

# 我国大学评价指标体系研究

王 琴

(中国科学院科技政策与管理科学研究所 北京 100080)

**【摘 要】** 通观我国大学评价的指标体系,虽然众多,但是迄今尚无较为统一与权威的评价指标体系,本文就是在这样一种情况下,进行大学评价指标体系的研究。首先,概述美国、英国等国大学评价指标体系和国内具有代表性的大学评价指标体系,然后分析了这些指标体系值得借鉴的地方和存在的问题并提出修改的意见,最后得出一种新的面向大学综合实力评价的指标体系架构。

**【关键词】** 大学;评价;指标体系

**【中图分类号】** G644

## A Study of the Index System for Appraising Chinese Universities

WANG Qin

**【Abstract】** Though there are so many indexes in our countries, no one can be the most authoritative. This paper is research under these cases. It firstly introduces the current situation of universities evaluation in the home and out. Then it analyses the main problems of them. In addition, it improves these index and a new index for appraising Chinese universities is put forward at last.

**【Key words】** university; appraising; index system of evaluation

### 1 引言

大学评价工作体现的是一种高校能力与实力的社会认可,是对高校办学水平、办学质量、办学效益的一次综合性考察,能增强社会对大学的了解,促进高校发展。而评价的指标实际上是一种价值观,能够引导高校不断改善工作,加强创新,提高教学能力及其社会声望。大学评价具有参谋功能即能对高校的现状进行诊断评估,也具有社会导向功能、学习功能,对高校还有管理功能和监督功能。在一定程度上,如果高校发展实在不能让人满意,可以根据评价的结果对高校的资源和其他的方面做一些调整,这就是制衡功能。但是大学评价又是一项很复杂的活动,它所包含的内容十分丰富,因而评价的结果也引来很多的争议。

大学评价的重要性引导人们不断探索其方法,方法主要包含四个方面:指标体系的设立、权重的选取、数据的收集、计算方法的选择。本文只讨论其中的指标体系问题。在分析目前国内外指标体系的基础上,总结了他们的可取与不足之处,提出了修改的

意见,最后得出了一个综合性的大学评价的指标体系架构。

### 2 国内外大学评价指标体系现状

在影响大学排行的诸要素中,指标发挥着基础的功能。指标的科学性、合理性与评价结果的信度与效度直接相关。缺乏科学合理的指标,评价就难以公正合理,也难以引导被评价者走向良性发展的轨道。指标只有建立在科学、合理的基础上,才能发挥其评价、引导的作用<sup>[1]</sup>。

世界各国都针对本国国情,研究出符合本国实际的大学排行榜,有些排行榜的指标体系在竞争中逐渐成为本国高等教育发展的参照指标<sup>[2]</sup>。

美国的《美国新闻与世界报道》在 1983 年率先推出全美大学排行榜,它所采用的评估指标体系由 7 个一级指标、16 个二级指标构成,主要反映高等院校教育投入和产出方面的指标,英国的《泰晤士报》和《泰晤士高等教育副刊》于 1986 年开始推出英国高校排名<sup>[3]</sup>,《泰晤士报》大学排名的指标每年作相应调整,由 1994 年的 15 项,调整发展为 2002 年的 9 项。

此外,还有像德国的《明镜》周刊、日本的《钻石周刊》、《商业周刊》、加拿大的《麦克林》、俄罗斯的《职业》杂志、《福布斯》、《新闻周刊》、《每日电讯》、《卫报》等也都对本国的大学进行过单项或综合的大学排名。总的来看,世界发达国家有影响力的重要教育评估机构和媒体大多都推出了各自国家的大学排行榜。在发达国家大学评价的影响下,发展中国家也纷纷仿效,推出本国的大学排行榜,如1997年印度《今日印度报》就对印度的大学进行了排名<sup>[4]</sup>。

在我国,1987年中国管理科学研究院科学学研究所评出了中国第一个大学排名。近年来,大学评价工作越来越受到社会的关注。据统计,目前我国有17个进行大学排名的单位,先后已经发布了30多个不同类型的大学排名。虽然参与评价的机构众多,但迄今尚无较为统一与权威的评价指标体系。

通过文献分析,笔者发现我国大学排名的指标体系是在不断发展完善的,从刚开始单一的指标体系,到后来的多指标和多维度的指标体系。如目前网大和广东管理科学研究院的武书连大学排名。1993年开始,广东管理科学研究院的武书连先生对7个大学排名,包括综合排名、单项排名及学科、专业、研究生院等排名,有2个一级指标:(1)人才培养;(2)科学研究。下设子类指标,根据专家调查法确定各项指标的权重。而后来的网大评出的中国大学排名榜,沿用《美国新闻与世界报道》体系,设6个指标:学术声誉、专家评分、论文、新生质量、师资、科研经费。每次各作相应调整<sup>[5-6]</sup>。

国内的大学排名在一定程度上激发了高等学校的办学活力。如科学论文SCI的排名就极大地刺激和促进了中国大学的科研和科学论文的发表。

### 3 我国大学评价指标体系分析

网大和广东管理科学研究院武书连的大学排名在所有的大学排名中影响力最大,对他们指标体系的讨论也最多。网大《中国大学排行榜》是考虑声誉、投入、产出的多方位评价,定性与定量相结合;重视相对量,强调招生的质量;既重视现在,除评价公布年度前1~2年的成绩外,又把积累性的指标,如两院院士、长江学者、博士点、硕士点、国家重点学科、研究基地等列入评价指标。武书连先生的排行榜全部使用原生数据,重视总量,特别是招生的数量,重视现在,只评价公布年度前2~6年的成绩。虽然他们二者每年都会对指标体系给予修改和完善<sup>[7]</sup>,但是这两个指标体系都还存在一定的问题,对

此,很多专家学者做了比较详尽的评论。

笔者归纳了一下,关于国内学者对这两个大学排名指标体系主要有以下几点看法:

#### (1)局限问题

大多数的学者都认为这两个指标体系中的指标不完善,甚至是非常局限。一些学者对部分指标提出质疑,如每千名学生人均国家重点学科数,他们认为国家重点学科多少反映了一个学校学科实力,与在校学生人数没有太大关系。

#### (2)规模与排名问题

现在的排名榜都很看重学校的规模,认为学校规模越大则排名应越靠前。如浙江大学、华中科技大学、中南大学等高校都因合并综合实力一下子跃居中国高校前列;而中国科技大学、南开大学、北京师范大学等院校因为没有合并而相对来讲综合实力排名有所下降。这些都是很不合理的。

#### (3)软指标问题

有学者认为,武书连的指标体系中,不仅硬指标不全面,而且软指标完全没有。如在他的指标中,没有学校声誉等指标。造成在指标体系上的片面性<sup>[8]</sup>。

#### (4)重理轻文问题

在中国大学排名榜中“重理轻文现象”广泛存在,体现文科院校科研和学术实力的指标在评价体系中很少体现,而且评价指标中人文科学相关的指标“权重”也过小,这直接导致了以文科为主的院校在排名中名次明显偏低<sup>[8]</sup>。

#### (5)前提假设问题

武书连的指标体系有一个虚拟了的假设,即“不同学科的科研人员平均具有相同创新能力”,大多数的学者认为这个假设根本立不住脚。因为按这个逻辑,学术大师与新助教是不存在区别的,但事实上科学创新能力的高低并不以研究人员的多少为必要条件。并且这个假设还可能导致这样一个结论,就是哪个大学的教学科研人数多,就能得分多,排名也自然会靠前。这显然是近乎荒谬的。按这个逻辑,各大学要想排名靠前,就必须搞人海战术了<sup>[9]</sup>。

笔者认为,还有些东西,被目前国内的学者忽视了或很少有人提及,指标体系中都没有考虑要计算负指标,即没有扣分项,这与社会的发展并不相符。近几年,高校频繁地出现招生等丑闻,如果指标中没有这一扣分项,只在权重上加以调整,则不能很好地引导高校注重声誉,防微杜渐;另外,指标体系也没有反映教学科研的质量,“质量”这一概念在高校评估中是很重要的,如果只注重一些表面的东西而忽

视这种本质的质量,笔者认为是不可取的。

#### 4 对指标体系的改进

大学评价指标体系是大学发展的指挥棒。建立大学评价指标体系最根本的目的是促进大学不断的发展。中国大学排行榜缺乏科学性,主要是由于在中国进行大学排名尚无一套比较科学、并符合中国高等教育实际的评价体系。即使是由教育部来排名,这个问题也同样存在。所以,在分析了前面提到的一系列问题后,笔者尝试着提出了一个新的指标体系。

针对指标局限问题,笔者根据实际情况,增加了一些指标,并减少了一些很模糊的指标,如生师比等,因为这样的指标很难确定到底多少才能算好;在解决规模与排名问题上,笔者不仅考虑了总量,而且还考虑了均量;在解决软指标问题上,笔者增加了声誉指标,并且把声誉指标分成了学术、社会、国际声誉,对不同的声誉分别进行统计;在解决重理轻文问题上,笔者在指标体系中增加了社会科学的指标和权重,在进行声誉调查时,增加了社会科学界的专家、学者。

因为网大只有两级指标,而它的一级指标和笔者的新指标体系及武书连的指标体系的二级指标相对应;它的二级指标和笔者的新指标体系及武书连的指标体系的三级指标相对应。所以在比较过程中,笔者做了相应的调整。

##### 一级指标比较

新指标体系	网大的指标体系	武书连的指标体系
声誉水平		人才培养
教育条件		科学研究
教育质量		
科研条件		
科研质量		

这套指标体系是笔者参照目前世界各国以及国内已有的科学评价体系,从国情出发,对我国高校的综合竞争力进行的较为全面系统的考评和计量。指标体系中比较明显地包含了投入、产出的理念。

(1)首先,现存的指标体系中值得借鉴的指标,基本上都在新的指标体系中体现出来了。

(2)新的指标体系一共是三级指标,对教育、科研、声誉等进行排名。考虑了数量与质量、规模与效益等方面的关系对各方面的情况进行了综合,同时以往评估的指标体系中没有考虑过负指标,虽然很多专家提到过,后来的指标体系改进也只是在权重上作了一些调整,不能从根本上解决问题。

##### 二级指标比较

新指标体系	网大的指标体系	武书连的指标体系
学术声誉	声誉	研究生培养
社会声誉	学术资源	本专科生培养
国际声誉	学术成果	自然科学研究
教育资源	学生情况	社会科学研究
校园硬件设施	教师资源	
经费情况	物资资源	
教师资源		
学科建设		
教育基地		
人才培养		
教师质量		
学生质量		
管理队伍		
经费情况		
科研人员		
科研基地		
重大项目		
技术转让		
学术成果		

(3)新的指标体系引进质量评估概念。办学质量由教育质量与科研质量分别评估,一方面保证了高校基本的办学条件和教育质量;另一方面又鼓励少数著名大学在高层次人才培养上进行学术竞争,形成社会影响,并在一定程度上发挥评估的导向作用,防止大学过分重科研轻教育。

在教育质量中增加了大学的学科建设。这种做法不仅增加了可比性,使评估结果更具说服力,而且还引导学校走内涵发展道路,办好自己的优势学科,同时一些名牌大学的劣势学科也不能获得高于实际的声誉,这进一步维护了评估的客观公正性,对非名牌大学起到了很好的激励作用<sup>[10]</sup>。

教育质量其实不仅仅包含着老师对学生的教育,也包含学校的管理水平,而这种管理水平在很大程度上是与教育管理者有关的。所以本指标体系加入了“管理队伍”评价指标,从各职能部门和院系部门的运行绩效、学生和教师的满意度、管理人员占有所有教职工人数的比例、是否企事业分开和后勤工作社会化、本届校长任职期间学校发展状况等五个方面加以评价。

(4)大学评价不能忽视总结性评价的“特殊功能”<sup>[11]</sup>,所以仍要进行必不可少的鉴定和教学核定,以检验教育实施的效果。但是大学评价更为注重的是形成性评价,即直接指向正在进行的活动的活动的评价。正如 Catherine A. Palomba 和 Trudy W. Banta 所说,“对结果进行评价意味着终结,而对过程进行评价则暗示着还有改进的时间和机会”<sup>[12]</sup>。

三级指标比较

新指标体系	网大的指标体系	武书连的指标体系
图书馆藏书量及生均藏书量	两院院士、知名学者、专家、大学校长和中学校长调查结果	博士生培养
生均校园建筑面积	博士点(每千名学生拥有量)	硕士生培养
生均教室占有面积	硕士点(每千名学生拥有量)	研究生创新环境
总量教育经费,生均教育费用	国家重点学科(每千名学生拥有量)	本科生培养
教授比例	国家重点实验室/国家工程(技术)研究中心(每千名学生拥有量)	本科生源质量
长江学者比例	国家人文社科重点研究基地(每千名学生拥有量)	SCI,S&N,EI,SCICITA,CSCD,CSCD-CITA,专利,国家级奖,省部级奖
院士比例	SCI	新华文摘,SSCI,A&HCI,CSSCI,专著,国家级奖,省部级奖
正高级以上人数比例	EI	
博导比例	SSCI	
博士学历人数比例	中国社会科学引文 CSSCI	
国家重点学科	录取新生质量(高考成绩)	
博士、硕士学位授权一级学科	研究生在全校学生中所占比例	
硕士学位授权学科、专业	教师中副教授以上人员的比例	
理科基地、文科基地、工科基地,	两院院士人数	
教育部重点实验室	长江学者特聘教授人数	
教育部工程研究中心	师生比(专任教师人数/学生人数)	
教育部社科重点研究基地	科研经费总量及专任教师和科研机构人员人均科研经费	
本科生毕业人数	图书总量及生均图书量	
博士点、硕士点、博士后流动站	校舍建筑面积及生均面积	
授权专利、专著情况		
高校青年教师奖		
国家杰出青年科学基金		
跨世纪优秀人才		
教育部人文社会科学奖		
学生就业率		
全国英语六级通过率		
国际级奖,国家级奖,省部级奖,发明情况,拥有专利、专著情况		
全国百篇优秀博士学位论文入选情况		
各职能部门运行绩效、院系部门运行绩效		
学生满意度、教师满意度		
管理人员占有所有教职工人数比例		
是否企事分开、后勤工作社会化		
本届校长任职期间学校发展状况		
总量科研经费,人均科研经费		
科研机构人员比例		
国家重点实验室,国家工程研究中心,国家实验室,国家批准的文科研究中心,人文社科重点研究基地		
获国家自然科学基金资助的项目,获国家社会科学基金的项目,973项目,863项目		
技术转让合同数		

续表

新指标体系	网大的指标体系	武书连的指标体系
当年成交金额总数		
SCI, S&N, EI, CSCD, SSCI, A&HCI, CSSCI, 及新华文摘收录论文数		
SCI、CSTPC 被引次数		
国家技术转移中心		
国家级大学科技园和文化科技园数量		
科技成果转化与产业化		
专家打分		
社会人员的认同度,学生的认同度,国内同行的认同度,校友的满意度,		
留学生、国外交换学生占全校学生比		
出席国际学术会议		
到国外进行访问研修,向国外派遣研究生		
补考率、违纪率		
事故率		
严重有损声誉的事件,在国家权威媒体上有过负面报道		

科研条件这个指标里增加了“技术转让合同数”和“当年成交金额总数”,符合现在发展了的大学变化情况,这是一个形成性评价。前面的“管理队伍”中“本届校长任职期间学校发展状况”也是一个发展性的指标,属于对过程进行评价。

(5)本指标体系把声誉情况分三部分考虑,即学术声誉、国内声誉和国际声誉。学术声誉定量打分,而国内声誉分为“社会认同度”和“校友满意度”两个子指标,“社会认同度”又分为“社会人员的认同度”、“学生的认同度”、“国内同行的认同度”三个子指标。这样的软指标一进入指标体系就更加完善与充实。而“国际声誉”部分主要参考国际同行的评价和留学生、国外交换学生占全校学生比。

(6)负指标主要是一些减分的指标。这些年来,随着高校的不断发 展,人们对高校的期望值越来越大,高校中的一些丑闻也受到了人们的关注。如果在指标体系中没有这方面的体现,则整个体系很不完整。因而,笔者在综合考虑各方面因素后,决定把负指标纳入其中。

本指标体系只是涉及中国大陆的一些高校的情况,作为一个评价方面的探讨。随着我国教育统计数据库的完善,指标体系还有待进一步的调整。另外,在做完今年大学排名榜的工作后,下一步需要结合实际情况和一些有用的模型,对高校的情况作一个比较科学的预测。

## 5 结束语

虽然我国的大学评价工作已取得了较大的进步,但是从整体上说,还不是很健全,尤其是缺乏影响大学排行的关键因素的评价指标。通观我国大学评价指标体系,虽然众多,但是迄今尚无较为统一与权威的评价指标体系。要做到真正科学的大学评价,就必须要有 一个由定性指标和定量指标构成的完善的科学评价指标体系。大学排名是一个复杂的系统工程,评价体系、指标选择不同,其评价结果将产生明显的差异。因此,完善评价指标体系一定要和评价目的、大学分类相结合,从而实现不同目的、不同类型大学的评价有各自不同的评价指标体系。

虽然目前的大学排名指标体系还需进一步完善,但只要我们能以科学求实的态度去做好大学排名的任何一个环节的工作,必然会对大学评价工作起到相应的推进作用。本文对大学评价体系的改进是笔者近年来研究的一些心得,不足之处肯定存在,仅当抛砖引玉,希望我国能尽快找到一个解决大学排名的科学的评价方法。

## 参考文献

- 1 戴跃侬.大学评价体系及其排行榜透视,扬州大学学报,2003(1)
- 2 中国需要什么样的大学排行榜? <http://www.southcn.com/news/china/zgkx/200401140420.htm>

3 中外大学排名综述 <http://edu.qq.com/a/20031114/000127.htm>

4 李越,叶赋佳.透视大学评价.教育发展研究,2002(2)

5 <http://rank2001.netbig.com>

6 武书连,吕嘉,郭石林.2004年中国大学评价.科学学与科学技术管理,2004(1)

7 蔡言厚,蔡莉.我国大陆“985”工程大学综合评价的探讨 <http://edu.tom.com/1268/1807/20041217-31041.html>

8 关于“大学排名”的问题 <http://cssd.acca21.org.cn/2003/news0220.html>

9 王处辉.论大学排名的科学性与严肃性问题.南开大学,[http://www.nbedu.net.cn/article/show\\_article.asp?ArticleID=9946](http://www.nbedu.net.cn/article/show_article.asp?ArticleID=9946)

10 刘晓红.国外高等教育评估制度对我国高教评估的启迪.北京大学学报,2004(1)

11 陈玉昆.中国高等教育评价论.广州:广东高等教育出版社,1993

12 Catherine A. Palomba & Trudy W. Banta: Assessment Essential Planning, Implementing and Improving Assessment in Higher Education. Jossey Bass Inc. Publishers 1999

作者简介

王琴,女,中国科学院科技政策与管理科学研究所 2004 级硕士研究生,研究方向:评价与管理、公共政策。

(上接第 12 页)

学目的的程序之一,这些程序的完成以及完成的好坏只是表明学生完成了专业学科规定的内容,但并不表明其学习能力。因此,评估按其重点可分成两大类:教学完成评估和学习能力评估。教学完成评估可以看成是教学大纲的组成部分之一,用来检查学生是否完成了教学大纲规定的学习内容;而学习能力评估则是对学生对所学知识的使用能力进行评估,也就是现在被愈来愈多的学校所用的效绩评估,这种评估方法的指导思想 and 操作方法具有以下的特点:

**指导思想:**评估的目的是通过特定的评估活动使学生有机会来展示自己的应用能力,同时去发现并提供让学生补习那些还没有掌握好的知识或能力的机会,也就是说,评估本身是给学生提供一个再学习的机会。通过评估,学生对自己的实际能力有一次比较直接的检验,可以对下一步的学习进行相应的调整。

**操作方法:**将评估的具体目的与参与评估的学生进行交流,评估测试的方法也要清楚地向学生做解释。评估者与被评估者的交流在很大程度上体现了当今时代特点的教学指导工作特征,即学生在学习中应确实起到学习主人翁的作用,不论是学习目的制定、学习过程的执行,还是学习结果的测试,都要最大限度地学生参与。在评估过程中,评估者希望被评估者能真正地将自己的实际水平和能力通过完成所规定的评估活动而表现出来,而与评估者交流评估目的和测试的方法则是要获得被评估者

更好的“配合”,创造一个“轻松”良好的评估测试环境,在这种环境中获得的评估信息具有较强的测试成绩真实性(authentic)和可靠性(reliable),同时测试成绩提供的多项测试点的信息,使得每次评估活动都可以达到多项评估目的。

现行的课堂(考试)测试对学生学习水平和能力水平的评估比较单一,测试结果很难提供一个准确的图像,造成这种情况有多种原因,有些原因是无法控制的,例如,学生学习水平和能力只有在学生毕业走上工作岗位后才能真正得到检验;有些原因是由于测试工具和方法缺乏科学性,不少课堂中的考题均由授课教师自己设计,这些教师尽管有很强的专业知识,但从未受过有关教学法的训练,缺乏有关测试理论的知识,设计出的考试题在其可行性(validity)和可信性(reliability)上有一定的缺陷。因此,仅仅依靠教学中的课堂考试(成绩)来作为主要对学生学习水平和能力水平的测试,会得出片面的评估结果,同时也会给课堂教学带来一定的负面影响。这样分析并不是要排斥教学中课堂考试的功能,如果课堂考试能与课堂教学目的相吻合(alignment between learning objectives and their assessment measures),那么课堂考试(成绩)会提供很有实际意义的评估信息。这就要求我们教育工作者积极开展有效的教学(法)改革。

作者简介

孙建荣,男,美国俄亥俄州瑞欧戈兰大学博士,美国高等教育中北部认证认可委员会咨询员/评估员。