志星.中国民航大学多维课堂教学质量评价体系研究[J].教发展与评估,2018,34(2):63-72.
- [5] 路丽娜,王洪才.质性评教:走出学生评教困境的理性选择[J].现代大学教育,2016(2):93-96.
- [6] 李雁冰.课程评价研究[D].上海:华东师范大学博士后研究报告,2000.
- [7] 梁竹梅.期中学生反馈教学咨询服务的发展研究——国外研究与本校探索[J].现代大学教育,2015(6):101-106.
- [8] 王国华,卓泽朋,周光辉.大数据背景下线上教学质量监控与评价体系的建构[J].淮北师范大学学报(哲学社会科学版),2020,41(03):107-111.
- [9] 陈倩倩.建构主义视角下高校课堂教学质量评价体系研究[J].襄阳职业技术学院学报,2022,21(02):81-85.
- [10] 许小雅.高职教育教学质量监控与评价体系构建[J].青年与社会,2019(15):72-73.
- [11] 郑革.高职教育教学质量监控与评价体系的建构途径[J].职教通讯,2016(35):35-37.
- [12] 何云辉.论五年制师范教育教学质量评价体系建构——以吉首大学师范学院为例[J].遵义师范学院学报,2013,15(02):66-68.
- [13] 孙坚,佟秀凤,卞德强.基于互联网的高校课堂教学质量评价方法改进对策[J].黑龙江科学,2022,13(07):132-133.
- [14] 游明伦,暨星球,姜山.“双高计划”视域下高职院校校本教学质量评价一体化改革探究——以贵州轻工职业技术学院为例[J].铜仁学院学报,2022,24(02):17-30.
- [15] 杜乐,周昊.OBE理论视角下公安院校实践教学评价体系的构建[J].教育观察,2022,11(1):38-41.
- [16] 赵馨蕊,周雨青,瞿林云,张志红.基于层次分析法的大学物理MOOC教学质量评价指标体系的建构[J].物理与工程,2018,28(04):45-56.
- [17] 胡建军.信息化背景下应用型本科高校教师教学质量评价的探索[J].上海教育评估研究,2022,11(01):26-29.
- [18] 赵方.高校课堂教学质量评价过程中存在的问题和对策[J].教书育人(高教论坛),2022(06):19-20.
- [19] 黄海笛,管程程.区块链技术在高校教学质量评价体系中的应用研究[J].现代信息科技,2022,6(02):160-163.
- [20] 韩青松,郭啸,刘伟诚.“双一流”建设背景下高水平大学研究生教学质量评价的必要性和重要性探究[J].大学,2022(02):71-75.
- [21] 石国玺,牛代买,周向军.OBE视域下基于层次分析法的课程教学质量评价体系构建[J].长春工程学院学报(社会科学版),2021,22(04):81-84.
- [22] 金莉,刘粉,马龙,张斌.基于可拓层次分析法的课程教学质量评价模型构建与应用[J].合肥师范学院学报,2021,39(06):94-98.
- [23] 吴丽月,侯国风.利益相关者视角下高职实践教学质量评价体系构建[J].岳阳职业技术学院学报,2021,36(06):7-12.
- [24] 刘莉莉,郭祖华,孙冬.教学改革与创新实绩融入教学质量评价的探索与实践——以河南工学院“3+1”教师教学质量评价为例[J].河南工学院学报,2021,29(06):41-44.
- [25] 谭丹凤.高职院校会计专业智慧化实践课程教学质量评价体系构建及应用研究[J].经济师,2021(10):231-233.
- [26] 周艳,蒋英,陈宇焱.高职院校课堂教学质量评价体系构建的问题及对策研究[J].中国多媒体与网络教学学报(中旬刊),2021(09):189-191.

基金项目:2019年广东省高等教育教学改革项目(综合类)“多维、质化、技术辅助的发展性高校教学评估改革方案”。

作者简介:袁长青(1970-),男,广东省和平县,广东外语外贸大学教学质量管理评估中心,主任、研究员、博士,主要从事教育管理与评估、教育文化研究。