

# 基于绩效技术的高校教师 教学水平评价系统方法研究<sup>\*</sup>

刘世清

(宁波大学 师范学院, 浙江 宁波 315211)

**摘要:**本文在全面分析当前教师教学水平评价中存在的问题及其产生原因的基础上,借鉴绩效技术和系统思想方法,提出一种基于绩效技术的评价方法-教师教学水平 TTF 系统评价方法。该方法要求评价工作的组织、策划者全面理解评价方法的内涵,正确地把握整体方案,制定科学的评价指标体系,严格控制评价过程,科学地处理原始数据,合理地应用评价结果,有目的地引导教师提高教学水平。

**关键词:**绩效技术;教学水平评价;TTF 系统评价方法

**中图分类号:**G40-058.1 **文献标识码:**A

笔者从分析当前高校教师教学水平评价中存在的问题入手,借鉴绩效技术的理论和方法,提出一种基于绩效技术的系统方法,从宏观的评价指导思想、评价系统的整体把握到微观的评价过程控制、结果的处理和解释等诸方面对该评价方法进行系统地论述,以期能解决目前在高校教师教学水平评价中存在的问题,推动教师教学水平评价向更科学,操作性强,信度更高的方向发展。

## 一、高校教师教学水平评价中存在的主要问题及原因分析

(一)评价主体一元化。教师教学水平评价主要由主持评价工作的决策机构、评价主体、评价客体和评价工具等 4 大要素构成。决策机构主要负责对评价工作的宏观策划、指导与监督;评价主体是对教学水平进行评判的人员,一般由学生、同行和督导员(教学专家)组成;评价客体也称评价对象,是被评价的教师和他所讲授的课程;评价工具即评价标准,它是评判教师教学水平的尺度。从理论上讲,评价主体应是多元化的,且他们的权重要经过科学的方法和程序来确定,并要在评价实践中不断进行调整和修改。但不少学校在评价中,评价主体只有学生,同行和教学专家几乎不参与。因为被评价的教师和课程太多,他们根本不可能全部参与,使得学生成为垄断的评判者,导致出现许多问题。这主要是学校主持评价的决策机构对教学水平评价的理论内涵和控制方法理解不到位所造成的。

(二)评价客体过于庞大。一所上规模的大学,授

课教师一般在 1000—2000 人,每学期就有 3000—4000 门课程要接受评价。这个工作一般要在 1—2 周内完成,工作量非常大,如果不采用科学的方法进行抽样评价,以减少工作量就会影响到评价的信度和效度。对一个被评价的教师来讲,如果他同时讲授几门课程或几个班级,就没有必要对每门课程进行测评,通过抽样或一定的形式测评 1 门,最多 2 门课程即可。从维护教师的权益方面来讲,他也有权要求一些有恶意的班级不能参加评价,从而给教师正常、严格管理学生留有一定的空间。从学生角度来看,同时评价 7—8 门课程,无论从负担上还是心理压力方面,都会影响到结果的信度和效度。

(三)评价标准的科学性受到质疑。评价标准是评价活动中衡量教学水平高低的工具,是评价工作的核心。但实际评价中使用的评价标准至少在两方面受到质疑:一是不同的评价主体对被评价对象所熟悉的侧面和把握的程度是不同的,比如学生在评价时,因为对教师上课的态度、语言表达、教态等方面有较长时间的接触,能够做到较准确地评判,而对教师掌握学科内容是否渊博并不一定能够准确把握。所以,要求评价标准应与评价主体对评价对象能够把握的方面和程度相吻合,有三套(学生用、专家用和同行用标准)相互衔接、相互配合、侧重点不同的子评价标准,从而构成一个评价标准的整体。但许多学校只用学生作为评价主体,可见,此标准非彼“标准”,值得密切关注。二是评价标准的制定需要有熟练掌握现代教育理论的专家参与,以准确地确定一级指标和二级指标的构成和划分,保障各级指标构

成的科学性;同时还要有教育统计和教育评价方面专家的参与,以保障指标权重分配的科学性和合理性。一些学校在这方面做的不够,评价标准的科学性值得商讨。

(四)评价过程的控制有许多漏洞。评价过程是否能进行有效地控制,将直接影响到所采集数据的准确性。评价时一般应由教学管理人员组织学生,分发、监督填写、回收评价表。但由于总体评价对象的数量太多,这些工作就由任课教师在上课前代劳。有的班级在填写时,有 1 人或 2 人替全班同学填表的;出现 99 分现象等等。这些现象的发生,大都是因为评价过程失控造成的。实际评价时应使被评价教师、学生和评价表格相分离。

(五)学生的顾虑和教师的妥协。学生在填写评价表时有一定的心理压力,主要是害怕对教师的评分太低会影响到教师编制考题的难易和对成绩的评定,事实上这种担心是确实存在的;同时,教师也有压力——除了要努力提高教学效果外,由于害怕学生对自己评分太低,使得教师在教学中不敢管理学生,教学过程难以有效控制,难以实现教学目标,提高教学效果;有的教师甚至要刻意取悦于学生,这是教师的悲哀,也是评价工作失败的表现。这样的评价结果教师不认可,学生也不认可,最终导致教师教学水平评价起不到应有的作用,反而起到反作用。

(六)评价结果不做可比性处理,进行简单应用。评价结果要可信可用,就要有可比性。比如教师甲讲授某门文科性课程,由一班学生评价的结果是 93 分,教师乙讲授某门理科性课程,由二班学生评价的结果是 88 分。这二个分数能说明教师甲比教师乙一定好吗?不能,因为它们不可比,讲授课程性质不同,评价主体不同,这样的评价结果怎么能可比呢?如果要进行比较,就要对这些数据进行可比性的处理。但有不少学校恰恰就是用这些原始分数来进行教师间的比较,并做出相应的结论,作为评选教学优秀的必备条件,这很难让教师信服和接受。

通过上面的分析我们可以看出,评价工作出现许多问题的原因,不是学生评价教师这种方法有错误,而是评价组织人员没有真正理解、领会和把握这种评价方法的内涵,使得评价工作在组织、实施、过程控制及结果处理和解释、应用上出现一系列问题,最终导致评价工作的失败。

## 二、基于绩效技术的高校教师教学水平系统评价方法——TTF 系统评价方法

我们在大量调查研究,特别是对两所高校进行个案研究的基础上,提出一种基于绩效技术的评价

方法——教师教学水平 TTF 系统评价方法。这种方法是以绩效评价思想与方法为基础,以系统理论为指导,借鉴当前评价方法的成功做法和经验教训后提出的。第一个 T(three)表示 3 个 1;即教师教学水平评价是一个系统整体;评价中一个重要的环节是评价过程的控制;评价要有一个科学严谨的评价标准。第二个 T(three)表示有 3 个评价主体;即学生、督导员(教学专家)和教师同行。F(four)表示评价结果要做 4 方面的处理;即要把原始分数转化为 Z 分数;要对评价的结果做出比较准确的解释;评价结果的应用要恰当,要在合适的范围内使用;评价的结果要及时、全面、准确地反馈给被评价教师本人。下面我们对于 TTF 评价系统作详细地论述。

### (一)TTF 评价系统宏观方案的组成与整体把握

对教师教学水平进行评价是为了实现 3 个提高:提高学校的整体教学质量;提高教学管理的科学化水平;提高教师的教学、业务水平,并提供量化的考核数据。真正使学生和教师成为教学过程的两个主体,在教学过程中实现良性互动式提高。

TTF 评价系统的构成:主要由评价指导与组织机构;学生;督导员和学科教学同行 4 部分组成。这 4 个部分缺一不可,特别是督导员和学科教学同行是不可缺少的,否则就只有学生成为评价主体了,会引发许多问题。指导机构主要由主管教学的校长总负责,由教务处相关科室具体操作,制定评价的指导思想、评价标准,进行时间安排,并宏观监督控制评价过程。除此之外,还有各院系主管教学的领导、教学管理人员和学生辅导员参与,他们具体负责本院系的评价工作和学生的组织协调工作。

评价对象(客体)的选择主要有两种方式:一是全体教师都是被评价的对象;二是部分教师、部分课程成为评价对象。从表面上看,让全校教师成为评价对象,似乎更能快速、有效地提高学校的整体教学质量,但实际上并非如此。由于被评价的教师数量过于庞大,督导员和学科教学同行人数和精力有限,实际操作时这种做法的可行性比较小。为了节省人力和物力,并且不降低评价的信度,常常采用抽样的方法来减少评价对象的数量,或进行典型抽样评价,比如对新任教的教师、要评职称的教师、申请参评教学优秀的教师、学生反映有较多问题的教师进行重点评价。

评价标准的制定:要使制定的评价标准具有较强的科学性,要注意以下 3 方面的问题;评价标准的制定人员要由各种相关专家组成,一般应有熟练掌握现代教育理论的专家、教育统计和教育评价方面专家参与;3 个评价主体所使用的评价标准其侧重

点应有所不同,以体现从不同的角度,由较权威的评判者全方位地评价教师的教学水平;对评价标准中的各级指标要用科学的方法赋予准确的权重;评价标准应根据评价的实践不断做出合理的修改与完善。

### (二) TTF 评价系统的评价过程控制与具体操作

TTF 评价系统的评价过程有两种实现方式:第一种是传统的形式,即把评价标准和相关信息印刷在纸介质上,通过组织学生填写评价表来实现对教师教学水平的评价。这种方式费时费力,评价过程较难控制,容易出现漏洞。第二种是数字化网络形式,把 TTF 评价系统开发成数字化软件系统,嵌接在校园网上,利用校园网的众多接点,实现快速、大量进行评价的功能。这种利用网络进行评价的过程相对容易控制,数据的输入与处理也能做到更快速、准确、有效。

### (三) TTF 评价系统的结果处理与解释

#### 1. 评价数据资料的处理程序与方法

对数据资料进行处理一般按以下程序进行:资料有效性鉴别、资料类型划分、评价主体信度检验、原始分数转换、绘制数据图表等。我们对收集到的 3 种数据资料(学生、专家、同行),首先要进行数据资料有效性鉴别,即对评价表填写的形式等方面是否符合评价的总体要求而做出是否有效的判断,把不符合要求的资料去掉。这样做既能节省人力物力,又能去掉无效资料以提高评价结果的可信度。

根据教师所从事学科的特点不同,我们有必要根据学科的性质对资料进行分类,这样能够增强评价结果的准确性,比如可以按学科性质分为文科组、理科组、工科组和医科组等;也可按课程的性质不同分为公共课、公选课、专业基础课和专业选修课等。当然,现在许多学校对教师、课程是一视同仁,既不按学科分类也不按课程性质分类,这样处理的结果是它的准确性要降低一些,但其可比较的范围大了,有利于在全校范围内进行横向比较。

在数据资料处理的前期,尽管我们对数据资料进行了有效性鉴别,但在有效资料中,有些资料有明显的倾向性,这主要是由于评价者的态度、能力、心理等方面存在的问题引起的,所以我们要对评价者个体的信度进行分析检验,以确定该评价者与评价主体(群体)间的差异程度,从而确定该评价者的评价资料是否可利用,以利于把态度不端正,有极端观点,与评价主体(群体)差异程度大的评价者的评价资料去除掉,提高评价结果的准确性和可信度。评价者个体的信度检验,一般采用肯德尔和谐系数公式来计算出信度系数:

$$\omega = \left[ \sum R_i^2 - \frac{(\sum R_i)^2}{m} \right] / \frac{1}{12} n^2 (m^2 - m)$$

式中:  $n$  = 被评价客体人数;  $m$  = 对被评价客体考评项目的个数;  $R_i$  = 第  $i$  个考评项目上所有评价客体的得分之和。 $\omega$  值越大说明该评价主体对所有评价客体的评价结果越可靠。

原始分数是评价主体给评价客体所评定的分数,是直接反映被评价对象考评结果的一种数据指标,不能直接用来比较,因为它存在以下一些缺陷:缺乏独立意义——当教师甲的原始得分是 90 分,我们无法判断这是一个高分还是低分,只有我们知道了与这个分数相关的整体和它产生的背景时,才能明确"90"分的含义;原始分数没有可比性——当教师乙的原始分数是 86 分,教师丙的原始分数是 78 分,我们很难说教师乙的教学水平就比教师丙高。因为这两个分数的评价主体是不同的,评价客体(教师本人和课程性质)是不同的,没有可比的基础。所以我们必须对原始分数进行处理和转换,通常的做法是把原始分数处理成  $Z$  分数或转换成  $T$  分数,转换公式是:

$$Z = \frac{x - \bar{x}}{s} \quad \text{或} \quad T = 10Z + 50$$

式中  $x$  代表原始分数,  $\bar{x}$  代表全部原始分数的平均分,  $s$  代表它们的标准差。转换以后的  $Z$  分数,就可以实现加减运算,实现不同课程、不同教师之间的比较,其可比性得到大大增强。

许多孤立、零乱、分散的  $Z$  分数,只能说明教师个体的教学水平,对全体教师、历年的资料、院系间的水平则无法表现出来,如果我们用平均  $Z$  分数作为纵坐标,以年份或院系单位为横坐标绘制图表,就可直观、清楚地展现出历年来教师教学水平的发展趋势,或者学院(系)与学院间  $Z$  分数的比较情况。

#### 2. 评价结果的解释

通过评价获得的  $Z$  分数或  $T$  分数,具有可比性,能够进行加减运算,这给我们解释和处理评价结果带来了极大的方便,也给人们理解评价结果提供了准确的依据。对评价结果的解释可以从以下两方面进行:一是  $Z$  分数(或  $T$  分数)只表示被评价个体的相对水平,不是绝对水平,比如我们只能说同一学校  $Z$  分数是 5 分的教师比 3 分的教师的教学水平高,而不能判断甲校  $Z$  分数是 3 分的教师就比乙校  $Z$  分数是 5 分的低。二是在同一学校中,如果我们确定了一个  $Z$ (或  $T$ )分数的常模值,那我们就可以给该校中的每一个教师的教学水平做出一个定性的解

释和说明。比如我们确定 T 分数 70 为及格,95 以上为优秀,那么就可以此为基准,对每一位教师的教学水平是否合格、是否优秀做出定性评价,并能比较准确地说明某一教师的教学水平在学校整体教师队伍中所处的地位。

### 3. 评价结果的应用

许多学校总是把评价结果作为教师评优的主要依据,这是可以的,但在应用时要保证满足以下几方面的条件:首先,评价结果只能应用于和教学相关的评比、评选中,而不能用于与教学无关的评比,例如可以用于评选教学优秀,但不能用于评选优秀教师和优秀教育工作者,因为评价标准中的评价指标只与教学水平有关,而与教师思想、道德、科研等基本不相关。其次,评价结果的 Z 分数必须是由学生、专家和教学同行 3 部分评价的综合结果,不能是某一方面的 Z 分数。第三,评价过程的控制是严格的,评价资料经过信度检验,且信度比较高。只有满足这些条件的 Z 分数,才可以作为与教学水平相关的评优活动的依据。

### 4. 评价结果的反馈

教师教学水平评价的宗旨之一是通过评价活动促进教师教学水平的提高,其中一个关键环节就是要把评价的结果准确、有效地反馈给教师。所谓准确就是把全校评价的整体情况及相关资料介绍给教师,特别是全校教师的平均 Z 分数、常模 Z 分数,该教师的 Z 分数,以及他在全校教师中的位置等。这

些必须让教师真正清楚和理解,它有助于教师理解和接受评价结果,教师会真正从内心认为这次评价是比较客观公正的,进而转变教学态度,积极进行教学改革和教学研究。所谓有效是指要清楚地、客观地、具体地指明教师在哪些方面具有优势,值得以后保持和发扬;哪些方面还存在不足,可能是由什么原因引起的,并给教师提供一些改进的意见和建议。不能就事论事,停留在评价阶段,没有反馈就没有提高。

TTF 评价方法是建立在信息技术、绩效评价和系统理论基础之上的,具有系统性、科学性、信度高、操作性强、利于教师教学水平的提高和教学管理的科学化等特点。评价中应注意全面理解和把握教师教学水平评价的系统性;避免教师与评价表格的直接接触;对学生过多评价教师和课程的心理压力和心理反应要予以关注;要合理解释与处理评价结果;严格控制评价过程。

### 参考文献:

- [1] 萧鸣政.现代人事考评技术及其应用[M].北京:中国人民大学出版社,1997.
- [2] 理查德·威廉姆斯.组织绩效管理[M].北京:清华大学出版社,2002.
- [3] 许传海,宋涛,张素.对教师教学质量评价的思考[J].中国高教研究,2003,(2):56-57.
- [4] 胡甲刚,汪成平.英国的发展性教师评价[J].中小学管理,2002,(10):21-22.
- [5] 陆兴发,包乌云.美国教师评价制度的改革与发展趋势[J].教育探索,2002,(10):93.

\* 本项目受浙江省 2003 年度哲学社会科学规划课题(NX03GL15)的资助。

## 简讯

### “多元智能理论及其在教学中的应用”国际论坛暨 DIC国际合作项目第三届年会在京召开

2004年5月19日至22日,“多元智能理论及其在教学中的应用”国际论坛暨DIC国际合作项目第三届年会在北京市京丰宾馆召开。本次会议由联合国教科文组织协会全国委员会、北京市教委、中国教育学会、“借鉴多元智能理论开发学生潜能实践研究”课题组联合主办,北京市丰台区人民政府、北京市教育学院等单位承办。

多元智能的创始人霍华德·加德纳博士在会上作了《对多元智能理论用20年来的反思》的报告,并回答了与会者的问题。来自美国、英国、日本、加拿大、澳大利亚、新加坡等国家的专家以及中国香港、澳门和台湾的专家分别在大会上作了报告,介绍多元智能在不同国家和地区的实践情况和各位专家所

研究、关注的热点问题。大会还安排了交流和参观活动。

中央电化教育馆“加强中国西部基础教育能力项目”的有关专家应邀出席了大会,并组织了《运用多元智能开发学生能力的互动式教学》研讨活动。

22日上午,在北京市第九十二中学进行了DISCOVER(Discovering Intellectual Strengths and Capabilities while Observing Varied Ethnic Responses)项目实验学校揭牌仪式,美国亚利桑那大学的琼·梅克教授作为项目负责人出席了揭牌仪式,并对现场听取的2节课进行了点评。据悉,北京市九十二中学将成为中国内地的第一所DIC项目学校。

(本刊记者 朱广艳)