

# 评教还是评人：普通高校体育课程学生评教有效性研究

蒲毕文<sup>1, 2</sup>

(1.广东金融学院 体育教研部, 广东 广州 510521; 2.华南师范大学 体育科学学院, 广东 广州 510006)

**摘要:** 采用问卷调查、数理统计等方法对普通高校体育课程学生评教有效性进行实证研究。研究发现, 学生评教既是对“教”的评价, 也是对“人”的评价, 评教分数具有区分度和有效性。具体而言: (1) “一表多课”学生评教体系造成体育课程评教分数低于其他课程; (2) 教学因素影响评教结果, 教得好(教学效果好的)教师评教分数较好; (3) 非教学因素“人”影响学生评教分数, 青年教师、讲师、人缘好、颜值好、技能好的女性教师更易获学生好评, 即体育课程教师存在差异; (4) 体育课程教学项目评教也存差异, 三大球、三小球优于武术、健美操、瑜伽等项目。研究认为, 构建不同课程评教指标、实现评教专业化发展, 加强学生评教宣传, 助力评教公正公平, 用好评教结果, 实现以评促教, 从而有效提升普通高校体育课程教学质量。

**关键词:** 学校体育; 学生评教; 体育课程; 普通高校; 有效性

中图分类号: G807 文献标志码: A 文章编号: 1006-7116(2024)06-0125-08

## Evaluation of teaching or individuals: A study on the effectiveness of student evaluation of teaching in physical education courses of college and university

PU Biwen<sup>1, 2</sup>

(1. Department of Physical Education, Guangdong University of Finance, Guangzhou 510521, China;

2. School of Physical Education, South China Normal University, Guangzhou 510006, China)

**Abstract:** This empirical study uses questionnaire surveys and statistical analysis to investigate the effectiveness of student evaluations in physical education courses at colleges and universities. The findings reveal that student evaluation of teaching assess both the "teaching" and the "individual", and evaluation scores demonstrate differentiation and effectiveness. Specifically: (1) the "one form for multiple courses" evaluation system leads to lower scores for physical education compared to other courses; (2) teaching factors influence evaluation results, with teachers who perform well receiving higher scores; (3) non-teaching factors, such as the teacher's youth, rapport, attractiveness, and skills, impact evaluation scores, showing differences among physical education instructors; (4) there are discrepancies in evaluation among different physical education projects, with team sports and racket sports outperforming martial arts, aerobics, and yoga. Thus, it is recommended to develop differentiated evaluation indicators for various courses, promote specialized development of evaluations, enhance publicity around student evaluations, ensure fairness and justice in evaluations, utilize evaluation results to improve teaching, and to ultimately enhance the quality of physical education courses in general universities.

**Keywords:** school physical education; student evaluation of teaching; physical education courses; college and university; effectiveness

学生评教是高校教师教学评价和教学质量保障制度的重要组成部分<sup>[1]</sup>。对教师的教学质量或教学效果进

行评价涵盖普通高校的所有学生、教师、课程, 评教数据较为完善, 用于反映整体教学质量<sup>[2-3]</sup>。普通高校

收稿日期: 2024-04-28

基金项目: 教育部人文社会科学研究项目(20YJCZH128); 广东省教育科学规划课题(2024GXJK144)。

作者简介: 蒲毕文(1982-), 男, 副教授, 博士研究生, 硕士生导师, 研究方向: 体育人文社会学。E-mail: 120447483@163.com

体育课程是以培养学生体育兴趣、掌握运动技能、发展运动特长、促进终身锻炼习惯养成<sup>[4]</sup>为目标,相较于普通本科其他课程教学,具有较强的实践性、示范性、身体参与活动等特点。然而,高校教学评价量表造成学生评教行为偏差、信息失真、信度效度较低<sup>[1, 5-7]</sup>,同时存在学生对评教认知不足,难以区分是对教学还是对教师评价等问题<sup>[8]</sup>。评教结果广泛运用于教师晋升、职称评审,对教师的教学工作有着重要影响<sup>[9]</sup>,导致教师的教学工作过于注重评教结果,忽视教学本质,影响体育课程的教学质量<sup>[10]</sup>。如何有效评价体育课程教学质量,体育课程学生评教是否有效可信,仍有待商榷。尤其是在体育课程评教中存在的具体问题:如“一表多课”学生评教体育课程是否有效(即体育课程评教是否与其他课程有区别)?教得好不如教师受学生欢迎好(即评课还是评人)?体育课程教师是否也有差异(评价的区分度)?体育课程不同运动项目课程评教是否存在差异?针对以上问题,本研究对广东地区 15 所高校(覆盖三批次录取到一批次录取学生)进行调研,探讨普通高校体育课程评教的有效性,为推动高校体育教学高质量发展提供参考建议。

## 1 研究设计

### 1.1 数据来源

通过与学生教师访谈、梳理各学校评教指标及借鉴相关研究文献等,制定出包括学生特征、体育课程评教、理论课程评教、实践课程评教、体育教师特征、体育教师教学因素、学生及体育课程评教认知等 7 个方面,共计 56 个选项的调查问卷(见表 1、表 2)。变量  $X$  上标对应类别序号,下标表示类别中的变量序号,比如  $X_4^1$  表示第一类学生特征的第 4 个问题,依此类推。

在数据获取方面,通过纸质问卷和网络问卷相结合的方式对广东地区 15 所非体育专业的普通本科层次学校(以下简称“普通高校”)学生进行发放。最终,共发放纸质问卷 800 份,回收 756 份,收到网络问卷 2 016 份,两者回收共计 2 772 份。剔除问卷关键变量缺失和全部答案统一的无效问卷,最终确定 1 986 份为实际有效问卷。问卷信度检验显示 Cronbach's  $\alpha$  系数为 0.932;效度检验 KMO 值为 0.962 且 Bartlett 球形度检验  $P < 0.001$ ,可见问卷信效度满足社会学调研要求<sup>[11]</sup>。具体问卷题项表述、变量赋值情况等(李克特 5 级评分)见表 1、表 2。

### 1.2 研究思路

研究主要探讨普通高校“一表多课”学生课程评教下体育课程评教是否与其他课程评教存在差异,学生是对教师教学的评价还是教师的评价以及不同运动

项目课程的差异性。具体思路及步骤:一是采用单因素方差分析对体育课程与其他课程(包括理论课程、实践课程)差异性进行检验,以验证“一表多课”学生体育课程评教与其他课程的差异。二是采用参数检验分析高分评教师和低评教师在教学因素方面是否存在差异,以验证学生评教是否是对教师教学的评价,即回答学生评教是否是对“教”的评价。三是采用回归模型并控制教学因素,对教师特征和学生特征等“人”的因素如何影响评教进行分析,以探讨学生评教中是否受“人”因素影响。

表 1 学生特征指标( $X^1$ )情况( $n=1986$ )

变量	情况
$X_1^1$ 学校类型	一本 27.62%, 二本 46.87%, 三本 15.51%
$X_2^1$ 就读年级	大四 12.52%, 大三 19.52%, 大二 25.29%, 大一 42.67%
$X_3^1$ 学生性别	男生 34.84%, 女生 65.16%
$X_4^1$ 专业类型	艺术 6.71%, 理科 44.81%, 文科 48.39%
$X_5^1$ 项目类别	三大球 30.72%, 三小球 32.83%, 其他项目 36.46%

表 2  $X^2$ 、 $X^3$ 、 $X^4$ 、 $X^5$ 、 $X^6$ 、 $X^7$  等变量情况( $n=1986$ )

变量	变量表述	平均值	标准差
体育课程评教	$X_1^2$ 课前准备	1.262	0.510
	$X_2^2$ 讲课认真程度	1.238	0.498
	$X_3^2$ 答疑和辅导	1.241	0.506
	$X_4^2$ 基本概念、原理讲解清晰	1.258	0.513
	$X_5^2$ 讲授内容重点突出	1.260	0.515
	$X_6^2$ 讲授内容与实践联系密切	1.329	0.580
	$X_7^2$ 教材或讲义适用程度	1.355	0.608
	$X_8^2$ 提供学习资源丰富程度	1.313	0.588
	$X_9^2$ 善于启发引导,沟通良好	1.262	0.533
	$X_{10}^2$ 学生能力培养	1.265	0.541
	$X_{11}^2$ 课后收获及推荐	1.279	0.564
理论课程评教	$X_1^3$ 课前准备	1.360	0.637
	$X_2^3$ 讲课认真程度	1.353	0.640
	$X_3^3$ 答疑和辅导	1.372	0.661
	$X_4^3$ 基本概念、原理讲解清晰	1.391	0.675
	$X_5^3$ 讲授内容重点突出	1.404	0.679
	$X_6^3$ 讲授内容与实践联系密切	1.438	0.694
	$X_7^3$ 教材或讲义适用程度	1.433	0.705
	$X_8^3$ 提供学习资源丰富	1.408	0.682
	$X_9^3$ 善于启发引导,沟通良好	1.410	0.699
	$X_{10}^3$ 注意学生能力培养	1.403	0.701
	$X_{11}^3$ 课后收获及推荐	1.427	0.718

(续表2)

X <sup>4</sup> 实践 课程 评教	X <sub>1</sub> <sup>4</sup> 课前准备	1.413	0.673
	X <sub>2</sub> <sup>4</sup> 讲课认真程度	1.404	0.677
	X <sub>3</sub> <sup>4</sup> 答疑和辅导	1.410	0.689
	X <sub>4</sub> <sup>4</sup> 基本概念、原理讲解清晰	1.434	0.714
	X <sub>5</sub> <sup>4</sup> 讲授内容重点突出	1.451	0.710
	X <sub>6</sub> <sup>4</sup> 讲授内容与实践联系密切	1.452	0.722
	X <sub>7</sub> <sup>4</sup> 教材或讲义适用程度	1.480	0.735
	X <sub>8</sub> <sup>4</sup> 提供学习资源丰富	1.448	0.724
	X <sub>9</sub> <sup>4</sup> 善于启发引导,沟通良好	1.442	0.712
	X <sub>10</sub> <sup>4</sup> 注意学生能力培养	1.426	0.705
	X <sub>11</sub> <sup>4</sup> 课后收获及推荐	1.463	0.732
X <sup>5</sup> 体育 教师 特征	X <sub>1</sub> <sup>5</sup> 性别 <sup>1)</sup>		
	X <sub>2</sub> <sup>5</sup> 年龄特征	1.735	0.557
	X <sub>3</sub> <sup>5</sup> 职称 <sup>2)</sup>	3.313	1.097
	X <sub>4</sub> <sup>5</sup> 教师人缘	3.454	1.303
	X <sub>5</sub> <sup>5</sup> 教师颜值	2.805	1.288
	X <sub>6</sub> <sup>5</sup> 运动技能水平	4.142	1.236
X <sup>6</sup> 教学 因素	X <sub>1</sub> <sup>6</sup> 课程教学难度	2.468	0.716
	X <sub>2</sub> <sup>6</sup> 教师迟到情况	1.236	0.484
	X <sub>3</sub> <sup>6</sup> 考勤严格程度	3.544	0.952
	X <sub>4</sub> <sup>6</sup> 答疑指导及交流情况	4.125	1.020
	X <sub>5</sub> <sup>6</sup> 课堂组织熟练程度	4.356	0.814
	X <sub>6</sub> <sup>6</sup> 讲解示范	4.186	0.878
	X <sub>7</sub> <sup>6</sup> 教学方式适合你吗	2.065	0.893
	X <sub>8</sub> <sup>6</sup> 训练项目知识拓展	3.296	1.054
X <sup>7</sup> 学生 收获及 对体育 课程评 教认知	X <sub>1</sub> <sup>7</sup> 学习收获(教学效果)	3.843	0.954
	X <sub>2</sub> <sup>7</sup> 您认为学校评教量表能 反映教学水平吗	3.659	0.860
	X <sub>3</sub> <sup>7</sup> 体育课程评教成绩	84.444	9.301

1)女教师占比 45.57%,男教师占比 54.43%; 2)因 410 份教师无职称,该项样本量为 1576

## 2 结果与分析

### 2.1 不同类型课程学生评教差异性分析

已有研究发现,单一量表对不同类型课程评价结果存在显著差异<sup>[12]</sup>,造成学生课程评价失真。体育课程以身体练习为主要手段,通过合理的体育教育和科学的体育锻炼过程,达到增强体质、增进健康和提高

体育素养为主要目标,且具有较强的实践性。为进行有效对比,将其他课程分为理论课程和实践课程。理论课程是指采用课堂教学,以课程知识点为主,教师讲解传授,学生识记和理解知识点一类课程,包括数学分析、中国近代史、保险学、宏观经济学、马克思主义哲学等课程。实践课程是指采用场景教学,以操作为主,教师实现示范、学生学习技能的一类课程,包括机械制图、自动化生产线实训、丰田维修技术、计算机实操、Stata 应用等。采用单因素方差分析检验体育课程、理论课程、实践课程评价是否存在显著性差异,具体结果见表 3。

通过表 3 可知,体育课程、理论课程、实践课程等评教结果显示,课前准备、认真程度、答疑辅导、讲解、内容社会实践练习、教材讲义选择、学习资源、教学启发、学生沟通、能力培养、收获、课程推荐等均呈现出显著性差异( $P<0.05$ ),体育课程评教分值最低,实践课程评教分值最高。由此可见,体育课程、理论课程及实践课程评教存在显著差异,不利于体育课程获取学生高分评教。

### 2.2 “教”的因素对学生评教影响

假设学生评教是对教学因素的评价,故提出以下两个研究假设:假设一、教师在体育课程学生评教中获取高分且教学因素指标  $x^6$  相对优越,说明学生评教是对“教学”的评价;假设二、在体育课程学生评教中获取高分和低分的教师,教学因素指标  $x^6$  不存在显著差异,或高分老师  $x^6$  评价成绩相对较差,说明学生评教不是对“教学”的评价。

为检验以上研究假设,将原有的数据分为高分区和低分区。为了使区分更为显著,去除中间的 386 份,将高分和低分各 800 份进行对比分析,采用方差齐性和均值  $t$  检验进行分析,具体结果详见表 4。

表 3 体育课程、理论课程、实践课程学生评教差异性分析表( $n=1\ 986$ )

变量	$M\pm SD$			$F$	$P$
	$X^2$	$X^3$	$X^4$		
$X_1^{2/3/4}$	1.262±0.510	1.360±0.637	1.413±0.673	31.025	0.000
$X_2^{2/3/4}$	1.238±0.498	1.353±0.640	1.404±0.677	38.973	0.000
$X_3^{2/3/4}$	1.241±0.506	1.372±0.661	1.410±0.689	40.216	0.000
$X_4^{2/3/4}$	1.258±0.513	1.391±0.675	1.434±0.714	40.643	0.000
$X_5^{2/3/4}$	1.260±0.515	1.404±0.679	1.451±0.710	47.820	0.000
$X_6^{2/3/4}$	1.329±0.580	1.438±0.694	1.452±0.722	20.033	0.000
$X_7^{2/3/4}$	1.355±0.608	1.433±0.705	1.480±0.735	16.813	0.000
$X_8^{2/3/4}$	1.313±0.588	1.408±0.682	1.448±0.724	21.417	0.000
$X_9^{2/3/4}$	1.262±0.533	1.410±0.699	1.442±0.712	42.986	0.000
$X_{10}^{2/3/4}$	1.265±0.541	1.403±0.701	1.426±0.705	35.181	0.000
$X_{11}^{2/3/4}$	1.279±0.564	1.427±0.718	1.463±0.732	41.050	0.000

表 4 方差齐性检验及均值  $t$  检验 ( $n=1\ 986$ )

对比指标	方差齐性检验				均值 $t$ 检验		
	$F$	Sig.	$t$	df	Sig.	均值差	
$X_1^6$	假设方差相等	1.106	0.293	-1.159	1 269.00	0.247	-0.04
	假设方差不相等			-1.159	1 145.73		
$X_2^6$	假设方差相等	147.280	0.000	-6.209	1 269.00	0.000	-0.15
	假设方差不相等			-6.209	1 198.32		
$X_3^6$	假设方差相等	5.076	0.024	2.689	1 269.00	0.007	0.13
	假设方差不相等			2.689	1 195.36		
$X_4^6$	假设方差相等	56.889	0.000	-8.643	1 269.00	0.000	-0.43
	假设方差不相等			-8.643	1 267.92		
$X_5^6$	假设方差相等	56.723	0.000	-7.732	1 269.00	0.000	-0.30
	假设方差不相等			-7.732	1 092.51		
$X_6^6$	假设方差相等	31.869	0.000	-3.193	1 269.00	0.000	-0.24
	假设方差不相等			-3.193	1 210.41		
$X_7^6$	假设方差相等	1.379	0.241	-0.588	1 269.00	0.557	-0.03
	假设方差不相等			-0.588	1 103.27		
$X_8^6$	假设方差相等	0.003	0.959	-2.184	1 269.00	0.029	-0.12
	假设方差不相等			-2.184	1 028.73		

通过表 4 可知, 方差齐性检验中  $X_1^6$ 、 $X_7^6$ 、 $X_8^6$  的  $F$  值统计量  $P>0.05$ , 属于方差齐性的情形, 而  $X_2^6$ 、 $X_3^6$ 、 $X_4^6$ 、 $X_5^6$ 、 $X_6^6$  的  $P<0.05$ , 属于方差不齐性。均值  $t$  检验结果显示,  $X_1^6$  课程难度、 $X_7^6$  教学方式适合等呈现显著性差异,  $X_2^6$  考勤严格程度呈现显著性差异且低分组优于高分组, 教师迟到、学生答疑指导、课程组织熟练程度、讲解示范清晰程度、训练项目拓展水平等评分高的教师均优于低评分教师, 说明高分组教师通过教学等吸引学生, 低分组通过考勤留住学生。由此可见, 学生评教是对教学的评价, 即学生评教是有效的。

### 2.3 “人”的因素对学生评教影响

在学生评教过程中, 非教学因素(如教师人缘、教师形象等)是否会对学生评教产生影响呢? 以下检验“人”的因素对学生评教的影响。

#### 1) 教师人缘对学生评教的影响。

“教得好, 不如人缘好”。吴立军等<sup>[8]</sup>关注教师人缘对影响学生评教的影响。现实中学生对体育教师熟悉程度远高于学校评教指标, 存在学生依据自身对体育教师的认知和熟悉程度进行评价<sup>[10]</sup>。体育教师群体中也常常抱怨“教的好不如与学生关系好”, 针对教师人缘是否对学生评教产生影响, 分别涉及“人缘”变量对学生评教影响模型和“人缘”与“教学效果”组合效应模型, 探讨教学效果和人之间的影响, 具体模型如下:

模型一引入“人缘”变量,  $y_i$  为学生评教成绩, 控制教学因素后将教学因素、学习收获等体育课程评教成绩进行汇总, 生成新的变量, 构成  $X_i^6 = \sum_{m=1}^6 X_{mi}^6$ ,  $X_i^7$  为教学效果,  $X_i^4$  为教师人缘, 具体模型定义如下:

$$\text{模型一: } y_i = a + a_1 X_i^6 + a_2 X_i^7 + a_3 X_i^4 + e_i \quad (1)$$

模型二是在控制教学因素后, 以教学效果差、人

缘好为基准, 构建教学效果好, 人缘好; 教学效果好, 人缘差; 教学效果差, 人缘好等 3 类教学效果与人缘组合的虚拟变量, 讨论教学效果与人缘组织如何影响学生评教, 具体模型如下:

$$\text{模型二: } y_i = a + a_1 X_i^6 + a_2 X_i^7 + \beta_1 d_{1i} + \beta_2 d_{2i} + \beta_3 d_{3i} + e_i$$

$$d_1 = \begin{cases} 1 & \text{教学效果好, 人缘好} \\ 0 & \text{其他} \end{cases} \quad d_2 = \begin{cases} 1 & \text{教学效果好, 人缘差} \\ 0 & \text{其他} \end{cases}$$

$$d_3 = \begin{cases} 1 & \text{教学效果差, 人缘好} \\ 0 & \text{其他} \end{cases} \quad (2)$$

本研究数据为截面数据, 为此采用 Robust 稳健估计, 降低模型自相关对模型评估的影响, 模型一、二回归分析结果详见表 5、6。

表 5 教师人缘对学生评教影响回归分析结果 ( $n=1\ 986$ )

	回归系数	标准误	$t$	$P$	95% CI
常数	66.129	1.704	38.799	0.000	62.788~69.470
教学因素 $X_i^6$	0.197	0.078	2.533	0.011	0.045~0.350
教学效果 $X_i^7$	1.896	0.285	6.662	0.000	1.338~2.454
教师人缘 $X_i^4$	1.749	0.185	9.438	0.000	1.386~2.112
$R^2=0.187$ , 调整 $R^2=0.186$ , $F=109.842(P=0.000)$ , $D-W$ 值=1.902					

表 6 教学效果、人缘组合虚拟变量对学生评教结果影响分析 ( $n=1\ 986$ )

	回归系数	标准误	$t$	$P$	95% CI
常数	66.477	1.839	36.141	0.000	62.872~70.082
教学因素	2.289	0.628	3.645	0.000	1.058~3.520
教学效果	2.615	0.337	7.769	0.000	1.955~3.275
$d_1$	5.496	2.688	2.045	0.041	0.229~10.764
$d_2$	1.820	0.515	3.534	0.000	0.810~2.829
$d_3$	1.977	0.733	2.697	0.007	3.414~-0.540
$R^2=0.163$ , 调整 $R^2=0.161$ , $F=62.868(P=0.000)$ , $D-W$ 值=1.899					

通过表 5 可知，模型一的教师教学因素和人缘因素变量均在 1% 水平显著，学习收获则在 5% 水平显著，且回归系数均为正数，表明“人缘”因素与学生评教结果显著正相关，说明“人缘”因素对学生评教产生影响。

表 6 数据显示， $d_1$ 、 $d_2$ 、 $d_3$  虚拟变量在以“教学效果差、人缘差”为基准，采用 Robust 稳健标准误回归分析可知， $d_1$ 、 $d_2$ 、 $d_3$  回归系数显著为正，其中  $d_1$ 、 $d_3$  显著水平在 1%， $d_2$  显著水平为 5%，说明与基准类型的“教学效果差、人缘差”，说明  $d_1$ 、 $d_2$ 、 $d_3$  评教成绩都显著好于基础类型。通过对回归系数分析发现  $d_1 > d_3 > d_2$ ，学生评教成绩从高到底依次为“教学效果好、人缘好” > “教学效果差、人缘好” > “教学效果好、人缘差” > “教学效果差、人缘差”。由此可知“教学效果好、人缘好”的教师学生评教成绩好，“教学效果差、人缘差”的教师学生评教成绩差，“教学效果差、人缘好”的教师学生评教成绩好于“教学效果好、人缘差”的教师，说明在学生评教活动中人缘好强于教学效果好。

2) 教师严格程度对学生评教的影响。

学生评教成绩与教师晋升、绩效考核等紧密联系<sup>[13]</sup>，由此造成教师不敢管、不管严，严师带来学生报复性低分评教<sup>[14]</sup>。现实调研发现，学生认为教师认真、严格要求并非会带来学生评教低分。采用考勤等手段约束学生，教师自身教学敷衍，教学效果差的老师容易被学生低分评价。到底是严师出低分还是教学效果差造成的低分评价呢？以“ $X_{3i}^6$  教师考勤严格”作为单独评教指标，在  $X_i^6$  构建组合变量是剔除  $X_{3i}^6$ ，加入教学效果  $X_i^7$ ， $y_i$  为学生评教成绩，构建模型三检验严师出低分现实情况；以“ $X_{3i}^6$ ”和“ $X_i^7$ ”组合成虚拟变量，以“考勤不严，教学效果差”为基准，构建模型四分析严师和教学效果是如何影响评教，具体模型如下：

$$\text{模型三： } y_i = a + a_1 X_i^6 + a_2 X_i^7 + a_3 X_{3i}^6 + e_i \quad (3)$$

$$\text{模型四： } y_i = a + a_1 X_i^6 + a_2 X_i^7 + \beta_1 d_{1i} + \beta_2 d_{2i} + \beta_3 d_{3i} + e_i$$

$$d_1 = \begin{cases} 1 & \text{考勤严格, 教学效果好} \\ 0 & \text{其他} \end{cases} \quad d_2 = \begin{cases} 1 & \text{考勤严格, 教学效果差} \\ 0 & \text{其他} \end{cases}$$

$$d_3 = \begin{cases} 1 & \text{考勤不严, 教学效果好} \\ 0 & \text{其他} \end{cases} \quad (4)$$

通过表 7 可知，模型三的教师教学因素、学习收获变量均在 1% 水平显著且回归系数均为正；考勤严格程度则在 5% 水平显著且回归系数均为负数，表明“严师”因素与学生评教结果显著负相关，说明“严师”因素对学生评教成绩产生负向影响，“严师出低分”。

模型 4 以“考勤不严，教学效果差”为基准，表 8 数据显示， $d_1$ 、 $d_2$ 、 $d_3$  均在 1% 水平显著，其中  $d_1$ 、 $d_3$  且回归系数为正， $d_2$  回归系数为负，即“考勤严格，

教学效果好”“考勤不严，教学效果好”的教师学生评教成绩均优于基准类型教师，“考勤严格，教学效果差”的教师学生评教分数显著劣于基准类型。体育教师教学效果好，学生评教分数较好，其中“考勤不严、教学效果好”教师学生评教优于“考勤严格、教学效果好”教师；“考勤不严、教学效果差”教师学生评教优于“考勤严格、教学效果不好”教师。由此可见，体育课程教学中严格教师出低分现象确实存在。

表 7 教师严格程度对学生评教影响回归分析结果 ( $n=1\ 986$ )

	回归系数	标准误	t	P	95% CI
常数	64.585	1.763	36.640	0.000	61.131~68.040
教学因素	0.447	0.088	5.057	0.000	0.274~0.620
教学效果	2.830	0.273	10.352	0.000	2.294~3.366
考勤	-0.207	0.218	-2.947	0.004	-0.304~-0.121

$R^2=0.149$ , 调整  $R^2=0.148$ ,  $F=80.676(P=0.000)$ ,  $D-W$  值=1.905

表 8 教学效果、严格程度组合虚拟变量对

学生评教结果影响分析 ( $n=1\ 986$ )

	回归系数	标准误	t	P	95% CI
常数	64.507	1.830	35.250	0.000	60.920~68.094
教学因素	0.454	0.089	5.132	0.000	0.281~0.628
× 考勤					
教学效果	2.662	0.304	8.751	0.000	2.065~3.258
$d_1$	0.305	2.447	6.682	0.000	0.181~0.571
$d_2$	-2.692	1.676	-3.607	0.008	-5.977~-0.592
$d_3$	0.549	2.780	6.704	0.000	0.079~1.077

$R^2=0.151$ , 调整  $R^2=0.149$ ,  $F=48.747(P=0.000)$ ,  $D-W$  值=1.904

3) 学生、教师个体特征对学生评教的影响。

学生评教追踪研究发现教师评教分数相对固化，评教分数较好的教师每次评教分数相对较高，评教分数较差的教师每次评教分数相对较低<sup>[15]</sup>。在普通高校体育课程评教过程中，好评给了谁、差评给了谁；学生也存在给予全体教师高分或者低分的实际情况，盲目打分，即谁打了高分，谁打了低分。在教学因素  $X^6$  和教学效果  $X^7$  基础上引入教师和学生个体特征，其中模型五引入学生个体特征，如  $X_{1i}^1$  (学校类型)、 $X_{2i}^1$  (年级)、 $X_{3i}^1$  (性别)、 $X_{4i}^1$  (专业)、 $X_{5i}^1$  (体育课程项目类别) 等，因大学本科仅大一、大二开设体育课程教学，仅对大一和大二进行比对，分析哪些类型的学生给了好评、哪些类型的学生给了差评。模型六引入教师个体特征，如  $X_{1i}^5$  (性别)、 $X_{2i}^5$  (年龄)、 $X_{3i}^5$  (职称)、 $X_{4i}^5$  (人缘)、 $X_{5i}^5$  (颜值)、 $X_{6i}^5$  (技能水平)、 $X_{7i}^5$  (课程难度)，分析哪些类型的教师获得好评，哪些类型的教师被打低分。具体模型如下：

$$\text{模型五： } y_i = a + a_1 X_i^6 + a_2 X_i^7 + a_3 X_{1i}^1 + a_4 X_{2i}^1 + a_5 X_{3i}^1 + a_6 X_{4i}^1 + a_7 X_{5i}^1 + e_i \quad (5)$$

$$\text{模型六: } y_i = a + a_1X_i^6 + a_2X_i^7 + \beta_1X_{li}^6 + \beta_2X_{li}^5 + \beta_3X_{2i}^5 + \beta_4X_{3i}^5 + \beta_5X_{4i}^5 + \beta_6X_{5i}^5 + e_i \quad (6)$$

通过表 9 可知,从学校类型来看,以本科一批为基准,本科一批、本科二批和本科三批存在显著差异( $P < 0.01$ )且回归系数为正,由此可见本科一批学生评教成绩最低,本科三批相对较好,最好为本科二批。从年级来看,以大学一年级为基准,大一、大二存在差异,但无显著性差异( $P > 0.05$ )且回归系数为负,大二学生在学生评教中给分更低。从性别来看,以男性学生为基准,男、女学生存在显著性差异( $P < 0.05$ )且回归系数为正,说明女生更愿意给出高分。从专业类型来

看,以文科学生为基准,文科、理科、艺术类均存在显著性差异( $P < 0.05$ )且回归系数为正,说明文科类学生评教成绩相对较低,艺术类学生在体育课程中评教成绩最高。从体育课程项目类型来看,以三大球为基准,三大球和三小球并无显著性差异( $P > 0.05$ ),武术、健美操、瑜伽等表演类存在显著性差异( $P < 0.05$ )且回归系数为负,说明表演类等体育项目教学成绩相对较低。由此可见,录取批次高、年级高、艺术类学生、学习武术、健美操、瑜伽等表演类课程的男生,容易在体育课程评教中打出低分。

表 9 学生个体特征对评教结果影响回归分析结果 ( $n = 1986$ )

	回归系数	标准误	<i>t</i>	<i>P</i>	95% CI
常数	61.166	1.888	32.398	0.000	57.466~64.867
教学因素	0.465	0.089	5.248	0.000	0.291~0.638
教学效果	2.809	0.267	10.508	0.000	2.285~3.333
一批次录取【参照项】学校类型					
二批次录取	2.984	0.529	5.645	0.000	1.948~4.021
三批次录取	1.862	0.697	2.671	0.008	0.496~3.229
大一【参照项】年级					
大二	-0.952	-0.509	-1.869	0.062	-1.950~0.046
男【参照项】性别					
女	0.898	0.412	2.178	0.029	0.090~1.706
文科【参照项】专业					
理科	0.194	0.417	3.465	0.042	-0.623~1.011
艺术类	1.508	1.701	3.886	0.036	-1.827~4.842
三大球【参照项】体育课项目					
三小球	0.199	0.458	0.434	0.664	-0.699~1.096
武术、健美操、瑜伽等	-1.438	0.476	-3.023	0.003	-2.370~-0.506

$R^2 = 0.173$ , 调整  $R^2 = 0.168$ ,  $F = 26.255 (P = 0.000)$ ,  $D-W$  值 = 1.915

通过表 10 可知,从性别来看,以女教师为基准,男女教师间存在显著性差异( $P < 0.05$ )且回归系数为负,说明男教师被低分评价。从年龄来看,以青年教师为基准,中青年教师存在显著性差异( $P < 0.01$ ),青老教师存在显著性差异( $P < 0.05$ )且回归系数均为负数,同时年龄越大回归系数负值越大,可见年龄大的教师易被低分评价。从职称来看,以无职教师为基准,助教、讲师、副教授在 5%水平显著,教授在 10%水平显著,助教和教授回归系数为负且助教负值更大,说明助教更易被学生评教给差评。从教师人缘来看,以教师人缘很差为基准,均存在显著性差异( $P < 0.01$ )且回归系数

均为正,人缘越好回归系数越大,说明人缘差的教师易被学生给差评。从教师颜值来看,以教师颜值差为基准均存在显著性差异( $P < 0.05$ )且回归系数均为正,颜值越高回归系数越大,说明颜值差的教师易被学生给差评。从教师技能水平来看,以技能水平差为基准,中低水平无显著性差异( $P > 0.05$ ),高水平存在显著性差异( $P < 0.05$ ),说明中低技能水平的教师易被学生给差评,高水平技能教师评教成绩优于中低技能水平教师,课程难度则无显著性差异( $P > 0.05$ )。由此可见,老教师、助教、人缘差、颜值差、技能水平低的男教师更易被学生评教给低分。

表10 教师个体特征对评教结果影响回归分析结果( $n=1986$ )

	回归系数	标准误	$t$	$P$	95% CI
常数	67.646	2.460	27.504	0.000	62.825 ~ 72.467
教学因素	0.236	0.105	2.239	0.025	0.029 ~ 0.443
教学效果	1.761	0.303	5.816	0.000	1.168 ~ 2.354
男【参照项】教师性别					
女	-1.009	0.420	-2.404	0.016	-1.832 ~ -0.186
青年教师【参照项】教师年龄					
中年教师	-0.590	0.462	-1.276	0.002	-1.496 ~ 0.316
老年教师	-1.281	1.125	-1.139	0.055	-3.486 ~ 0.923
无职称【参照项】教师职称					
助教	-0.883	1.418	-0.623	0.033	-3.662 ~ 1.896
讲师	0.409	0.834	0.491	0.023	-1.225 ~ 2.043
副教授	0.181	0.910	0.199	0.042	-1.602 ~ 1.965
教授	-0.134	0.947	-0.142	0.087	-1.991 ~ 1.723
很差【参照项】教师人缘					
差	2.504	1.199	2.089	0.037	0.154 ~ 4.854
一般	5.571	1.210	4.605	0.000	3.199 ~ 7.942
较好	6.993	1.212	5.772	0.000	4.619 ~ 9.368
很好	6.032	1.280	4.712	0.000	3.523 ~ 8.541
很差【参照项】教师颜值					
差	1.007	0.724	2.774	0.006	0.589 ~ 2.426
一般	2.002	0.783	2.175	0.030	0.168 ~ 3.236
较好	2.398	0.910	2.953	0.003	0.557 ~ 4.239
很好	2.688	0.939	2.553	0.011	0.904 ~ 4.471
很差【参照项】教师运动技能水平					
差	0.599	0.527	-1.137	0.256	-1.631 ~ 0.433
一般	0.082	1.160	0.071	0.943	-2.192 ~ 2.356
较好	0.239	0.505	2.473	0.036	-0.751 ~ 1.228
很好	0.667	0.507	-3.316	0.008	-1.661 ~ 0.327
不难【参照项】课程难度					
一般	-0.360	0.691	-0.521	0.602	-1.715 ~ 0.995
很难	-0.667	0.507	-1.316	0.188	-1.661 ~ 0.327

$R^2=0.211$ , 调整  $R^2=0.200$ ,  $F=15.004(P=0.000)$ ,  $D-W$  值=1.906

### 3 结论及建议

普通高校“一表多课”学生评教中,体育课程评教分数显著低于理论课程和实践课程。就“教学”因素而言,教学效果较好的教师评教优于教学效果差的教师,说明普通高校体育课程评教具有区分度,是对教师教学的评价,学生评教有效。就非教学因素“人”而言,教师人缘显著影响学生评教,在现实体育课程评教中“拼人缘”“严师出低分”等现象确实存在;引入学生、教师个体特征研究发现,录取批次高、年级高、艺术类、学习武术、健美操、瑜伽等课程男生更易打低分,老教师、助教、人缘差、颜值差、技能水平低的男教师更易被学生评教给低分。非教学因素“人”影响学生评教,存在对“人”的评价,体育课程项目差异存在,三大球、三小球优于武术、健美操、瑜伽等表演性项目。

基于以上结论,提出如下建议:

第一,构建不同课程评教指标,实现评教专业化

发展。对“一表多课”存在的体育教师评教分数偏低的现实问题,依据课程特征制定差异化评教标准、提高体育课程教学质量、建立多维度的评教体系等,改善体育课程评教偏低现状。一是制定差异化评教标准,针对体育课程的特点,制定更符合其性质的评教标准,如增加评价体育课程教学方法的创新性、课程的互动性和参与度等评教内容,降低教材、讲义等评教内容。二是建立多维度的评价体系,构建一个包含学生、同行教师、教育专家等多方参与的评价体系,以更全面地评估和提升教学质量。

第二,加强学生评教宣传,助力评教公正公平。针对“拼人缘”“严师出低分”等问题,一是强化学生评教的教育和引导,让学生充分理解评教的重要性和目的,强调评价应基于教师教学内容、方法、态度等,而非个人喜好或人际关系。二是实现教师绩效评价多元化,不仅考虑学生评教结果,同时考虑教师的教学准备、课堂管理、教学创新以及同行评议。三是构建

合理的教学考勤政策,充分考虑考勤和学生评教政策的公正性,避免因放松考勤或评分方式而影响学生对教师的评价。

第三,用好评教结果,实现以评促教。研究结果显示学生评教分数具有区分性且整体有效,由此可见学生评教是提高体育课程教学质量、实现教学管理的有效手段。现阶段学生评教仍是教师绩效考核、职称晋升等证明手段,学生评教促使体育教师教学功能弱化,为此应有效利用学生评教结果,促进以评促教,如体育课程教学过程中低分教师(如分数后5%)应组织专家进行现场评教和诊断,提出针对性建议,指定优秀有经验的教师传帮带,提升教学水平。

### 参考文献:

- [1] 周继良. 高校学生评教行为偏差影响因素的实证研究——基于制度分析的视角[J]. 高等教育研究, 2018, 39(2): 59-72.
- [2] 叶赋桂, 田静, 罗燕. 美国高校教师评价的变革及其动因研究[J]. 教育学报, 2008(5): 74-81.
- [3] 郭丽君. 走向为教学的评价:地方高校教学评价制度探析[J]. 高等教育研究, 2016, 37(6): 68-73.
- [4] 谢燕歌, 洪浩. 普通高校公共体育课程教学指导思想探索与思考[J]. 北京体育大学学报, 2014, 37(1): 94-99.
- [5] 周继良, 龚放, 秦雍. 高校学生评教行为偏差及其与学科类别、学校类型和学业自评的关系——基于南京和常州十所高校的实证调查[J]. 高等教育研究, 2017, 38(10): 64-74.
- [6] 周继良, 秦雍. 高校学生评教行为偏差基本类型及其与学生相关背景特征的关系[J]. 复旦教育论坛, 2018, 16(6): 65-73.
- [7] 韩映雄, 周林芝. 学生评教的信度、效度、影响因素及应用风险[J]. 复旦教育论坛, 2018, 16(6): 74-81.
- [8] 吴立军, 田启波. 评“教”还是评“人”?!——基于学生评教的有效性研究[J]. 高教探索, 2020(8): 57-65.
- [9] 马浚锋, 罗志敏. 偏离、异化与重塑:学生评教本真性变奏——基于高校学生质性评教结果[J]. 高教探索, 2023(3): 48-55.
- [10] 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.
- [11] 风笑天. 社会研究方法[M]. 北京:中国人民大学出版社, 2013: 56.
- [12] 隗金水. 广州体育学院研究生课堂教学质量评价的研究[J]. 广州体育学院学报, 2011, 31(6): 98-102.
- [13] 王洪才. 高质量高等教育体系的基本内涵、主要特征与实践路径[J]. 现代教育管理, 2023(4): 1-9.
- [14] 李欣, 张威. “多元交互”在线教学评价指标体系构建的内涵、框架及路径[J]. 现代教育管理, 2022(10): 101-107.
- [15] 王玉刚, 柳兴国, 安强身. 学生评教的“廉价投票权”问题及治理[J]. 高教探索, 2015(8): 71-75.