

高校学生网络评教影响因素研究

聂亚菲, 刘国买

(福建工程学院 管理学院, 福建 福州 350118)

摘要: 抽取4个院系、12个专业, 约37 724个原始数据, 经过异常值处理后, 获得34 022个有效个案。在信度和效度检验的基础上, 采取均值相等性的稳健性检验和 Tamhane 方法进行方差不相等情况下的多重比较, 探索了学生网上评教影响因素。研究结果表明, 学生年级、学生成绩、教学班级数、教师年龄、授课门数、课程性质等非教学因素显著影响教师教学综合评价。

关键词: 网络评教; 综合评价; 影响因素

中图分类号: G647

文献标志码: A

文章编号: 1672-4348(2014)02-0176-04

Study on influential factors of students' on-line evaluations of teaching in universities

Nie Yafei, Liu Guomai

(School of Management, Fujian University of Technology, Fuzhou 350118, China)

Abstract: 37,724 raw data from 12 specialties of 4 departments in certain university are extracted, out of which 34,022 effective data are obtained after removing abnormal ones. On the basis of reliability and validity tests, multiple comparisons under unequal square variance are made by the robustness test and Tamhane test (based on equality mean). Influential factors affecting on-line teaching evaluations are explored. It is indicated that non-teaching factors such as students' grade, scores, classes, teachers' age, number of courses, and the nature of the courses significantly affect the comprehensive on-line evaluation of teaching.

Keywords: on-line teaching evaluation; comprehensive evaluation; influential factor

学生评教是教学质量监控体系的重要组成部分, 目的在于通过学生了解教师教学工作状态, 沟通教与学的信息, 促进教师内部竞争机制的运转, 改进教学方法, 提高教学质量。然而评价指标体系简单、制度粗糙, 评价结果滥用, 使学生网络评教有效性饱受质疑。

为更好地认识学生评教的利弊并通过控制影响因素来改进学生评教工作, 评教影响因素的研究一直是研究重点。部分研究肯定学生特征(主观偏好^[1]、学生态度^[2-4]、学生成绩、高低年级^[3])、课程特征(热门程度、教学条件^[5]、课程类别和课程性质^[3,5,6])、教师特征(教师学历、教师

职称^[5])对学生评教的影响。此外, 福克斯教授效应(教员的表达对学生评教具有压倒影响)^[7]、师生关系、评教时间安排、问卷评价标准体系等对学生评教也有较大的影响^[6]。也有研究认为学生评教受非教学因素的影响较小^[4], 年级差异和文理科差异对评教的影响不显著^[6]。

现有研究多为调查研究, 样本有限, 可信度不足, 且没有涉及信息化管理。本研究在教学管理信息化背景下进行大样本数据研究, 探索学生网络评教的影响因素, 以期丰富和弥补现有研究的不足。

收稿日期: 2013-09-10

基金项目: 教育部规划课题(FFB090523)

第一作者简介: 聂亚菲(1975-), 女(汉), 湖南衡阳人, 讲师, 硕士, 研究方向: 高等教育管理研究。

一、高校学生网络评教典型系统的评价

1. 指标体系

评价指标包括“为人师表,教书育人,严格管理”、“备课充分,精神饱满,语言生动,有感染力”等10个。每个指标都用五级定序变量来度量。

2. 单指标评价

学生对某门课程教师教学情况作出等级评价,后台程序按“好~100、较好~75、一般~50、较差~25、差~1”,将评价等级转化为“100、75、50、25、1”相应分值的数值变量。计算该指标的简单平均,作为该单项指标的对应分值,转化为百分数作为“满意度”。

3. 单个学生对某课程教学评价

所有指标的简单算术平均作为该学生对该教师某门课程教学的评价。

4. 任课教师某课程教学评价

所有学生某门课程的简单算术平均作为该教师某门课程的教学评价。

5. 任课教师教学评价

所有课程的简单算术平均为该任课教师教学评价。

由于简单、易行、操作性强,教学管理信息系统往往集成了学生网络评教系统,从而具有典型性。然而该评教系统评价指标简陋单一,难免忽略了许多主观和客观的信息。

(1)所有学生评教的简单算术平均掩盖了学生个体因素如评价态度、评价心理、个性特征等差异化因素对评价结果的影响。

(2)所有课程的简单算术平均掩盖了学生群体因素如年级、专业、班级规模等差异化因素所导致教学评价的差异。

(3)所有教师横向评价掩盖了开课重复次数、学期授课总门数等、课程因素如选修、必修、学科等差异化因素对评价的影响。

拟采取均值相等性的稳健性检验和 Tamhane 方法进行方差不相等情况下的多重比较研究验证上述假设。

二、研究过程

1. 抽样与数据处理

抽取5个入学年度,4个院系,12个专业,约

37 724 个原始数据,进行描述统计后发现:“课程代码”空缺1 716 个、“入学时间”空缺25 个、“学号”空缺25 个、“教师职称”空缺1 137 个案等共3 702 个异常个案。删除异常个案,最终获得34 022 个有效个案。此外,学生成绩中尚有2 727 个缓考、免修、缺考、成绩以等级评价的情况。有效样本基本情况略。

2. 高校学生网络评教系统指标体系的信度和效度检验

任何量表在使用前必须进行信度和效度检验。前者检验量表的可靠性或者结果一致性,后者检验量表能够测量出所要测量特性的程度。通过分半法计算内部一致性系数,利用克伦巴赫(Cronbach)公式计算 α 系数来估计测验内部一致性。在量表的信度检验中,Cronbach $\alpha = 0.90$,标准化 Cronbach $\alpha = 0.90$,意味着对于学生教学评价,测度量表只有10%的内容未涉及,信度很高。

KMO 检验和 Bartlett 球形检验。对测度量表进行 KMO 检验和 Bartlett 球形检验,检验的 KMO 值为0.954, $p = 0.000 < 0.05$,达到显著水平,测度量表具有结构效度,并且说明样本数据适合进行因子分析。在主成分分析和因子分析中,共同因子累计解释总变异的52.7%;因子载荷约为0.7,聚敛效度和区分效度基本满足相关要求,测度量表基本通过效度检验,适合做后续的分析。

3. 教师教学综合评价结果相关影响因素挖掘

(1) 影响因素内部分组及其描述统计

教师总评分的影响因素复杂,拟从学生年级、学生成绩、授课班级数、授课门数、课程性质等方面挖掘影响教师教学评价的因素。

影响因素内部分组及均值变化规律描述统计,如图1。

1)学生年级。入学年度包括2006~2010共5年数据,转换为大学五至一年级,把教师总评分分5组。统计表明,随着进入大学时间的增长,学生对老师的评价成开口向上的抛物线:一年级学生(2010年入学)较高,大学四、五年级的学生给出评价分均值最高,标准差较小,其他年级则比较低且标准差较大。

2)学生成绩。在0至100之间,按差(0~40)、不及格(40~60)、一般(60~70)、中等(70~80)、良(80~90)、优秀(90~100),把教师总评分

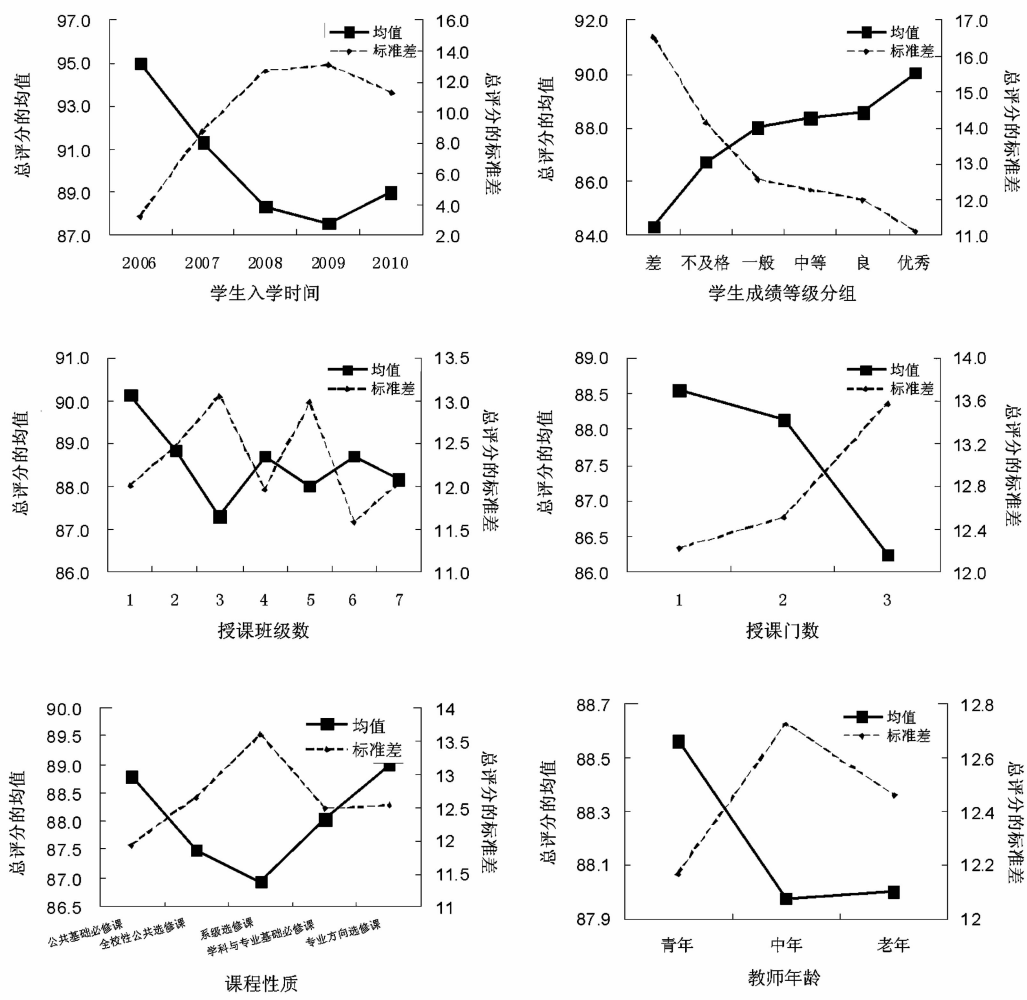


图 1 教师总评分的均值随各影响因素分组变化图

Fig. 1 Mean square value of teachers' overall on-line evaluation scores vs influential factors group

分 6 组。统计表明,学生成绩从低到高,给出评价平均分先快速上升,后平稳,再上升,标准差逐渐减少,反映成绩差和不及格的学生评教分数低于其他类别,而成绩优秀的高于其他类别,一般、中等和良三个等级均值无显著差异。

若考虑缺考、缓考、免修等成绩信息,也有类似结果。缺考与差、不及格的评价均值无显著差异,免修和缓考与一般、中等、良的评价总分均值无显著差异。

3) 授课教学班级数。按照学生数每 30 个人为一个教学班,30 人以下为第 1 组,180 人以上为第 7 组,把教师总评分分 7 组。统计表明,1 个教学班的评价均值显著高于多个教学班,标准差相差无几,多个教学班之间差别不明显。

4) 授课门数。按照授课 1、2、3 门把教师总评分分 3 组。统计表明,授课 1 至 2 门的评价分均值高于 3 门课程,随着授课门数的增加,教学评

价加速降低,标准差随授课门数的增加而下降。

5) 课程性质。按照“学科与专业基础必修、公共基础必修、全校性公共选修、系级选修、专业方向选修”,把教师总评分分 5 组。统计表明,系级选修课教学评价均值最低,标准差最大,专业方向选修课和公共基础必修课教学评价均值最高,全校性公共选修课与学科与专业基础必修课评价居中,而标准差与其他类别相差无几。

6) 教师年龄一般在 25 至 60 之间,按青年(35 岁以下)、中年(35 至 45 岁)、老年(45 岁以上)分类,把教师总评分分 3 组。统计表明,青年教师的教学评价均值最高、标准差最小,中老年教师的教学评价均值较低且基本相当。

(2) 影响因素内部分组教师总评分均值相等性检验

采用不要求方差齐性的均值相等性的稳健性检验,检验的 p 均小于 <0.05 ,拒绝各个影响因素

内部各组间均值相等的零假设,认为教师评价均值在授课学生年级、学生成绩、授课教学班数量、授课门数、课程性质和教师年龄上存在显著差异,教师评价显著受到上述非教学因素的影响。

(3)影响因素内部分组教师总评分均值多重比较

均值相等性的稳健性检验表明,影响因素内部分组的教师总评分均值不相等,拟通过多重比较探索哪些组间均值不相等,从而发现变化规律。采用 Tamhane 方法进行方差不相等情况下的多重比较,多重比较表格略。各影响因素给出的显著性检验结果如下:

1)学生年级。除了三年级与一年级的检验 $p > 0.05$,其余年级均值比较的 $p < 0.05$,认为评价在学生年级上存在显著差异。

2)学生成绩。除了一般、中等和良三个等级的 $p > 0.05$,其余等级的 $p < 0.05$,认为评价在学生成绩上存在显著差异。

3)授课教学班数量。教学班级数为 2、3、4、5、6、7 组的 $p > 0.05$,1 个教学班级的 $p < 0.05$,认为评价在单个与多个教学班间存在显著差异。

4)授课门数。授课 1、2、3 门的 $p < 0.05$,认为评价总分在授课门数上存在显著差异。如果显著性水平定为 0.01,则授课 1、2 门的 $p > 0.01$,授课 3 门的 $p < 0.01$,不同于授课 1、2 门,认为评价在授课门数上存在显著差异。

5)课程性质。公共基础必修课与专业方向选修课,全校性公共选修课与系级选修课、学科与专业基础必修课,系级选修课与全校性公共选修课的 $p > 0.05$,而其余 $p < 0.05$,认为评价在这些课程类型上存在显著差异。

6)教师年龄。中年和老年教师的 $p > 0.05$,

青年与中老年教师的 $p < 0.05$,认为评价在教师年龄上存在显著差异。

三、结语

通过抽取 4 个院系、12 个专业,约 37 724 个原始数据,异常值处理后获得 34 022 个有效个案。在对高校学生评教指标体系测度量表进行信度和效度检验后,采取均值相等性的稳健性检验和 Tamhane 方法进行方差不相等情况下的多重比较探索了学生网络评教的影响因素。研究结果表明学生年级、成绩、教学班级数、课程性质、授课门数、教师所属年龄段是影响教学评价的重要非教学因素,验证了前文所提假设。对于高校学生评教工作建议如下:

(1)对评价分值进行异常值处理。建议排除极端分值,根据余下约 90% 的学生评价分值进行评价并给出评价极差、均值与方差,以消除学生个体因素如评价态度、心理等因素对评价结果的影响;加强对学生评教行为本身的评价与分析,对评教行为“异常”的学生有针对性地进行沟通与辅导。

(2)学生评教只能作为教师教学评价的一个方面,决不可滥用。一是因为学生评教系统存在上述诸多非教师能力和努力程度所能左右的客观因素影响评价结果;二是因为学生并不实际具备评价教师教学的能力,大多数评价只是出于个人体会、兴趣偏好。

(3)区别对待教师所属年龄段、授课门数、课程性质等非教师个人能力、努力程度所能影响的外部因素。适当限制教师授课门数并实现同年级、同年龄段、同课程性质评价。

参考文献:

- [1] 廖明,姜峰,郭燕锋.大学生对教学质量的主观偏好与学生评教体系的完善[J].中国高等教育,2012(18):49-50.
- [2] 常亚平,陈亮,阎俊.高校“学生评教”误差形成机制研究——基于学生态度的视角[J].高教探索,2010(1):80-86.
- [3] 闫志明,张光旭,张立新.影响高校学生网上评教的因素研究[J].中国电化教育,2007(3):45-47.
- [4] 李楠,王欣.高校人力资源管理中学生评教的影响因素分析[J].人口与经济,2012(3):27-32.
- [5] 韩明,陈启山,王鹏辉.教师与课程特征对高校学生评教分数的影响[J].华南师范大学学报:社会科学版,2010(4):44-47.
- [6] 周艳华,解鸿泉.影响高校学生评教的因素及应对策略[J].教育理论与实践,2009(33):6-8.
- [7] 王永林.学生评教的特性及其影响因素初探[J].教育科学,2005(1):28-30.