

本科教学工作水平评估问题论析

刘天佐

(湖南农业大学, 湖南 长沙 410128)

摘要: 结合国外特别是美国高等教育的质量认知与质量认证模式, 从评估的策划、标准与结论三个方面, 论述了中国普通高校本科教学工作水平评估存在的主要问题。一是组织体系中的模式与进度问题, 表现为评估模式与高校管理体制存在偏差; 评估速度与高校质量建设周期不相协调。二是指标体系中的条件与质量问题, 表现为过分强调资源, 将声誉与条件等同于教育质量; 过分强调教师教学, 将教学水平等同于学生学习的质量。三是评价体系中的效度与信度问题, 表现为强调专家团队实地考察, 降低了评估工作的效率; 评估结果无区分度, 混淆了公众的视听。

关键词: 教学评估; 组织体系; 指标体系; 评价体系; 高校

中图分类号: G642.42

文献标志码: A

文章编号: 1009-2013(2011)02-0064-05

Problems in undergraduate teaching level assessment

LIU Tian-zuo

(Hunan Agricultural University, Changsha 410128, China)

Abstract: Grounded on the quality perception and quality certification model of higher education assessments in foreign countries, especially United States, the paper analyses problems lying in the undergraduate teaching level assessment in China from three aspects: strategical planning, standards and conclusions of the assessments. Viewed from the aspect of strategical planning, the problem is that the assessment mode deviates from management system of colleges and universities. Considered from the assessment standards, the problem is evaluation subjects over-emphasizing teaching resources and equates the reputation and conditions as educational quality. Looked from perspective of the conclusions of the assessments, the problem is that the validity and reliability are unsatisfied.

Key words: teaching assessment; organization system; index system; assessment system; colleges and universities

自1985年中国开展高等教育评估工作以来, 经过不断的改革与实践, 高等教育评估这个系统工程日渐完善, 评估工作取得的成效十分显著, 但同时也存在不少问题。钟秉林教授对上一轮本科教学工作水平评估工作进行总结时指出, (评估在取得成绩的同时) 也存在许多不容忽视的问题, 一是评估指标体系单一, 缺乏明确的分类指导功能; 二是评估方式单一, 过于依赖专家进校考察方式降低了评估的效率; 三是评估结果优秀率过高, 与社会公众的印象反差较大; 四是评估工作的效能有待加强, 评估后的整改工作和成果巩固以及经验推广工作相对薄弱, 评估过程中对社会舆论的正面引导不够。^[1]

笔者拟从组织体系、指标体系与评价体系等三个方面, 对中国高校本科教学工作水平评估存在的问题进行论述。

一、组织体系: 模式与进度的问题

组织体系是评估工作的总策划。组织体系应当解决评估的主体、对象、模式与节奏等评估工作的基本问题, 并进一步解决评估的策划者、组织者、参与者、实施者、评估模式与评估节奏安排等问题。然而, 在本科教学工作水平评估的组织体系中, 质量评估的模式选择与进度安排存在明显问题。

1. 评估模式与高校管理体制存在偏差

当前国际高等教育质量评估可以归纳为4种主要模式: 一种是以美国为代表的质量认证模式, 是一种以自我评估和同行评估为基础的质量保障机制,

收稿日期: 2011-03-12

作者简介: 刘天佐(1964—), 男, 湖南衡阳人, 教授, 主要从事教育财政与高校财务管理研究。

政府一般不直接参与高校评估,而是通过认证机构来实现对教育质量的间接控制,目的是为公众提供公信力高的高校质量证明。二种是以英国为代表的院校审核模式,是一种类似财务审核方式的质量保障机制,通过对高校内部的质量保障机制进行考察和评议,为学科基准(学术标准)制订与内部质量控制提供指导。三种是以法国为代表的政府评估模式,是一种以政府机构评估为主导对高校质量评估的质量保障机制,评估目的是增进大学的责任,督促其提高教育质量。四种是以日本为代表的自我评估模式,是一种以高校自检自评为重点的质量保障机制,该模式由高校内部人员进行自评,目的是发现自身在教学与科研等方面的不足。上述四种模式的机制不同、标准不一、方式各异,但在评估模式的选择上,却充分体现了不同办学主体在评估中的职能分工,充分赋予了高校对质量保障的主要权责,注重依靠与发挥专家优势,并注意吸收社会公众参与。^[2]

受历史文化传统的制约和中央集权管理体制的影响,中国政府部门一直是大学质量保障的主体,他们不仅代表权力,而且代表资格与信用(孙建荣《中美高等教育质量保障的比较研究》,在2008年院校研究高级研讨班专家会上发言材料)。中国主要是政府通过制定政策和规章自上而下调整高等教育行为,高等教育质量面临的问题和挑战,归因于学术认知与传统文化以及政府和高等教育之间的关系。笔者认为,由教育部发起、策划并组织实施的本科教学工作水平评估方式的选择,既不同于国外高等教育的评估模式,也不符合中国高等教育“由中央与地方共管、以地方政府管理为主”的管理体制现状。众所周知,中国高等院校的教育质量评估工作是在教育大众化背景下实施的,当时面临的问题是怎样调动中央与地方两级政府办学的积极性、加强质量控制、改善办学条件,从而达到提高教育质量的目的是,而这同样也是教学工作水平评估要重点解决的矛盾。但是,在教育部决策者的思维中,尽管地方政府作为高等教育一级管理主体,但教育部却没有让地方政府参与到评估的组织工作中去,试图通过高校的自评自纠与专家现场考察,来解决扩招后高校的质量保障等问题。由于中

央教育行政部门在评估过程中的“越位”现象,导致了地方教育行政部门在评估中的“缺位”,从而使“以评促建”指导思想中地方政府没有实施“评”的权力,也导致了地方政府在“以评促建”中“建”的责任没有较好地履行(据笔者统计,由于地方政府财政有限,高校扩招期间生预算内教育经费支出呈明显下降趋势,高校为改善办学条件与准备评估工作,直接负债达2 000亿元)。

2. 评估速度与高校质量建设周期不相协调

西方国家的高等教育大众化是一个自然的过程,一方面,国家经济高速发展需要大量的人才,因此推动了高等教育大众化的进程;另一方面,大众化是在国家的教育制度和财政能力均已做好了充分准备的情况下自然进行的。美国著名教育学家马丁·特罗指出:“美国高等教育规模的扩展不取决于政府的规划或政策,美国早在有众多人注册入学之前就建立了大众高等教育系统的组织和结构框架,所需的只是发展,而取决于社会对中等后教育的需求和关于中等后教育机构的入学标准与学费。”(Martin Trow, *Comparative Perspectives on Higher Education Policy in the UK*, Oxford Review of Education, 1998年第1期)

中国高校的扩招是在特定历史背景下实施的中央政府行为(1998年亚洲金融风暴事件后,以汤敏为代表的一批经济学家认为,扩大高校招生计划可以拉动国内建筑等行业的消费需求,中央政府为了扩大内需,做出了高校扩招的决定),也可以说是在政府财政承受能力、社会对高校扩招的认识以及高校办学条件等方面均不具备的情况下实施的政府行为。由于高校扩招与高校管理体制调整几乎同时进行,一方面,许多合并高校还处在磨合期,合并院校在机构设置、人员安排、学科规划、专业设置等方面还没有形成整体意见,学校工作处于百端待举之中;高校的人才培养模式由原来的精英教育向大众化教育转变,这是中国高等教育从未经历过的事情,教育举办者与教育承办方还没有领悟到大众化背景下人才培养的规律。另一方面,教育扩招速度与政府财政投入力度极不相称,2006年全国高校的招生人数是1998年的5倍,高校的毛入学率由1995年的7.2%增长到2006年的22%,但财政性

教育经费占GDP的比例一直徘徊在3%左右,同期政府的高等教育财政拨款只增加了2倍,财政性教育经费占学校办学经费的比例由扩招前1998年的64%下降至2006年的不足43%。^[3]在这种情况下教育部急于在全国各高校推进本科教学水平工作评估,希望以此来促进政府增加教育投入与促进高校人才培养质量的提高,但由于责任主体定位不准以及评估节奏把握不好等原因,结果出现了事与愿违的现象,给评估高校造成了巨大的压力与长期的“硬伤”。高校评估的节奏与大学建设周期之间存在反差。^[4]首先,在高校进行大规模数量扩张以后,质量建设有一个过程,尤其是合并与新设立的大学更需要一个磨合阶段,所以,评估推进的速度应该与高等教育和大学本身建设的进度相一致。如果评估节奏过快,即使评估方案设计得非常合理,在执行过程中也容易走样。其次,教学评估本身也有一个成长与完善的过程,特别是周期性教学水平评估,面对如此众多且种类不一的高等学校,更需要一个逐渐成熟与修正的过程。因此,评估本身与评估节奏之间也需要有一种比较好的协调与互动。否则,容易出现欲速不达的现象。

二、指标体系:条件与质量的问题

指标体系是评估工作的标准。指标体系的制订应以当时高等教育面临的问题为导向,以问题的合理解决为目标,充分反映教育举办者、承办者与社会公众的意愿。指标体系既是专家小组的工作依据,又是高校建设的行动指南。然而,在教育部本科教学工作水平评估指标体系的设计过程中,由于急于改善办学条件的迫切愿望以及对教育质量理解与把握上的模糊认知,导致指标体系的设置不太合理。

1. 过分强调资源,将声誉与条件等同教育质量

关于资源与质量的关系,奥斯汀在1993年研究表明,这种关联是不存在的,或者甚至是负相关(William E Knight, Ph.D. Quality Assurance Its Support by Institutional Research in American Higher Education)。他在2002年又进一步提出了两种不同的质量观:一种是质量的声誉与资源观;二种是才能发展观。他认为最好的高校能够最大限

度地促进学生的学习与发展,声誉及拥有的资源和学生学习、发展之间并没有明确的关联。在这种理论的指导下,美国高校质量认证强调结果而不是过程,重视学生的学习结果更甚于重视院校的资源。

中国教育部在本科教学水平评估指标的設置中,对于资源与办学条件指标的設置存在两个方面的主要问题。一是过分强调资源与办学条件在指标体系中的份额;二是指标体系設置没有区分度,形成“一把尺子量天下”的局面。在评估指标体系中,涉及资源与办学条件的一级指标包括“师资队伍、教学条件与利用”两个方面,二级指标、主要观测点以及等级标准的制订也过于具体,同时又没有考虑各高校的建设将受到条件现状、财力水平与人才培养周期等限制。在这种评价标准的指导下,全国高校掀起了一股建新校区之风、大兴土木之风与人才流动之风,也导致部分地方政府与高校放宽职称评定标准、扩大职称评定指标,以满足评估对于办学资源与条件的需要。这种影响对于办学资源基础较差的高校尤为突出,它们不仅要在评估期间加快基础设施建设,而且要大量引进人才、添置仪器设备与图书资料。根据笔者的测算,以评估期间的价格水平,将“师资队伍、生均校舍面积、生均教育行政用房、生均仪器设备与生均图书资料”等几个办学条件指标按“合格标准”进行建设,折合成“生均标准”所需经费在5万元以上。由于中国东部、中部与西部存在明显的经济发展差距,地区之间财力分布的不平衡性导致教育投入的巨大差异。在高等教育系统中,既有办学类型与办学层次上的差异,也存在办学地域与办学条件的差异,经济发达的东部沿海地区以及中央财政供给高校,教育投入相对充裕,而广大中西部地区高校的政府投入则严重不足,多数学校财务一直赤字运行,更有部分学校几乎到了“山穷水尽”的地步,连教师工资与正常运行的经费均无法保障。而本应在评估投入中发挥重要作用的地方政府,由于其作用与责任没有在评估中得到体现,使其对高校的投入无法得到保障。面对巨大的经费需求与评估压力,高校选择向银行贷款等手段筹措资金,到评估工作基本结束时,全国高校银行贷款余额超过2000亿元(据笔者了解,在高校扩招以前,全国高校只有发展校办产

业所形成的极少量债务,评估工作结束时所形成的债务,均是评估期间建设任务所产生),从而对于评估高校造成了极大的创伤。正如中国科技大学朱清时校长 2008 年 4 月 5 日接受记者采访时指出:“(评估)初衷是好的,但是这个指标只适合不多的学校和少数类型的学校,对于条件相差比较远的或类型完全不同的学校,造成一些长远的硬伤,还有是把这些学校迅速推向泡沫化。”^[5]

2. 过分强调教师教学,将教学水平等同于学生学习的质量

受文化认知、质量观念与管理体制的影响,美国教育部的职能是有限和间接的,其质量保障机制是通过认证机构来实现对教育质量的间接控制,目的是为公众提供公信力高的高校质量证明。美国中部地区教育认证委员会对高校有效的质量调查界定为:鉴定重要的学生成果,在控制重要的背景变量之后,找出这些成果在学生进入大学后是如何随着时间的变化而得到提高的;确定学生不同的大学经验是如何促成了这种成长变化的;根据这些变化对政策和实践进行积极地改革。由此可见,美国社会对于评估的功能由原来只是学校内部用以改进教学的一种手段,到现在则被改装成服务于公众信息需要的工具,希望通过评估活动来帮助高校向社会展示其教育成果(程程《市场竞争中的高校评估及其范式的更新》,2008 年院校研究高级研讨班专家会上发言材料)。

大学的学习到底由什么构成,对这一问题的不同认识构成了不同的教育质量观。中国本科教学工作水平评估指标体系设置中,对教师与其课堂教学的评估所占权重较大,涉及对教师考核的指标有“师资队伍”、“教学管理”与“学风”三个一级指标以及“师资队伍数量与结构”、“主讲教师”、“管理队伍”与“教师风范”等二级指标,超过评估指标总数的 45%,体现的是一种“声誉与资源质量观”。在“主讲教师”二级指标中,将“主讲教师资格”、“教授、副教授上课情况”以及“教学水平”作为三个观测点,这种对教师和其教学水平的评价使得评估看起来更像是教师评估,并给人一种印象,即优秀教师和他们教学水平就等于学生学习的质量(孙建荣《中美高等教育质量保障的比较研究》,

在 2008 年院校研究高级研讨班专家会上发言材料)。评估指标体系对学生学习效果设置了“教学效果”一个一级指标,二级指标包括“基本理论与技能”、“毕业论文与毕业设计”、“思想道德修养”、“体育”与“社会声誉”等指标,其基本理念是课程考试分数通常被认为是学生学习的一个指标,而毕业论文则可提供学生学习的另一类证据。在评估实施过程中,专家组对学生学习效果的评估,主要是通过“随机抽查”毕业论文以及对随机挑选学生进行限时考核(如计算机应用技能、英语和基础科学等)来完成的。这种测试由于受试学生事先经过训练、监考不严、教师提醒与阅卷问题等因素的干扰,很难得出真实的学生学习情况,从而使这项关键性的测试流于形式。专家组对学生毕业论文的评估包括论文的选题、框架、内容、格式以及评分等要素,评估学校往往是将学生完成论文的最后文本作为档案保存,没有提供教师指导与学生修改论文的过程给专家组查阅。这种忽视教师对学生论文指导过程的检查,是本科教学工作水平评估的一个明显不足。

三、评价体系:效度与信度的问题

评价体系是评估工作的结论。评估工作过程应当在规范、高效、客观与公正的前提下进行,应当在没有各种因素干扰的前提下才能得出正确的结论,评估结论应当能够满足政府部门、评估学校与社会公众对于评估学校教育质量的信息需要。然而,在本科教学工作水平评估过程中,评估工作的效度与信度均不同程度地存在问题。

1. 强调专家团队实地考察,降低了评估工作的效率

美国高校的评估与认证工作是在一种“悄无声息”的状态下进行的,其在评估过程实施中的基本理念是基于“把评估结果与学生投入和经验的数据结合起来,评估才是最有效的”。美国教育部通过全国高校数据库系统收集学校办学状态数据,由认证机构派出几名工作人员到评估高校进行数据核查,他们既没有热烈的欢迎仪式与高规格的接待,也没有各种形式的报告会与汇报会,是在一种大多数人均不知不觉中进行的评估活动。学校的主要责

任是及时、客观地提供办学状态数据信息,否则将被取消享受的政府资助资格。

受传统习俗与文化认知的影响,中国的本科教学工作水平评估是在一种“领导高度重视、事先准备充分、师生全员参与与接待服务周到”的热烈气氛中进行的,评估学校的宗旨是“一切为了评估”。这种评估方式不仅影响了评估的效率与结论的客观性,而且还导致机会主义的滋生,助长了学术不正之风。首先,参与评估过程的专家人员构成问题。评估专家组的成员大多数是学校的行政管理人員和部分教师,他们有多年的教学与管理工作经验,很容易从本校与本学科的角度,将个人的经验与感觉带入评估工作之中,混淆了评估与咨询之间的界限,从而影响评估结论的有效性。同时,专家组成员基于本学科背景与工作经验提出的问题及建议,很容易被单位领导接受,在一定程度上“妨碍”了评估院校维持原本已有的特性。^[6]其次,专家组在实地考察中太多的仪式性活动,更降低了评估过程的有效性,引起了来自学术界和公众的尖锐批评。这些形式主要表现为提供一些后勤服务、特权与接待的礼遇等,一些必需的活动(如开会、访谈、收集信息等),这些都被作为实地考察的一部分,使得评估带有“主观偏见”,对社会正确理解评估的过程产生了消极的影响。同时,专家组在实地考察期间,主要通过收集信息与依靠评估标准等方式来做出判断,收集信息的方式主要是通过事先准备会议人员与发言材料的报告会来获取的,由于参会人员的发言内容事先经过领导的引导,专家组以此得出的结论难免受其影响。

2. 评估结果无区分度,混淆了公众的视听

从20世纪80年代中期开始,美国一些州的州长以及一家著名的教育研究所分别发表报告,要求大学建立“问责”制度,明确公布其教学绩效标准,向社会报告学生学习成果,并针对评估结果采取相应的行动(Ewell, 2005)。从那以后,政府不断代表公众发布指令,要求高校扩大评估的范围,以应对社会问责。美国的高等教育质量保障机制经过了几十年的规范与完善,应该都是比较健全的,但到了2006年,美国联邦教育部长斯佩林斯任命的高等教育未来委员会发布的报告对高等教育中“问责机制

的严重缺失”提出严厉批评,该委员会认为,美国现有的中学后教育系统中根本没有一个通盘的规划,以保证从内部能够产生足够的问责机制并对外界发布有用的公共信息;高校起码应为高等教育的“消费者们”建立一个简明易懂的数据库,帮助学生及其父母比较不同大学的学生在校期间学业收获(程颢《市场竞争中的高校评估及其范式的更新》,2008年院校研究高级研讨班专家会上发言材料)。这一批评不只是给很多学校现行的评估项目打了不及格,而且也宣告了评估专家们多年研究工作的失败。

中国政府由教育部评估中心组织的本科教学工作水平评估,从2003年起“五年一轮”,到2005年底,已经评估171所本科院校,其中93所获得优秀、66所获得良好、12所获得合格;优秀、良好、合格的比率分别为54.4%、38.6%和7%(刘凤泰《高度重视评估整改,巩固发展评建成果》,在本科教学工作水平评估整改工作研讨会上的讲话,2006年7月16日)。关于中国高校评估的公信力问题,纪宝成校长2008年3月26日在《人民日报》发表评论文章指出:“有的高校办学质量明明比较差,社会上反映的问题也很多,但最后评出来的结果,可能还相对很不错。这就源于评估体系、方法过于单一僵化,评估的方法基本差不多,评估结果自然也差不多。如果太注重形式,许多评估也就都流于形式了。”笔者认为,中国高校评估的公信力明显欠缺:一是评价结论设置导致机会主义的问题。由于评价等级没有区分度,无论是清华、北大,还是一般大学评定的等级均是按“优秀”、“良好”与“合格”三个等级标准进行评定,容易混淆公众的视听,给一些高校产生投机的心理。因此,在一部分评估院校领导的理念中,评估既是一次挑战,也是一次展示与发展的机遇,尤其是对于部分地方院校,在评估中能够获得“优秀”,不仅可以与地方名校处于同一个“档次”,而且可以获得与国内名校相同“级别”的“国家级招牌”,这将为今后招生宣传与争取“学位点”等资源创造有利条件,许多高校不仅将“迎评工作”作为学校一切工作的“中心”,而且设立专门的“迎评办公室”,目标

(下转第89页)