

对外汉语教学同行评估的实证调查与分析

刘德联 李海燕

提要 本文指出教学评估在对外汉语教育事业中应起到的作用以及通过各种形式的教育评估建立科学系统的教育评估体系的意义,分析了对外汉语教学同行评估的现状,阐述了实行同行评估的必要性,通过对对外汉语教学同行评估所进行的实证调查与分析,分析了同行评估与学生评估的差别,指出二者是不能相互替代的。

关键词 对外汉语教学;教学评估;同行评估;信度

Investigation and Analysis about Peer Evaluation in TCSL

Liu Delian & Li Haiyan

Abstract: This paper points out the importance of the teaching evaluation in TCSL, and the construction of scientific evaluation system in various ways is meaningful. It analyzes the present state of peer evaluation in TCSL, discusses the necessities of peer evaluation, shows the difference between peer evaluation and student evaluation and that they cannot be replaced by each other.

Key words: TCSL; teaching evaluation; peer evaluation; validity

零、绪论

对外汉语教育是近年来中国的高等院校里逐步发展起来的一门新学科。早期的高校对外汉语教育,只是为那些希望入系学习但语言不过关、听不懂院系课程的留学生设立的带有补习性质的课程班,现在已经发展成为多方位的,包括预科教育、本科教育、研究生教育以及非学历教育的学科,形成特有的一整套对外汉语教学体制。

对外汉语教育事业的全面开展,呼唤对外汉语教师素质的提高和教学活动的规范化,因而有必要建立一整套科学系统的教育评估体系。

科学的教育评估体系的建立,对评估对象来说,通过有效的信息反馈及交流,可以起到诊断作用(指出优缺点以及如何改进)、导向作用(引导评估对象的行进目标)、鉴定作用(衡量评估对象是否称职)和激励作用等等。

科学的、完整的教育评估体制应当包括领导评估、学生评估、同行评估以及自我评估等各种形式。对评估对象所做的结论,应当在各类评估的综合评价后得出。其中,同行评估的权重应当占有相当的比例。

“研究结果仍然支持这样的观点——由同行来评价教师的学术水平比学生评更恰当。(J. A. 森特拉 1992, 8) 在各类评估中,最能给予教师教学水平以科学评价的当数同行评估。”

一、同行评估的现状

目前,对外汉语教育评估还未能在各高校中普遍开展。在已经实行教育评估制度的高校中,也极少有同行评估,大多是学生评估。不可否认,学生评估对教育者的督促作用是很明显的,从总体上也应当说是公正的、客观的,但是毕竟有一定的主

观因素在里面,仅仅靠学生评估的结果来评判一名教师的教学水平,往往是不全面的,在某些方面甚至是不公平的。有时候教师的严格管理(如较严格的记分、较多的作业等)会让一些学生感觉不舒服;有时候教师一些行之有效的教学方法在短时间内也得不到学生的认可,往往出现老师一心要搞好教学、

山梨県立大学 国際政策学部 国際コミュニケーション学科

Department of International Studies and Communications, Faculty of Global Policy Management and Communications, Yamanashi Prefectural University

提高学生的汉语水平而学生不领情的现象。而个别教师会为了讨好学生采取过于宽松的教学管理（如不认真记考勤、考试都给高分等等）。

在一些高等学校，虽然没有明确提出同行评估这一概念，但是已经在这方面进行了一些有益的尝试。主要形式有以下几种：

1. 建立教学督导制度。

一些高校选派有丰富教学经验的老教师担任教学督导。教学督导随堂听课，课后对教师在课堂的教学行为进行点评和必要的指导，并在适当的时候开办讲座，根据一些教师教学中出现的问题，向全体教师做教学指导。

2. 建立平行班负责人制度。

一些高校留学生较多，出现几个教学班共用一本教材的现象，于是建立平行班负责人制度。平行班负责人要通过听课、共同备课、轮流出考题等方式指导年轻教师和外聘教师。

3. 录制教师完整课堂教学活动。

在一些班级进行课堂教学实况录像，然后组织部分教师观摩，对教学中的问题并进行研讨。

4. 组织研究生进行课堂教学观摩。

组织对外汉语教学相关专业研究生进行课堂教学观摩，可以说是一种变相的同行评估。对外汉语教学相关专业的研究生，在学习了一定的课堂教学理论之后，会用一种评判的眼光来观察教师的课堂教学活动，会在理论指导下，对教师的课堂实践进行思考，提出问题。

二、有关同行评估标准的制定

同行评估不同于学生评估，其特点在于：

1. 同行评估更注重教师的素质，如教师的内在修养，包括知识结构、教学态度；教师的外在形象，包括着装、仪表、精神状态；教师的语言表达能力，包括规范的语音、适当的语速等等。
2. 同行评估更注重从业务角度对教师的课堂活动进行评价。过去有一句话说：“内行看门道，外行看热闹。”一个有经验的评估者，能够结合教学理论对被评估者的教学活动归纳、总结，给予准确的评价。
3. 同行评估更注重教学方法的评估，包括课堂教学的步骤，时间的掌握，教学技能的运用，课

堂活动的组织等等。

4. 同行评估更具有客观性。由于学生的利益与教师的管理有一定的相关性，学生在评估时常常有所顾虑，主观性相对来说较强。同行评估由于评估者与被评估者没有根本的厉害冲突，评估结果会更加客观。

要搞好同行评估，首先需要制定科学的、完整的评估标准，使同行评估有章可循。制定这样一个标准，并使它具有较高的信度和效度，不是一件简单的事情。“教育学指出，教学过程的本质是一个特殊的认识过程。这种特殊性反映在三方面：（1）教师的指导性；（2）认识的间接性；（3）教学的教育性。对于高等教育而言，还有另外三个方面，即学生学习的独立性、专业性和开拓性，而这三个方面又受到教师学术水平和教学水平的深刻影响。所以，评估高等学校的教学质量的关键是‘教师的学术水平’和教师根据教学规律和自己的教学水平所安排的‘教学过程的质量’，这是两项影响教学质量的关键性指标，由此派生的指标体系便是科学的指标体系。”（许建钺 1992，4）

要制定同行评估的标准，还需要从定量和定性两个方面来考虑。如果能够将评估指标量化，会使评估更具有客观性，也便于进行数学统计，对被评估者定级。但是，“由于教育活动的极端复杂性和人们对存在于教育活动中各项因素相互关系的认识相对来说还很肤浅，在评估中用纯客观的方法有时也会把复杂现象作简单化处理，导致评估者有意或无意地忽视一些在目前情况下较难定量的和缺乏客观资料的因素。而这些因素对教育活动来说也许具有极为重要的影响，离开了这些因素的评估，教育评估是不全面的。”（许建钺 1992，4）因此，在评估中也要加强定性分析，使被评估者了解自己在教学中被肯定的方面和需要改进的方面，对他今后的教学工作产生促进作用。

为制定这样一个标准，本文作者在参考本校学生评估项目、兄弟院校的学生评估以及同行评估项目、国外教育评估项目的基础上，试拟出一份同行评估项目并在校对外汉语教师中进行了问卷调查，根据反馈意见对同行评估项目进行修订，初步制定出同行评估表，然后在一些教师的课堂教学中进行试评，由教学督导、调研者在听课时填写同行

评估表,对不同教师、不同课程的同行评估结果进行统计分析,根据试评中出现的问题对同行评估项目进行再修订,最后制定出同行评估的项目标准。我们希望能将这一成果运用于课堂教学中,在使用中不断改进、完善,通过同行评估这一手段,使之更有效地服务于教学实践,提高对外汉语教学质量。

三、有关同行评估项目的调查

确定科学合理的评估项目是决定评估有效性的关键。“指标是一种具体的、可测的评估项目,因此,抓住被评客体的本质属性并把它分解为可测项目,就是指标体系设计的根本方法。”(许建钺 1992, 4)

为使我们的同行评估项目更趋于合理,我们设计了一个有关同行评估项目的调查问卷。调查问卷中所列的项目是在以下评估项目的基础上综合而成的:

1. 美国一些大学的教学评估项目;
2. 美国一些专家对大学师生的调查结果;
3. 美国教育测试服务中心(ETS)的“学生教学评估报告表”(SIR);
4. 国内一些大学的学生评估项目;
5. 国内一些大学的同行评估表;
6. 国内一些大学的课堂教学质量评估表。

我们最初列出的项目较多,分类也比较细,但是我们不可能将所有的项目都列入调查表中。在选取整理的过程中,我们注意遵循以下原则:

1. 评估项目的说明要简明扼要,具有直接可测性,所规定的内容是可以通过实际观察和测定获得明确结论的。
2. 所选评估项目具体可行,每一位教师认真执行都可以做到。
3. 所选评估项目具有引导性,教师可以以此为目标,约束自己的课堂行为。
4. 所选评估项目是课堂教学中不可缺少的。

我们的调查问卷包括下面几方面的内容:

- (一) 课前准备。具体内容包括:做好课前教学准备、熟悉教学内容、课上讲明教学安排、准备适当的补充材料或练习等等。
- (二) 教师素质。具体内容包括:着装、课堂行为、音量、发音、语速、板书、应变能力、学科知识等等。
- (三) 教学方法。具体内容包括:教学步骤、组织

学生参与课堂活动、合理利用课堂时间、举例、难度控制、讲解、辅助教学手段、讨论活动的组织、启发式教学、激发学生的学习兴趣、了解学生的理解程度、将教学内容与学生已有经验知识的结合等等。

(四) 教学效果。具体内容包括:按计划完成教学内容、给学生均等的机会、学生的满意度、突出教学难点、课堂气氛等等。

(五) 加分项目。具体内容包括:多媒体、网络、影视教学资源的利用、课堂教学有创意等等。

问卷调查在北京大学对外汉语教育学院的教师中进行。在我们所收回的29份有效问卷中,老师们对前四类项目中的选项赞同率较高,占调查人数的72.4%到96.5%,认为“无必要”的仅占3.4%到20.7%。对第五项看法有较大分歧。我们在最后的定稿中没有列入这一类。

很多教师在调查问卷中对评估项目的措辞以及评估内容提出自己的意见与建议。我们根据广大教师的意见和建议,精简合并了部分项目,增加了部分内容,并根据各个项目的权重,设计了各类项目的计算分数,使我们的评估项目更加合理,更便于操作。

四、同行评估的实证调查与分析

在有关评估项目的调查研究基础上,我们初步设计了作为同行评估的《对外汉语课堂教学同行评估表》,在北京大学对外汉语教育学院的部分兼职教师的课堂教学中试用,并对此进行了实证调查与分析。我们的研究目的是希望通过实证调查与分析来验证同行评估与学生评估之间有无差异,有多大差异,是否能检验出学生评估中发现不了的教学问题。

我们之所以选择在兼职教师的课堂教学中试用,是因为北京大学对外汉语教育学院多年来一直对兼职教师进行教学跟踪管理,以保证兼职教师的教学质量,这种管理是通过聘请教学经验丰富的退休教师担任教学督导来完成的,但是由于缺少系统的量化指标,主要是靠督导的听课印象做出某种评价。在2006年秋季学期我们请教学督导老师使用我们研制的《对外汉语课堂教学同行评估表》跟踪评价一些兼职教师的课堂教学,然后把评估结果与该学期末的对同一名教师的学生评估结果进行比较,分析评估表的信度,对此进行了实证研究工作。

为了对教师的教学素质和能力有更全面的评价,我们设计的《对外汉语课堂教学同行评估表》在广泛征求教师意见的基础上,设计了四大类别(教学准备、教师素质、教学方法、教学效果)、11个评估项目和26个评估细目,对评估细目的评价有“好、一般、不好”三个级别,对评估项目的评价有“A、B、C、D、E”五个级别,每个级别都赋予了相应的分值,所有的细目评价都为“好”的话,总分为100分,所有的项目评价都为A的话,总分也为100分。详细情况请参看本文后附表1。把评估内容分为项目和细目的目的有两个,一是通过细目帮助评估者更好地把握评估内容,使课堂评估内容更具体,增强评估的针对性和可操作性;二是希望比较项目和细目评价的一致性情况,为同行评估量表的修改提供依据,以使量表更加简便、科学。

2006年秋季学期教学督导老师使用《对外汉语课堂教学同行评估表》对11位兼职教师的课堂教学进行了评估,每位教师的评估时间为两节课。这11位教师的课堂教学在课型上有汉语精读课、读写课和专业汉语课,覆盖了初、中、高级。我们分别计算了督导的细目评分总分和项目评分总分,在期末时又收集了对这11位教师该门课的学生评估分数,满分也为100分。数据见表1:

通过对表1数据的计算,得到督导细目评分与学生评估总分、督导项目评分与学生评估总分的相关系数分别为:0.114和0.052,相关程度很低。督导细目评分和督导项目评分的相关系数为0.954,相关程度很高。

2007年春季学期,本文研究者参考督导老师

的意见对同行评估量表进行微调,四个类别和11个项目没变,把“教师素质”类中“内在修养”项目下由原来的5个细目合并为4个细目,把“教学方法”类中“课堂活动”项目下由原来的3个细目改为4个细目,这样适当增加了课堂活动的比重,项目和细目的总分未变。详细情况参见本文后附表2。督导老师使用修改后的评估量表在2007年春季学期的19名兼职教师的课堂教学进行了评估,之后我们再次收集了期末该班学生的评估分数,数据见表2:

通过对表2的数据的计算,得到督导细目评分与学生评估总分、督导项目评分与学生评估总分的相关系数分别为:0.191和0.177,相关程度也很低。督导细目评分和督导项目评分的相关系数为0.92,相关程度很高。这与2006年秋季学期的分析结果基本一致。

根据上面对表1和表2的数据的分析结果,可以看出,同行评估和学生评估基本上不相关。这也证明了本文在第二节中所谈的同行评估不同于学生评估,两者的性质、出发点和目标都不相同,是不能互相取代的。我们也发现,大部分教师的学生评估结果好于同行评估,我们选取督导评估分与学生评估分相差较大的详细查看了学生评估中学生对该教师的评语,发现其中出现较多的评价是“老师很善良、很关心我们、给我们很多资料、很开心、老师声音很好听”等,在同行评估比较差的班级里也常常出现某些学生给教师满分的情况。可见,很多学生对老师的评价更多的出于感情和直观感受,他们很难判断老师讲解的语法点是否是正确的,老师的教学方法和环节是否高效等等。而很多教师满足

表1:2006年秋季学期对11名教师的督导评估和学生评估情况

	督导细目评分总分	督导项目评分总分	学生评估总分
教师A	70.25	70.00	84.44
教师B	83.00	85.00	86.36
教师C	69.50	74.00	78.80
教师D	80.25	79.00	82.40
教师E	81.75	86.00	74.10
教师F	91.00	91.00	88.80
教师G	63.25	75.00	77.20
教师H	50.25	55.00	86.20
教师I	78.00	81.00	88.80
教师J	89.00	91.00	90.80
教师K	55.50	67.00	89.25

表 2: 2007 年春季学期对 19 名教师的督导评估和学生评估情况

教师序号	督导细目评分总分	督导项目评分总分	学生评估总分
1	70.25	68.00	91.43
2	83.00	85.00	90.18
3	70.50	66.00	71.83
4	74.50	74.00	88.80
5	74.50	77.00	74.44
6	39.75	44.00	88.83
7	69.50	76.00	94.67
8	80.25	84.00	91.00
9	89.00	81.00	99.00
10	91.00	91.00	97.00
11	70.00	80.00	87.00
12	81.75	86.00	92.20
13	55.50	67.00	90.00
14	79.75	84.00	75.00
15	63.25	75.00	80.00
16	50.25	55.00	90.40
17	69.50	74.00	91.60
18	78.75	85.00	94.70
19	78.00	81.00	92.60

于学生的溢美之词或者不明确学生评估分数低的原因所在,就很难在教学能力和教学方法上更高地要求自己,使自己的教学水平得到真正提高。在这种情况下,同行评估作为学生评估的补充就显得尤为重要。

表 1 和表 2 的数据分析显示,督导细目评估和项目评估的相关程度很高,因此为了使同行评估更加操作简便,在评估量表中可以仍保留评估细目作为参考,但评估者只需给评估项目打分,在计算总评分时也只需计算评估项目总分即可。

另外,为了检验同行评估的内部一致性即评估信度,本文的两位作者曾同时听了两位教师的课并进行评估,结论一致性很高,信度基本可靠。

五、相关讨论与结论

通过以上的讨论和实证分析,我们可以对同行教学评估得出以下结论:

1. 同行评估与学生评估是从不同角度实行的,二者是不能相互代替的。同行评估对提高被评教师的课堂教学水平的作用大于学生评估。
2. 同行评估的内部一致性较高,信度可靠。有研究显示,同一班级学生对同一位教师的评估存在较大差异,评估得分较高的教师,学生之间

的评估基本没有差异。评估得分相对较低的班级,相对差异较大。除了学生因学习目的不同而对教师的要求不同、文化背景的不同对师生情感交流的影响、学生不同的学习态度等等,都会对评估的得分产生影响(刘德联、李海燕,2004),而采用同行评估可以有效地避免学生评估的这种不足。

在本文的调查过程中,我们也听到少部分教师对同行听课评估感到有压力,这是可以理解的。因此本文建议,同行评估应以探讨课堂教学规律、提高课堂教学水平为主要目的,可主要用于对新手教师的教学评估之中,作为学生评估的补充。

参考文献

- J. a. 森特拉(美)著,许建钺译,《大学教师工作评估》,北京航空航天大学出版社,1992
- 蔡德章、聂万勇、高献忠,建立发展性课堂教学评估体系,《中国高等教育评估》2006年3期
- 刘德联、李海燕,对外汉语教学评估偏差分析及纠偏对策,《语言教学与研究》2004年第5期
- 许建钺 《高等学校教育鉴定与水平评估》,中国科学技术出版社,1992
- 赵慧、沈玉顺,《课堂评估技术在我国高校教学实践中的效果研究》,2006年3期

附表 1：对外汉语课堂教学同行评估表（06 版）

姓名：_____ 教龄：_____ 课型：_____ 班级：_____ 级 _____ 班 教材：_____

类别	项目	细目	细目评价	项目评价
教学准备	目的明确	能让学生了解教学计划	好□ 一般□ 不好□	A□ B□ C□ D□ E□
	材料充分	熟悉教学内容	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		能适当补充相关练习材料	好□ 一般□ 不好□	
教师素质	外表形象	着装得体、仪表整洁	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		精神饱满、行为规范	好□ 一般□ 不好□	
	内在修养	专业文化知识讲解准确清楚，基本功扎实	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		国内外文化知识丰富，能与学生进行文化交流	好□ 一般□ 不好□	
		有责任心、认真、自信、耐心、宽严适度	好□ 一般□ 不好□	
		善于与学生沟通，有一定的应变能力	好□ 一般□ 不好□	
		所谈话题得当，言辞得体	好□ 一般□ 不好□	
	语言表达	讲解简单明了	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		发音标准清楚、语言规范	好□ 一般□ 不好□	
		语速适合学生水平	好□ 一般□ 不好□	
教学方法	教学步骤	层次分明、重点突出	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		教学环节合理必要，详略得当	好□ 一般□ 不好□	
		时间掌握适度	好□ 一般□ 不好□	
	课堂活动	能以多种方式调动学生学习积极性	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		教学互动性较强，能有效组织交际实践活动	好□ 一般□ 不好□	
		使学生练习机会均等，对学生一视同仁	好□ 一般□ 不好□	
	教学技能	板书准确清楚，布局合理	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		有效使用身体语言、教具及现代化教学设施	好□ 一般□ 不好□	
		教学设计有创意	好□ 一般□ 不好□	
	因材施教	能针对课型特点使用适当方法	好□ 一般□ 不好□	A□ B□ C□ D□ E□
能适应学生水平和要求，循序渐进		好□ 一般□ 不好□		
教学效果	课堂气氛	气氛活跃、秩序良好	好□ 一般□ 不好□	A□ B□ C□ D□ E□
	学生反应	能使大部分学生掌握所学内容	好□ 一般□ 不好□	A□ B□ C□ D□ E□

附表 2：对外汉语课堂教学同行评估表（07 版）

姓名：_____ 教龄：_____ 课型：_____ 班级：_____ 级 _____ 班 教材：_____

类别	项目	细目	细目评价	项目评价
教学准备	目的明确	能让学生了解教学计划	好□ 一般□ 不好□	A□ B□ C□ D□ E□
	材料充分	熟悉教学内容	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		能适当补充相关练习材料	好□ 一般□ 不好□	
教师素质	外表形象	着装得体、仪表整洁	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		精神饱满、行为规范	好□ 一般□ 不好□	
	内在修养	专业文化知识讲解准确清楚	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		所谈话题得当	好□ 一般□ 不好□	
		有一定的应变能力	好□ 一般□ 不好□	
		认真、自信、和蔼、耐心、言辞得体	好□ 一般□ 不好□	
	语言表达	发音标准清楚、语言规范	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		语速适合学生水平	好□ 一般□ 不好□	
		讲解简单明了	好□ 一般□ 不好□	
教学方法	教学步骤	层次分明、重点突出	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		教学环节合理必要，详略得当	好□ 一般□ 不好□	
		时间掌握适度	好□ 一般□ 不好□	
	课堂活动	能以多种方式调动学生学习积极性	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		教学互动性较强	好□ 一般□ 不好□	
		能有效组织交际实践活动	好□ 一般□ 不好□	
		使学生练习机会均等	好□ 一般□ 不好□	
	教学技能	板书准确清楚，布局合理	好□ 一般□ 不好□	A□ B□ C□ D□ E□
		有效使用身体语言、教具及现代化教学设施	好□ 一般□ 不好□	
		教学设计有创意	好□ 一般□ 不好□	
因材施教	能针对课型特点使用适当方法	好□ 一般□ 不好□	A□ B□ C□ D□ E□	
	能适应学生水平和要求，循序渐进	好□ 一般□ 不好□		
教学效果	课堂气氛	气氛活跃、秩序良好	好□ 一般□ 不好□	A□ B□ C□ D□ E□
	学生反应	能使大部分学生掌握所学内容	好□ 一般□ 不好□	A□ B□ C□ D□ E□