

高校教学评价制度对教师教学绩效影响的实证研究：基于全国63所高校的数据检验

陶政宇, 李楠

(首都经济贸易大学, 北京 100070)

摘要:教学质量是高校的立校之本, 要提高教学质量, 需要对教师的教学工作进行有效的评价。虽然我国高校教师的教学评价开展了近四十年, 但现行评价制度的效果如何, 能否对提升教师的教学绩效产生积极作用, 较少有学者对此展开实证分析。基于全国63所高校的匹配数据, 使用分层线性回归模型, 实证分析评价制度的构成及其对提升高校教师教学绩效的影响。当评价制度的构成具备以下四个特点时, 能够提升教师的教学绩效: 第一, 评价主体不宜选择全面评价; 第二, 评价内容要面对教学的全过程; 第三, 评价方式不能只有定性评价; 第四, 评价结果要反馈具体问题。

关键词:高校教师; 教学评价制度; 教学绩效; 匹配数据; 分层线性模型

中图分类号: G645.11

文献标志码: A

文章编号: 1003-2614(2024)06-0030-08

DOI: 10.19903/j.cnki.cn23-1074/g.2024.06.019

一、问题提出

高校是培养高层次创新人才的重要基地, 教学是高校的基本职能, 教学质量决定了高校的教育质量。要提高教学质量, 需要对教师的教学工作进行评价。随着高等教育的普及化、高等教育规模的扩大化, 我国高等教育实现了从精英化阶段到大众化阶段再到普及化阶段的转轨, 教学质量成为政府、教育界乃至社会大众关注的焦点。2020年10月, 中共中央、国务院印发了《深化新时代教育评价改革总体方案》, 其中对高等教育质量、教学评价等问题给予高度关注。高等教育高质量发展, 最根本的就是回归高校教学与人才培养的本职职能, 从而实现人的高质量发展^[1]。而教学评价正是保障教学运行规范、提高教学管理水平、提升教育教学质量的重要手段, 对高校的教学管理实践具有指导作用, 直接影响着大学未来发展的走向。我国的高校教师教学评价开始于20世纪80年代中期^[2], 发展至今, 越来越多的高校把教学评价作为教学管理的重要手段, 也作为学校考核教师的一项重要内容。高校的教学评价制度在我国运行了近四十年, 在借鉴西方相应制度的基础上, 结合本土情况形成了自己的制度体系。有些做法在借鉴的过程中可能存在偏颇, 比如关于评价结果只反馈分数; 有些坚持下来的做法也可能并未达到预期的效果, 比如360度评价主体的选择。尽管学术界已经对高校教师教学评价制度构成展开研究, 但对于评价制度的效

果, 较少有学者对此进行实证分析。本研究将基于制度效果的视角, 采用全国63所高校的匹配数据, 实证分析教学评价制度的构成及其对提升教师的教学绩效产生的作用以及教学评价制度的完善途径, 以期对教学评教制度的高质量发展提供启发与参考。

二、文献回顾

综观高校教师教学评价的相关研究, 可以将研究成果具体分为评价目的、评价主体、评价方式、评价内容、评价结果等五个模块。

第一, 关于评价目的, 可将其分为基于管理目的的“奖惩性评价”和基于发展目的的“发展性评价”^[3]。尽管学者们都认可教学评价的最终目的是提高教学质量, 但对于教学评价的目的究竟为何, 不同学者所持观点并不相同。陈明学和郑锋^[4]、郭丽君^[5]认为应当将奖惩性评价目的和发展性评价目的结合起来使用。杨平^[6]、陈春莲和唐忠^[7]主张采用发展性评价目的。在高校落实教学评价制度过程中, 评价结果主要作为教师奖惩的依据而存在^[8]。

第二, 关于评价主体, 学者的看法主要集中在以下方式上: 以学生为主体的单维度评价^[9]; 以学生和同行(或专家)为主体的二维度评价^[10]; 分别以学生、同行(或专家)、领导(或管理者)为主体^[11], 或以学生、同行(或专家)、教师为主体的三维度评价^[12]; 多维度评价^[13]。因为任何一个评价主

收稿日期: 2023-08-27

基金项目: 2017年度教育部人文社会科学研究青年基金项目“我国高校教师教学评价制度的效果及重构研究”(编号: 17YJC880055)。

作者简介: 陶政宇, 首都经济贸易大学劳动经济学院博士研究生, 讲师, 研究方向: 高校教师发展、高校思想政治教育与管理; 李楠, 首都经济贸易大学劳动经济学院副教授, 硕士研究生导师, 研究方向: 高校教师绩效管理。

体都有其局限性,多维度评价可以实现互补,所以越来越多的学者主张采用多维度评价主体^[14]。考虑到多维度评价在实际工作中并不容易操作,受客观条件的限制,无法保证对教师的评价是客观公正的,在评价过程中会采用二维度评价或者三维度评价。还有部分学者认为,由于学生是教学的直接受益者,使用学生评教有一定优势^[15],但由于学生对事物的认识多处于感性阶段,缺乏理性思考,同时考虑到教学的特点,应该重视教师自评和同行评价,完善多元主体参与的教学评价体系^[16]。

第三,关于评价方式,美国高校教师教学评价分为学生评价、同行评价、教师自我评价和系主任评价。当以学生为评价主体时,评价方式有等级评价、书面评价、面谈等方法^[17];以同行为评价主体时,主要采用考察授课材料和进行课堂访问等方式^[18];以教师本人为评价主体时,则使用书面报告、教学档案袋等方法^[19];系主任评价一般采用书面评价的方式^[20]。国外高校在教学评价中常采用定性与定量结合法^[21]、形成评价法等评价方法^[22]。在我国高校的实践中,教学评价方法侧重评价量表和听课,多以定性研究方法对教学评价制度改革进行理论反思^[23]。

第四,关于评价内容,我国高校主要集中在教学态度、教学内容、教学方法、教学效果等四个方面^[24]。也有学者认为除上述教学的共性表现以外,还要关注教师的个性化教学行为^[25],尽可能避免出现用一份评价表评价不同类型教育教学活动的情况^[26]。当前的教学评价指标习惯于从学校管理者的需要出发,缺少对学生学习的关怀^[27]。随着教育的核心逐渐向学生回归,教学评价范式的转变方向应从学生需求角度出发,以学习为中心开展^[28]。不同于美国高校,我国高校的教学评价内容呈现出共性多于个性、统一性多于多样性的特点。正因如此,在评价内容上,可以考虑根据专业性质、课程特点和不同成长阶段教师的差异设计不同的评教标准,以体现出教学评价内容的多元性和差别性^[29]。

第五,关于评价结果,研究发现,教师认为,有利于自身朝着更好的方向进行改变的信息才是最有用的。比起分数或等级,有针对性的建议会对教师的教学改进更有促进作用^[30]。在操作层面,尽管高校在教学评价量表中会设有开放式问答的主观题目,但由于该题目不是必填内容,学生会选择不填写。学校的教学评价结果往往以分数或等级的形式呈现,没有反馈有益于教师改进教学的评价结果,如每个评价细目的具体得分、不同教师间评价结果的横向对比信息、教师个人历年评价结果的纵向对比信息等^[31]。为使教师更好地了解自身不足、明确未来改进方向,学校应当根据教师的差异采用恰当的形式提供具体的评价结果信息。

通过对相关文献的整理发现:关于高校教师教学评价制度的研究多集中于制度的应然研究,即教学评价制度应该是什么样的,对评价制度的各部分构成仍存在不同的观点。基

于此,本研究着重于以下两个方面内容:一是对教学评价制度的效果进行研究,分析评价制度如何构成才会有利于提升教师的教学绩效水平;二是对教学评价制度的效果展开定量研究,在问卷调查的基础上,分别对学校 and 教师发放问卷,将二者的数据进行匹配,通过分层线性回归模型,分析不同评价制度对教师教学绩效的影响作用,定量地说明教学评价制度的实施效果。

三、理论模型的构建

(一)数据来源

本文使用的数据来源于首都经济贸易大学课题组于2018年9月开始对高校教学评价制度的效果展开的调查,主要收集教师的个人基本信息、教师对教学评价制度的评价、不同学校的教学评价制度构成等信息。

为了从学校和教师两个层面探究影响教学绩效水平的因素,课题组分别发放了学校问卷和教师问卷,对两个层次的问卷进行数据匹配。教师问卷通过问卷星的方式开展调查,学校问卷则是根据教师问卷的填写情况,直接联系相应学校的教务部门进行填写,每个学校只填写一份学校问卷。调查采取多阶段抽样方法,首先从东、中、西部地区随机抽取63所高校,根据该校教师数量再对教师进行随机抽样。调研共回收有效学校问卷63份、有效教师问卷908份。63所高校来自16个不同省市,调查的受众面较广,包括中央部属本科院校、省属本科院校等。由于学校的层次不同、特点不同,因此调查结果具有一定的代表性和说服力。

(二)变量与测量

1. 因变量:教师的教学绩效水平

关于教学评价制度的效果,目前学界没有统一的说法,本文将其定义为教学评价制度的实施是否提升了高校教师的教学绩效,以此来衡量教学评价制度的效果。教学绩效的指标可以用教学质量(结果)和教学投入(过程)来表示。由于目前可测量的教学质量指标有学生评教的成绩和成绩变化趋势,可测量的教学投入指标为教师的备课时间,因此本文的因变量包括学评教成绩、备课时间和成绩变化趋势。参考以往研究^[32]的做法,学评教成绩和备课时间分别根据教师问卷中的两个题项界定,这两个题项分别为:“近一学年的学评教分数是多少分?”“近一学年每课时的平均备课时间约为多少小时?”成绩变化趋势根据教师问卷中的一个题项界定,该题项为:“近五年的学评教成绩变化趋势是什么?”取值1、2、3,分别代表“下降”“基本持平”“上升”。

2. 自变量

自变量包括个体和学校两个层次的变量。个体层次的变量包括教师的性别、教龄、学历、职称、教学获奖情况、科研项目情况、论文数量、科研时间。关于教师性别、教学获奖情况、科研项目情况、学历、职称变量的测量,在教师调查问卷

中分别询问:“您的性别是什么?”“近五年是否获得校级及以上级别的教学效果奖项?”“近五年是否主持校级及以上级别的科研项目?”“您的最高学历是什么?”“您的职称是什么?”相应设置为有序或无序的分类变量。关于教龄、论文数量和科研时间变量的测量,在调查问卷中分别询问:“您的教龄是多少年?”“近五年在核心及以上期刊发表的论文数量有多少篇?”“近一年平均每周科研时间约为多少小时?”将其作为连续变量直接纳入模型。

学校层次的变量包括教学评价制度的评价主体、评价内容、评价方式和反馈情况。关于评价主体数量变量的测量,在学校问卷中询问:“学校教学评价的主体情况是?”取值1、2、3,

分别代表“单维度评价”“多维度评价”“全面评价”。关于评价内容变量的测量,在学校问卷中询问:“学校教学评价的内容是?”取值1、2、3、4,分别代表“只评价课堂内容”“既评价课前准备,也评价课堂教学”“既评价课堂教学,也评价课后安排”“既评价课前准备,也评价课堂教学及课后安排”。关于评价方式变量的测量,在学校问卷中询问:“学校教学评价的方式是?”取值1、2、3,分别代表“定性评价”“定量评价”“定量评价和定性评价相结合”。关于反馈情况变量的测量,在学校问卷中询问:“学校对教学评价结果的反馈情况是?”取值1、2、3,分别代表“没有任何反馈”“没有反馈具体问题”“反馈具体问题”。主要变量定义及数据信息来源见表1。

表1 主要变量定义及数据信息来源

变量定义	数据信息来源及说明
因变量	
学评教成绩	回答“近一学年的学评教分数是多少分?”
备课时间	回答“近一学年每课时的平均备课时间约为多少小时?”
成绩变化趋势	回答“近五年的学评教成绩变化趋势是什么?”:1=下降,2=基本持平,3=上升
自变量:个体层次	
性别	选择“您的性别是什么?”:0=女,1=男
教龄	回答“您的教龄是多少年?”
学历	选择“您的最高学历是什么?”:1=本科,2=硕士研究生,3=博士研究生
职称	选择“您的职称是什么?”:1=讲师及以下,2=副教授,3=教授
教学获奖情况	回答“近五年是否获得校级及以上级别的教学效果奖项?”:0=否,1=是
科研项目情况	回答“近五年是否主持校级及以上级别的科研项目?”:0=否,1=是
论文数量	回答“近五年在核心及以上期刊发表的论文数量有多少篇?”
科研时间	回答“近一年平均每周科研时间约为多少小时?”
自变量:学校层次	
评价主体	回答“学校教学评价的主体情况是?”:1=单维度评价,2=多维度评价,3=全面评价
评价内容	回答“学校教学评价的内容是?”:1=只评价课堂内容,2=既评价课前准备,也评价课堂教学,3=既评价课堂教学,也评价课后安排,4=既评价课前准备,也评价课堂教学及课后安排
评价方式	回答“学校教学评价的方式是?”:1=定性评价,2=定量评价,3=定量评价和定性评价相结合
反馈情况	回答“学校对教学评价结果的反馈情况是?”:1=没有任何反馈,2=没有反馈具体问题,3=反馈具体问题

(三)方法与模型

本文采用分层线性模型(Hierarchical Linear Modeling, HLM),该方法的目的是基于一系列非同一个层次的自变量,对因变量的值进行估计,由此可见,分层线性模型适用于具有层次性的研究内容。虽然国内有关探讨教学评价制度效果的研究尚未普遍使用该方法,但本研究旨在分析高校教师教学评价制度的效果,研究内容和数据本身具有层次特性。教师教学评价制度的效果不仅受到教师个人特征的影响,也受到教师所在学校的学校特征的影响。此外,HLM可针对学校之间教师的教学绩效水平的变异作分析,而传统的回归分析则无法做到。本文将个体层次变量、学校层次变量逐步加入模型,进行HLM分析,以验证研究假设。

首先,零模型是不加任何变量对模型方差进行分解,以判断是否有必要构建分层线性模型。零模型如下所示:

$$\text{Level 1 方程: Effect}_{ij} = \beta_{0j} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

$$\text{Level 2 方程: } \beta_{0j} = \gamma_{00} + \mu_{0j} \quad \mu_{0j} \sim N(0, \tau_{00})$$

其中:Effect_{ij}是j学校的i教师的教学绩效水平;β_{0j}为j

学校的平均水平;r_{ij}是教师个体的随机误差,表示j学校的i教师与学校平均水平的差异;γ₀₀为总体平均水平;μ_{0j}是学校的随机误差,表示j学校平均水平与总体平均水平的差异。

其次,在零模型中加入个体层面变量和学校层面变量,构建如下模型:

$$\text{Level 1 方程: Effect}_{ij} = \beta_{0j} + \beta_{1j}T_{1ij} + \beta_{2j}T_{2ij} + \dots + \beta_{nj}T_{nij} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

$$\text{Level 2 方程: } \beta_{0j} = \gamma_{00} + \gamma_{01}S_{1j} + \mu_{0j} \quad \mu_{0j} \sim N(0, \tau_{00})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}S_{1j} + \mu_{1j} \quad \mu_{1j} \sim N(0, \tau_{11})$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}S_{1j} + \mu_{2j} \quad \mu_{2j} \sim N(0, \tau_{22})$$

.....

$$\beta_{nj} = \gamma_{n0} + \gamma_{n1}S_{1j} + \mu_{nj} \quad \mu_{nj} \sim N(0, \tau_{nn})$$

其中,T和S分别表示教师特征变量和学校特征变量。

四、数据分析结果

(一)描述性统计分析

表2报告了主要变量的描述性统计结果。由表2可知,

教师近一学年学评教成绩的平均值约为 91.0 分,近一学年每课时的平均备课时间约为 5.5 小时。关于近五年的学评教成绩变化趋势,3.41% 教师的的成绩呈下降趋势,66.63% 教师的的成绩基本持平,29.96% 教师的的成绩呈上升趋势。

关于个体层次的变量,在受访者中,有 39.8% 的男性,教龄约为 12 年,近一年平均每周科研时间约为 16 小时,近五年在核心及以上期刊发表的论文数量约为 4 篇。在受访者中,最高学历是本科、硕士研究生、博士研究生的比例分别为 3.96%、30.29%、65.75%。关于教师的职称,46.48% 是讲师及以下,43.61% 是副教授,9.91% 是教授。38.7% 的受访者在近五年获得过校级及以上级别的教学效果奖项,70.5% 的受访者在近五年主持过校级及以上级别的科研项目。

关于学校层次的变量,学校教学评价主体的统计结果如下:27.53% 的学校是单维度评价,65.65% 的学校是多维度评价,6.83% 的学校是全面评价。学校教学评价内容的统计结果如下:30.51% 的学校只评价课堂内容,33.26% 的学校既评价课前准备也评价课堂教学,7.71% 的学校既评价课堂教学也评价课后安排,28.52% 的学校既评价课前准备也评价课堂教学及课后安排。学校教学评价方式的统计结果如下:17.84% 的学校采用定性评价,4.19% 的学校采用定量评价,77.97% 的学校采用定量评价和定性评价相结合的方式。学校对教学评价结果的反馈情况的统计结果如下:5.40% 的学校没有任何反馈,23.35% 的学校没有反馈具体问题,71.26% 的学校会反馈具体问题。

表 2 主要变量的描述性统计(在个体层次展示的部分均值和标准差数据为男教师)

变量定义	均值	标准差	最小值	最大值	样本量
学评教成绩	91.036	4.778	60	100	908
备课时间	5.515	6.161	0	20	908
成绩变化趋势	2.265	0.513	1	3	908
个体层次					
性别	0.398	0.490	0	1	908
教龄	12.137	8.373	0	35	908
教龄平方	217.333	275.620	0	1225	908
学历	2.618	0.562	1	3	908
职称	1.634	0.656	1	3	908
教学获奖情况	0.387	0.487	0	1	908
科研项目情况	0.705	0.456	0	1	908
论文数量	3.671	4.032	0	20	908
科研时间	16.141	14.484	0	60	908
学校层次					
评价主体	1.793	0.549	1	3	908
评价内容	2.343	1.186	1	4	908
评价方式	2.601	0.773	1	3	908
反馈情况	2.659	0.577	1	3	908

(二)教学评价制度对教师的学评教成绩的影响

零模型实际就是方差分析,通过计算组内相关系数(Intra-class Correlations, ICC)可以发现,零模型(模型 A1)的 ICC^①为 16.19%,也就是说教师教学绩效水平的差异有 16.19% 来自学校间的差异,模型的检验结果也证明较之简单线性回归模型,本研究更适合采用分层线性模型(体现在表 3、表 4 和表 5 中的“其他控制变量”上,分别用“×”和“√”来表示不适合和适合)。接下来,在两层的分层线性回归模型中,让教师个体层次模型的截距随不同学校而发生变化(即随机截距)。在具体建模策略上,本文通过逐步在模型中加入变量的方式考察学校制度对教师的学评教成绩的净效应。在模型 A1 的基础上,本文又构建了 2 个模型,分别为加入个体层次的非完全模型(模型 A2)和加入学校层次的全模型(模型 A3)。表 3 报告了模型估计的结果。

模型 A2 显示,第一,教龄越长的教师的学评教成绩越高。随着教龄的增长,教师的教学经验不断丰富,对课程更加熟悉,把握也更加充分,学评教成绩会随之升高。第二,较之本科学历,学历为硕士研究生、博士研究生的教师的学评

教成绩更高。目前高校中的绝大多数教师,特别是中青年教师,拥有硕士、博士学位是必要条件,只有极少数年长教师由于历史原因对其学历没有特别要求,但由此可以看出,教师是一个需要终身学习的职业,只有不断学习成长,才能跟上时代的发展,带给学生前沿的知识,得到学生的认可。第三,获得过校级及以上级别的教学效果奖项的教师的学评教成绩更高,主要原因是获得过教学效果奖项的教师通常都是教学效果好、学评教成绩一直名列前茅的教师。第四,在核心及以上期刊发表的论文数量越多、主持过校级及以上级别的科研项目的教师的学评教成绩更高。

模型 A3 在模型 A2 的基础上加入了学校层次变量。结果显示,个体层面变量的影响(主要是显著性、方向、大小等三个方面)基本不变,保持稳定。关于制度的影响,表现在以下三个方面:

第一,关于教学评价主体,可以发现:与单维度评价相比,多维度评价或全面评价都会使得教师的学评教成绩显著下降。导致这一结果的可能原因有两个:一方面,随着评价主体的增加,由于打分标准不一致,教师的学评教成绩相对

分散,所以评价主体越少,分数反而越高;另一方面,在本次调研中,有1/4多的学校采用单一学生评教的方式考核教师的教学质量,而单一学生评教的分数普遍较高。这也可以解释为:当教学的评价主体只有学生时,教师只需要考虑学生的需求,所以学评教的成绩普遍较高;当教师要面对更多的评价主体时,不仅需要考虑学生的需求,还要考虑其他主体的要求和偏好,这会导致他们对单维度评价的关注和投入降低,从而导致其学评教成绩显著降低。单纯或过度使用学生评教也可能产生一些负面的结果或问题,如大多数采用以学生为主要教学评价主体的高校会出现明显的分数膨胀现象^[33]。

第二,关于教学评价内容,结果显示:与只评价课堂教学相比,“评价课前和课堂”“评价课前、课堂、课后”这两种情况,教师的学评教成绩都会显著提高。本次调查发现,有将近1/3的高校只评价教师的课堂教学质量,而如果将课前准备和课后作业均纳入评价体系,教师的学评教成绩反而更高。由此也可以看出,教师在人才培养的过程中,并非只需要完成课堂教学,课前准备和课后作业环节都能看出教师对于教学及学生学习效果的重视。如果高校教师的教学评价只是对课堂教学进行评价,则无法全面反映教师教学的真实情况。这也会让教师错误地认为,自己只需要把课上好,不需要在其他方面投入更多的时间和精力。事实上,在大学教育中,教师除了要向学生传授知识,师生间的思想交流、师生关系及其相互合作,也是促进学生发展和专业成长的重要条件。评价内容应该具有系统性,贯穿教学的全过程,包括教师的课前准备、课堂教学的实施和课余时间的辅导^[34]。

第三,关于教学评价结果的反馈情况,研究显示:与没有反馈相比,有反馈但未反馈具体情况,教师的学评教成绩显著下降。由此可以看出,只反馈分数并不能对教学质量的提升产生积极作用,反而有不利影响。在本次调研中,有1/4多的高校没有给教师任何评教反馈或虽有反馈但没有反馈具体问题,如果反馈给任课教师的信息只有分数或评价等级,教师无法从这些简单、孤立的数据中找到提高教学质量的方法。

表3 估计学评教成绩的分层线性模型

	模型 A1	模型 A2	模型 A3
教龄		0.140 ** (0.066)	0.139 ** (0.066)
教龄平方		-0.002 (0.002)	-0.002 (0.002)
学历(以本科为参照)			
硕士研究生		2.153 ** (0.850)	2.132 ** (0.853)
博士研究生		2.090 ** (0.877)	2.075 ** (0.874)
教学奖项情况		0.987 *** (0.319)	0.981 *** (0.319)
科研项目情况		0.645 * (0.369)	0.690 * (0.367)
论文数量		0.101 **	0.104 **

		(0.046)	(0.046)
评价主体(以单维度评价为参照)			
多维度评价			-1.351 ** (0.651)
全面评价			-2.359 ** (1.033)
评价内容(以只评价课堂教学为参照)			
评价课前和课堂			2.422 ** (0.943)
评价课堂和课后			-0.996 (1.151)
评价课前、课堂、课后			1.380 ** (0.693)
反馈情况(以没有反馈为参照)			
不反馈具体问题			-1.852 * (1.108)
反馈具体问题			-1.201 (1.163)
其他控制变量	×	√	√
常数项	90.924 *** (0.318)	86.522 *** (0.967)	91.620 *** (2.552)
样本量	908	908	908
组数	63	63	63

注:*、**、***分别代表在10%、5%和1%置信水平下显著。

(三)教学评价制度对教师的备课时间的影响

为探究教学评价制度对教师备课时间的影响,本文构建了3个模型,分别为零模型(模型B1)、只有个体层次变量的非完全模型(模型B2)、加入学校层次变量的全模型(模型B3)。表4反映了因变量为备课时间的分层线性模型分析结果。零模型(模型B1)显示,校间的方差成分占总方差的比例(即组内相关系数ICC)为0.1103^[35],故需要采用分层线性模型进行分析。

模型B2只包括个体层次因素,结果显示,随着教龄的增加,教师的备课时间呈现先减少后增加的趋势。对于大多数教师来说,随着教龄的增加,其对主授课程的熟悉程度也会增加,所以备课时间呈现减少的趋势。随着信息化时代的到来,旧有的课程也会遇到瓶颈和挑战,无论是从内容、形式上都需要推陈出新,对教师来说需要投入更多的时间重新打造课程,由此开始出现备课时间增加的趋势。

模型B3在模型B2基础上加入了学校层次因素。可以发现,个体层次变量的影响(主要是显著性、方向、大小等三个方面)基本不变,保持稳定。在学校层次,与单维度评价相比,当教学评价主体为全面评价时,教师的备课时间会显著减少。这主要是因为当评价主体为全面评价时,教师需要面对的群体增多,特别是加入同行评价和领导评价时,还要考虑人际关系的因素,这势必会分散教师的精力,减少教学投入。有研究指出,与很多同事进行人际交往会显著降低教学绩效^[36]。第二,关于教学评价的方式,与定性评价相比,定量评价或定量评价和定性评价相结合的方式均会使教师的

备课时间出现显著上升。关于教师教学评价的方式,很多研究都提出定性评价方式的重要性,这主要是出于对发展性评价目的的认可,认为定性评价能够帮助教师发现问题,指明发展方向,有助于发展性评价目的的实现。本研究也发现,定量评价同样有其存在的必要性。比如,对教师的评分以及排名可以让教师找到自己在群体中的相对位置,能够激发教师努力教学。从备课时间的增加可以看出,教师除了重视自己教学质量的提升,还看重与同事教学质量的比较。

表 4 估计备课时间的分层线性模型

	模型 B1	模型 B2	模型 B3
教龄		-0.346*** (0.089)	-0.342*** (0.089)
教龄平方		0.007*** (0.003)	0.007*** (0.003)
评价主体(以单维度评价为参照)			
多维度评价			-1.301 (0.808)
全面评价			-4.677*** (1.223)
教学评价的方式(以定性评价为参照)			
定量评价			3.900*** (1.423)
定量评价和定性评价相结合			3.053** (1.325)
其他控制变量	×	√	√
常数项	5.731*** (0.372)	7.333*** (1.268)	5.030 (3.550)
样本量	908	908	908
组数	63	63	63

注: *、**、*** 分别代表在 10%、5% 和 1% 置信水平下显著。

(四) 教学评价制度对教师的学评教成绩变化趋势的影响

本文使用 orderedlogit 模型分析教学评价制度对教师的学评教成绩变化趋势的影响。由表 5 可知:

第一,随着教龄的增加,教师的学评教成绩变化呈现先下降后上升的趋势。随着教龄的增加,教师对所授课程越来越熟悉,不会花费更多的时间在备课上,课程的形式和内容也不会有太大变化,有可能使得成绩变化有所下降。直到学生对课程的满意度比较低,教师意识到必须对课程作出改变时,就会重新投入时间和精力在课程建设上,通过在课程中加入新的形式或内容等方式提高课程质量,其学评教成绩也会逐渐上升。第二,主持科研项目的教师的学评教成绩变化呈现上升趋势。科研对教学质量的提升有促进作用^[37],这不仅体现在教师之间的教学质量对比上,也同样适用于教师个人的教学质量对比。由此也更加印证了洪堡时代的思想——科研是最好的教学。第三,学校对教学评价结果的反馈,与没有反馈相比,在有反馈但未反馈具体情况时,教师的学评教成绩变化出现下降趋势。关于教学反馈,如果只是反馈分数和排名,而没有反馈具体问题,则不利于教师的教学

质量提升。这种情况不但对不同教师间的教学质量比较如此,对同一教师的教学质量变化也同样如此。

表 5 估计成绩变化趋势的排序模型

	模型 C1	模型 C2
教龄	-0.095*** (0.031)	-0.094*** (0.032)
教龄平方	0.002* (0.001)	0.002* (0.001)
科研项目情况	0.344* (0.177)	0.343* (0.180)
反馈情况(以没有反馈为参照)		
不反馈具体问题		-0.836* (0.467)
反馈具体问题		-0.717 (0.475)
其他控制变量	×	√
样本量	908	908
Pseudo R ²	0.019	0.027

注: *、**、*** 分别代表在 10%、5% 和 1% 置信水平下显著。

五、讨论与启示

制度学家指出,许多组织行动都被认为是理所应当的,没有人好奇一项活动为什么开始或为什么应该继续,每个人只会认为这是唯一且理所当然的方式。组织倾向于遵循那些已经被其他组织确立合法性的管理结构和实践,管理者也会陷入同样的陷阱,只做“大家正在做的”。目前各高校都制定了教学评价制度,但制度的构成却千差万别,无论从评价主体、评价内容、评价方式还是评价结果的反馈都没有统一的做法。究其原因,主要是高校在实施制度的过程中,逐渐忽略了其功能,即教学评价制度是否真正提升了教师的教学绩效水平。本研究从教学评价制度的效果(即教学评价制度是否达成了其功能)出发,以提升教师教学绩效水平为基础,通过比较不同制度构成对教师教学绩效的影响发现,当高校的教学评价制度具备以下特点时,能更好地提升教师的教学绩效水平。

第一,评价主体不宜选择全面评价。本研究显示,当评价主体为全面评价时,教师的教学质量和备课时间都会显著下降。虽然美国高校对教师的教学评价普遍采取全面评价,国内也有学者提出将 360 度绩效考评法用于高校教师的教学评价,但不论是国外的全面评价,还是国内的 360 绩效评价体系,都有其适用的情景与特点,而我国高校教师的教学评价也有其自身的独特性,不能简单地照搬照用,否则将适得其反。目前,几乎所有的高校都在开展学生评教,学生评教有其存在的必要性和重要性,但同时也有它的局限性,比如:学生的评教能力有限;评教结果往往出现“集体高分”,且区分度很小^[38];当评教结果用于人事目的时,教师可能会具有评价标准偏移的倾向^[39]。虽然学生可以作为教师教学评价的主体之一,但不能是唯一,与此同时,可以考虑加入专家等评价主体来弥补学生评价的缺陷。

第二,评价内容要面对教学的全过程。本研究显示,当评价内容为教师教学的全过程时,教师的教学质量显著提升。一直以来,各高校对于教师教学的评价关注的都是教师的课堂教学质量,忽视了教师课前准备和课后安排。精心设计、认真准备课程以及学生课外的自主学习和课下的师生互动均是教育教学活动的重要环节,同样影响着人才培养质量,因此将教育教学活动开展前的准备工作和课堂教学之外的活动纳入教学评价的范畴是非常必要的^[40]。高校应该重新界定教学评价的内容,考虑由专家和学生共同参与评教,发挥各自的专长和优势,课堂教学以学生评价为主,课前准备和课后安排以专家评价为主,学生没有能力评价的课堂教学内容也可以由专家来评价,二者的结合共同构成教师的教学评价,使评价更加科学合理。

第三,评价方式不能只有定性评价。本研究显示,当评价方式含有定量评价时,会加大教师对教学的投入。当教师既知道自己的教学评分,也了解院系的平均分时,就能清楚二者之间是否存在差距,从而推动其内在动机充分发挥作用,提升教学表现^[41]。定性评价有其存在的必要性,定量评价也有其存在的重要性,二者应当共同构成教学评价方式。

第四,评价结果要反馈具体问题。本研究显示,如果教学评价结果的反馈只有分数而没有具体问题,比没有任何反馈对教师的教学质量更有负向影响。在管理过程中,往往将评价结果以分数或排名的方式呈现。对于教师来说,只有分数的反馈结果无助于教师的教学反思与改进。只有反馈具体问题,才能让教师清楚地了解自己在教学中存在的具体问题是什么,才能对症下药,切实地改进教学,提高教学质量。高校的管理部门和教学管理者不仅要及时了解教师的需求,为教师教学的改进提供必要的支持,也要善于利用评价结果,为教师专业发展提供持续动力。

注释:

①在实际研究中,如果 $ICC > 0.059$,就有必要采用多层线性模型进行分析。

参考文献:

- [1] 刘振天,胡更新,赵志强. 新时代高等教育高质量发展的理论意涵与多维路径[J]. 湖北社会科学,2023(9):146-152.
- [2] 尹天光. 基于大数据分析的高校教师教学发展性评价体系构建研究[J]. 教育理论与实践,2021(27):42-45.
- [3][4] 陈明学,郑 锋. 发展性教师教学质量评价的创新与实践[J]. 中国大学教学,2017(5):78-80+88.
- [5] 郭丽君. 高校教学评价:基于组织制度的分析[J]. 高等教育研究,2017(7):78-82.
- [6] 杨 平. 发展性教学评价理念下应用型院校课堂教学质量评价的现存问题与出路[J]. 教育与职业,2019(1):59-62.

- [7] 陈春莲,唐 忠. 教师教学评价体系的构建与实施——基于“五维一体”发展性评价的改革思路[J]. 中国高校科技,2020(10):29-32.
- [8] 赵 锋. 新时代大学教师教学评价:反思与重构[J]. 高校教育管理,2023(4):114-124.
- [9] MILLER - YOUNG J, SINCLAIR M, FORGIE S. Teaching Excellence and how it is Awarded: A Canadian Case Study [J]. Canadian Journal of Higher Education, 2020(1):40-52.
- [10][23] 赵 锋. 研究型大学教师教学评价制度研究——建设一流本科教育的视角[J]. 大学教育科学,2021(5):77-86.
- [11] 张 傲,廖睿智,陈奕含,等. 基于分数膨胀现象的教学评价改进方法[J]. 高教发展与评估,2023(6):32-41+120-121.
- [12] 包水梅,陈嘉诚. 政策工具视角下我国“双一流”建设高校教学评价政策偏好研究[J]. 高校教育管理,2022(5):40-51.
- [13] 臧琰琰. 大学教师评价的理论遵循和应然选择[J]. 黑龙江高教研究,2021(4):85-90.
- [14] 田贤鹏,姜淑杰. 高质量发展视域下高校学生评教:问题表征与制度变革[J]. 现代教育管理,2022(9):74-83.
- [15] 卫建国,汤秋丽. 新时代高校教师教学评价改革与创新论析[J]. 黑龙江高教研究,2023(2):33-37.
- [16] 周群英,刘晓雪. 教学学术发展与大学教师教学评价制度创新[J]. 当代教育科学,2019(1):35-38.
- [17][20] 苏 肖,冯玉芳. 论美国高校教师教学评价[J]. 河北师范大学学报(教育科学版),2010(12):109-112.
- [18] 宋 焯,陈金芳. 美国高校教师同行教学评价的原则、程序及启示[J]. 中国高等教育,2022(17):62-64.
- [19] 叶玲娟. 以教学学术评价教学型教授何以可能? ——基于高校教学型教授职称评审条件的分析[J]. 国家教育行政学院学报,2022(3):87-95.
- [21] 刘隽颖. 美国一流研究型大学教师教学评价中的基本原则——以加州大学系统为例[J]. 清华大学教育研究,2020(4):88-96.
- [22] 李 超. 高校学生评价变革进展述评[J]. 外国教育研究,2018(7):105-117.
- [24][34] 梅雄杰,计国君. 新时代高校教师教学评价体系:人性审思与重构之向[J]. 中国高教研究,2023(12):40-47.
- [25] 牟智佳,刘珊珊,陈明选. 循证教学评价:数智化时代下高校教师教学评价的新取向[J]. 中国电化教育,2021(9):104-111.
- [26] 蔡 蕾. 基于OKR模式的我国一流大学建设高校人才评价改革路向研究[J]. 教育发展研究,2021(7):7-12.
- [27] 刘 理,宁立伟. 高校学生评教的制度正义及其实现[J]. 中国高教研究,2022(8):84-89.
- [28] 贺武华. 以学为中心的高校“教”与“学”质量评价体系改进[J]. 江苏高教,2019(3):21-25.
- [29] 苏 强,蔡晓卫. 效率还是价值:高校教师评价制度的逻辑

- 辑困境及其张力调适[J]. 华东师范大学学报(教育科学版), 2023(6):133-141.
- [30][38]刘秀梅,张雪遥. 高校教师教学质量评价审思[J]. 高教探索, 2022(4):56-60+82.
- [31]李正元,陈嘉诚. 高等教育高质量发展视角下大学教学评价制度的改革进路[J]. 兰州大学学报(社会科学版), 2022(6):118-127.
- [32]秦玉友,赵忠平,曾文婧. 义务教育教师教学工作时间结构研究——基于全国10省20市(县)的数据[J]. 教师教育研究, 2017(4):39-45.
- [33]STROEBE W. Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis [J]. Basic and Applied Social Psychology, 2020(4):276-294.
- [35]谢宇. 回归分析[M]. 修订版. 北京: 社会科学文献出版社, 2013:301.
- [36]景丽珍,杨贞兰. 同事关系对高校教师工作绩效的影响[J]. 高等教育研究, 2013(5):39-45.
- [37]陶政宇,李楠. 高校教师科研对教学反哺性的实证研究:基于全国63所高校的数据检验[J]. 黑龙江高教研究, 2023(6):36-41.
- [39]赵颖,哈巍. 分数膨胀的影响因素与学生选课策略[J]. 清华大学教育研究, 2020(6):63-74.
- [40]游森,秦建平,王逸尘. 卓越教师教学行为评价指标研究[J]. 中国教育学刊, 2021(12):72-78.
- [41]钱梦岑. 学生评教对高校教师教学行为的影响——基于国内某高校学生评教结果的实证研究[J]. 中国高教研究, 2018(10):81-86.

The Effect of Teaching Evaluation System on Teachers' Teaching Performance: Based on 63 Chinese Colleges and Universities

TAO Zheng-yu, LI Nan

(Capital University of Economics and Business, Beijing 100070, China)

Abstract: Teaching quality is the foundation of higher education. In order to improve the quality of teaching, it is necessary to evaluate teachers' teaching work effectively. Although the teaching evaluation of college teachers in China has been carried out for nearly 40 years, a few scholars have conducted empirical analysis on the effect of the current evaluation system and whether it can play a positive role in improving teachers' teaching performance. Based on the matching data of 63 colleges and universities in China, this paper uses hierarchical linear regression model to empirically analyze the composition of evaluation system and its impact on improving the teaching performance of college teachers. It is found that if the evaluation system has the following characteristics, it can improve the teaching performance of teachers. First, the evaluation subject should not choose comprehensive evaluation. Secondly, the evaluation content is the whole process of teaching. Thirdly, the evaluation method cannot only be qualitative evaluation. Last, the results of the evaluation should be reflected back to teachers on specific issues.

Key words: college teachers; teaching evaluation system; teaching performance; matching data; HLM