

世界一流大学的建设与评价： 国际经验与中国探索^{*}

周光礼 蔡三发 徐贤春 王顶明

摘要：“双一流”建设的目标是建设世界一流大学，建设世界一流大学离不开国际标准与国际参照。在全球两大高等教育体系中，盎格鲁-撒克逊体系传统上遵循市场逻辑，奉行自由主义理念，认为世界一流大学是自由竞争的产物。罗马体系遵循行政逻辑，奉行国家主义理念，强调世界一流大学的产生是政府指定的。东亚地区的高等教育具有浓厚的欧洲大陆体系色彩。与此相适应，盎格鲁-撒克逊体系实行院校自主的改进性评估，罗马体系实行政府主导的问责性评估。“双一流”建设需要国际可比指标，我们可以通过世界大学排行榜的元评估解析出有用的评价标准。“双一流”建设的元评估包括技术性维度和合法性维度。通过对QS、THE、U.S.News、ARWU的元评估，发现世界大学排行榜在技术性标准上存在缺憾，在合法性标准上严重不足。因此，不能直接将其作为“双一流”建设动态监测和成效评价的依据。

关键词：世界一流大学；“双一流”建设；大学排名；元评估

“双一流”建设的目标是建设世界一流大学及推动高等教育强国建设。建设世界一流大学和高等教育强国离不开国际标准与国际参照。因此，全方位搜集国内外各类大学评价体系，具体分析其评价定位、关键指标、数据渠道及发展趋势，有利于提炼世界一流大学评价的国际可比指标；全面梳理全球主要国家和地区世界一流大学生成的历史进程，总结其政策目标和实践经验，有利于寻找高等教育强国建设的国际参照。本文围绕世界一流大学的建设与评价，总结和提炼了全球主要国家和地区世界一流大学的建设经验；介绍了全球主要的大学评价体系，分析了其评价标准与方法；讨论了世界大学排名的元评价，旨在为建构中国特色的“双一流”建设动态监测与成效评价指标体系提供启示与借鉴。

一、世界一流大学建设的国际经验

根据政府与市场在高等教育资源配置中的作用不同，全球高等教育可以分为两大体系：盎格鲁-撒克逊体系（英美体系）和罗马体系（欧洲大陆体系）。盎格鲁-撒克逊体系传统上遵循市场逻辑，奉行自由主

义理念，认为世界一流大学是自由竞争的产物。尽管在美国有加利福尼亚州、德克萨斯州这样以规划引导的区域高等教育体系，但是整个高等教育系统的分层原则是以各个院校相互竞争为基础的，正如马丁·特罗所言，“各个院校通过在市场上竞争获得有助于提升学术名望的条件，如学术声誉、著名教授、研究经费、捐赠等，从而提高院校地位。在这种地位体系里，决定一所院校名望和生存机会的因素是在上述竞争市场上的成功与否”^{[1]182}。罗马体系遵循行政逻辑，奉行国家主义理念，强调世界一流大学的产生是政府指定的。这种高等教育系统的分层原则是政府主导的，大学等级是由政府分配给各院校和高等教育各部门的职能、权利、特权和资源决定的。“其分配方式是政府政策和其他措施的反映，而政府正是通过这些政策和措施来控制由其提供经费的院校的学术地位和生死存亡。”^{[1]182}由此可见，与英美国家的高等教育治理不同，政府行为在欧洲大陆国家的高等教育治理中扮演了重要角色。欧洲大陆国家的世界一流大学建设是以政府名义在高等教育发展过程中实施的战略行动。

^{*} 本文系教育部人文社会科学研究基金项目“学科生态视野中的我国一流工程学科发展战略研究”（18JJDGC037）的研究成果

东亚地区的高等教育具有浓厚的欧洲大陆体系色彩,这是因为强调政府主导是后发型国家赶超早发型国家的共同选择。与英美这些老牌的资本主义国家相比,东亚国家都是后发外生型国家。东亚地区的世界一流大学建设最初都是模仿和借鉴西方大学。值得指出的是,世界一流大学是一个东方式的概念,源于中国高等教育的重点建设政策。世纪之交的“985工程”将世界一流大学由一个政策概念转变为一个学术概念。当然,中国建设世界一流大学的动因既有外部国际竞争的压力,也有建设创新型国家的内在需求。正如习近平总书记所言,“我们对高等教育的需求比以往任何时候都更加迫切,对科学知识和卓越人才的渴求比以往任何时候都更加强烈”^[2]。随着中国经济的崛起,中国的世界一流大学建设具有很强的“外溢效应”。当中国启动“985工程”建设时,欧洲大陆国家基于罗马传统的文化基因和国际竞争的现实需求也纷纷启动各自的世界一流大学建设工程。如德国于2006年启动了“卓越大学计划”,投入资金19亿欧元以提升大学竞争力;法国于2011年启动了“卓越计划”,旨在建设5~10所世界一流的研究型大学;俄罗斯于2013年启动了“俄罗斯大学竞争力增强计划”,力争2020年前使至少5所俄罗斯大学进入世界大学排名TOP100。英美国家虽然实行高等教育治理的“市场模式”,但受其他国家建设世界一流大学的刺激,政府依然加强了世界一流大学建设的政策引导。如美国联邦政府主要通过科技政策引导世界一流大学建设,英国以及其他英联邦国家则出于提高高等教育产业的全球竞争力而大力扶持顶尖研究型大学。东亚国家和地区反响最为强烈。如日本于2002年实施了“21世纪卓越中心计划”(COE Project),2014年又启动“超级国际化大学计划”(“Top Global University Project”);韩国于1999年启动了推进一流大学建设的“面向21世纪的智力韩国计划”(“BK21计划”),2008年又公布了“建设世界一流大学计划”实施方案,即“WCU计划”,以加快培育世界级高水平优秀大学。台湾地区于2006年启动了两期“迈向顶尖大学计划”。泰国于2010年启动了国家研究型大学项目(2010—2012年),菲律宾于2010年启动了公立高等教育改革的路线图^[3]。

欧洲大陆体系的特点是强调高等教育发展中的政府行为,英美体系的特点是让市场在资源配置中起决定性作用。这两大体系对世界一流大学建设成效的评价也不一样,欧洲大陆体系实行政府主导的问责性评估,英美体系主要依靠院校自主的改进性评估。问

责性评估是一种不对等的社会关系,总是与责任联系在一起。问责性评估常常伴随惩罚性行为,故很难得到大学的积极配合。改进性评估强调评估的目的不是问责,而是促进大学发展,这种“自愿问责机制”能够得到大学的全面配合^[4]。在评估的技术与方法上,问责性评估和改进性评估有两个共同点:一是组建专家团队,进行定性评估;二是委托第三方评估机构,进行定量评估。建设世界一流大学离不开国际可比指标,世界大学排名榜就是这一需求的直接产物。自从上海交通大学首开世界大学排名以来,全球有各种世界大学排名榜30多个,其中THE、QS、U.S.News、ARWU最具影响力。从这4家世界大学排名榜来看,英美体系的成就远胜于欧洲大陆体系。这一现象为我们反思政府主导世界一流大学建设提供了契机。事实上,世界一流大学可以分为两个层次,一个是指标意义上的世界一流大学,一个是文化意义上的世界一流大学。政府主导对于建设指标意义上的世界一流大学可能是有用的,但是对于建设文化意义上的世界一流大学其作用可能有限。真正的世界一流大学是文化意义上的,因为这种大学具有一定的高度和宽度,高度上超越功利而求“止于至善”,宽度上超越国界而求“世界大同”。指标意义上的世界一流大学是功利主义的产物,以竞争而非合作为目的;文化意义上的世界一流大学以构建人类命运共同体为己任,是真正的世界一流大学。应该说,中国当前已经不缺乏指标意义上的世界一流大学,中国缺乏的是文化意义上的世界一流大学。所谓文化意义上的世界一流大学,就是大学植根于独特的文化土壤,有自身的使命和价值观,有深沉而远大的抱负,能够为其他国家、其他文明提供实质性的启示和借鉴。

世界一流大学的生成有两种不同的逻辑,一种遵循自然演化的逻辑,另一种遵循人为建构的逻辑。英美体系的牛津大学、剑桥大学、哈佛大学、耶鲁大学是在漫长的历史演进中形成的,是在一定的文化土壤中“长出来”的,是物竞天择、适者生存的结果,这是一种典型的自然演化逻辑。东亚地区的首尔大学、新加坡国立大学、东京大学、京都大学、清华大学、北京大学在政府主导下建设而成,是政府指定、人为设计的结果,这是一种典型的人为建构逻辑。人为建构逻辑在短期内或许是有效的,但长时段边际效应却是递减的。大学要走得远、走得长必须遵循自然演化逻辑,这是组织生命的源泉,也是世界一流大学生成的真谛。中国建设世界一流大学,必须结合建设逻辑与演化逻辑,以建设逻辑创建指标意义上的世界一流大学,以

演化逻辑培育文化意义上的世界一流大学,最终由建设路径过渡到演化逻辑,从简单的定量比较走向内涵发展。

二、世界一流大学评估的指标体系与评价方法

(一)问责性评估与改进性评估

欧洲大陆体系强调世界一流大学发展中的政府行为,政府在一流大学评估中发挥主导作用。近年来,随着博洛尼亚进程的推进,政府职能已经由直接控制转变为间接控制。当前的主要做法是组建一个半官方的评估机构或者委托一个经过政府认证的第三方机构负责具体项目的评估。一般来说,其评估分四个阶段:①评估机构组织专家委员会对大学提供的项目草案进行评审,评审通过的项目需提交各自项目的详细论证方案,进入第二轮的评审。②国际同行专家评议。将第一轮通过的项目送交国际同行专家进行评议。其中,同行专家50%以上来自国外著名大学或研究机构。③确定拟资助的正式名单。根据国际同行专家的评审意见,参照世界大学排行榜等国际可比指标,评估机构提出候选项目名单,提交专家委员会评审,专家委员会通过评审提出拟资助项目的正式名单。④资助委员会提出资助方案。资助委员会由相关政府机构负责人、专家委员会成员构成,最终决定资助方案。这种评估本质上是问责性评估。英美体系主张世界一流大学发展中的自由竞争,大学主导的“改进性评估”在质量保障中发挥主要作用。改进性评估是院校自主发起的、以改进自身为目的的评估。一般每5年开展一次学科国际评估。改进性评估包括两个方面,一是邀请国际顶尖同行,对学校的办学条件、教师水平、学术表现等进行为期1周的现场考察,形成专家诊断报告;二是委托第三方机构进行文献计量评估,对学校5年来的学术成果进行统计分析,形成定量分析报告。

这两份报告分别提交学校领导使用。这种评估本质上是“自愿性问责评估”。如果说欧洲大陆体系比较强调外部质量保障体系,那么英美体系比较强调内部质量保障体系。

当前这两种体系出现了相互借鉴和相互融合的趋势。欧洲大陆体系开始重视大学内部质量保障体系,英美体系开始强调外部质量保障体系。实现外部质量保障和内部质量保障的有效衔接是两大体系发展的共同趋势。

(二)大学评估与世界大学排名

当前,在中国“双一流”建设中影响较大的第三方评估有:教育部学位中心的学科评估、美国基本科学指标(Essential Science Indicators,ESI)前1%学科、英国泰晤士报发布的THE世界高等教育排名、英国Quacquarelli Symonds集团发布的QS世界大学排名、美国《美国新闻与世界报道》发布的U.S.News世界大学排名、中国上海交通大学发布的ARWU世界大学学术排名。这些第三方评估机构的评估定位各不相同。教育部学位中心的学科评估旨在监测中国的研究生教育质量、上海交通大学的ARWU排名是监测“985工程”建设成效的副产品、英国的THE、QS和美国U.S.News的世界大学排名旨在为学生以及家长选择留学目的地提供简便的参考、ESI旨在为学者揭示科学发展趋势以及22个学科的研究前沿。下面重点介绍中国高校普遍对标的4家世界大学排行榜。(见表1)

1. QS世界大学排名指标体系。QS的世界大学排名的指标体系包括:学术同行审查(40%)、全球雇主评价(10%)、师生的比例(20%)、教师人均引文(20%)、国际化(10%)。(见表2)

2. THE世界大学排名指标体系。泰晤士报(THE)的世界大学排名的指标体系包括:教学(30%)、研究(30%)、被引频次(30%)、国际视野(7.5%)、工业收入(2.5%)。(见表3)

3. U.S.News世界大学排名指标体系。U.S.News

表1 4家世界大学排名的基本情况

	THE	QS	U.S.News	ARWU
发布机构	泰晤士高等教育	QS公司	美国新闻与世界报道	上海交通大学 世界一流大学研究中心
所在国别	英国	英国	美国	中国
开始年份	2010	2004	2014	2003
排名类别	世界大学排名 金砖国家与新兴经济体排名 世界大学声誉排名 建校50年以下100所大学排名 亚洲大学排名 世界大学学科排名	世界大学排名 世界大学教师水平排名 世界大学学科排名 区域排名 最佳留学城市排名	世界大学排名 美国大学本科排名 美国大学研究生排名 地区大学排名 国家大学排名 世界大学专业排名	世界大学学术排名 两岸四地大学排名 世界大学学科领域排名
排名长度	前800	前800	前750	前500
排名公布	前200名排名次,之后按区间	全部排名次	全部排名次	前100排名次,之后按区间
发布日期	每年10月	每年9月	每年9月	每年8月
数据来源	高校报送数据 Elsevier公司的Scopus数据库针对学者的调查	高校报送数据 Elsevier公司的Scopus数据库针对学者与雇主的调查	高校报送数据 Thomson的Web of Science数据库针对学者的调查	SCIE和SSCI高被引科学家的《Nature》《Science》Thomson的Web of Science数据库

表2 2018年QS世界大学排名指标体系

指标构成	指标含义	权重
学术同行审查	QS公司将大学的“声誉”视为大学排名中最重要的评估指标。得分来自同行评议人的评审统计结果,学者们被要求选出在近3年内他们认为在其领域表现最优秀的学校,不包括他们自己所在的学校。	40%
全球雇主评价	调查对象是在全球或者重要国家招聘毕业生的雇主。得分来自在全球或者重要国家招聘毕业生的雇主反馈,雇主们被要求选出他们认为培养出最优秀毕业生的大学。	10%
师生的比例	师生比例	20%
教师人均引文	5年期间的引文总数除以一个大学专任教师的数量。2016年排名的引文数据来源 Elsevier 的 Scopus 数据库的 23000 个学术期刊在 2010-2014 年发表的文章在 2010-2015 年被引情况分析,文章类型包括:期刊文章、会议论文、评议、书籍及书籍出版章节。 ①在 2016-2017 年度排名, QS 分析了 1030 万篇研究论文和 6630 万被引频次;②教师数据来自各学校。	20%
国际化	国际学生的比例和国际教师的比例	10%

表4 2018年U.S.News世界大学排名指标体系

指标构成	英文原意	权重
全球研究声誉	Global research reputation	12.5%
区域研究声誉	Regional research reputation	12.5%
出版	Publications	10%
专著	Books	2.5%
学术会议	Conferences	2.5%
归一化引用影响力	Normalized citation impact	10%
总被引次数	Total citations	7.5%
前10%高被引文献	Number of publications that are among the 10 percent most cited	12.5%
前10%高被引文献占比	Percentage of total publications that are among the 10 percent most cited	10%
国际合作	International collaboration	5%
国际合作论文发表占比	Percentage of total publications with international collaboration	5%
前1%高被引论文	Number of highly cited papers that are among the top 1 percent most cited in their respective field	5%
前1%高被引论文占比	Percentage of total publications that are among the top 1 percent most highly cited papers	5%

表3 2018年度THE世界大学排名指标体系

指标构成	指标含义	评价方法	权重
教学	学习环境	声誉调查	15%
		师生比	4.5%
		博士生占比	2.25%
		师均博士学位授予数	6%
		师均收入	2.25%
研究	体量、收入与声誉	声誉调查	18%
		师均科研收入	6%
		师均科研论文量	6%
被引频次	研究影响力	归一化影响因子(FWCI)	30%
国际视野	员工、学生和研究生	国际国内学生比	2.5%
		国际国内教师比	2.5%
		合作者为外国人的论文比	2.5%
工业收入	创新-知识转移	师均来自工业界的科研收入	2.5%

表5 2018年度ARWU指标体系

一级指标	二级指标	代码	权重
教育质量	获诺贝尔奖和菲尔兹奖的校友折合数	Alumni	10%
	获诺贝尔科学奖和菲尔兹奖的教师折合数	Award	20%
教师质量	各学科领域被引用次数最高的科学家数量	HiCi	20%
	在《Nature》和《Science》上发表论文的折合数	N&S	20%
科研成果	被科学引文索引(SCIE)和社会科学引文索引(SSCI)收录的论文数量	PUB	20%
	上述五项指标得分的师均值	PCP	10%

注:1.对纯文科大学,不考虑N&S指标,其权重按比例分解到其他指标中。

2. ARWU指标在2003年略有不同,分为Nobel、HiCi、N&S、PUB、Faculty5项。

(三)世界大学排行榜的积极作用与消极作用

世界大学排名指标体系包括:全球研究声誉(12.5%)、区域研究声誉(12.5%)、出版(10%)、专著(2.5%)、学术会议(2.5%)、归一化引用影响力(10%)、总被引次数(7.5%)、前10%高被引文献(12.5%)、前10%高被引文献占比(10%)、国际合作(5%)、国际合作论文发表占比(5%)、前1%高被引论文(5%)、前1%高被引论文占比(5%)。(见表4)

4. ARWU指标体系。“世界大学学术排名”(ARWU)指标体系包括:教育质量(10%)、教师质量(40%)、科研成果(40%)、师均表现(10%)。(见表5)

这4家世界大学排行榜的共性特点有6个。①推崇大型综合性研究型大学模式;②评价覆盖的国家和大学数量越来越多(从100强到500强);③大学的区域性和专业性排名逐渐增多;④排名指标逐渐多样化、科学化和客观化;⑤广泛运用声誉调查和文献计量分析;⑥市场化的企业主体以第三方身份进行大学排名。

尽管美国开展大学排名已经100多年,但世界大学排名却是中国人的创造,其初心是衡量“985工程”的建设成效。世界大学排行榜的积极作用主要体现在3个方面。①提供了大学办学质量的考核方法与评价体系,有利于利益相关者进行决策;②提供了展现大学实力的窗口,促进了全球大学之间的竞争;③大学排名发挥指挥棒的作用,推动了综合性研究型大学模式在全球扩散。

世界大学排行榜的消极作用也主要表现为3个方面。①大学排名体系自身存在缺陷,具体包括大学排名指标权重设置的模糊性;重视量化指标,没有兼顾非量化指标;注重“投入-产出”评价,忽视“发展过程”评价;重视规模数量,忽视内涵式发展;重视理工类高校,轻视人文类高校。②大学排名实施中的问题,具体包括:数据统计处理缺乏必要的审查和监督、易被商业化污染,形成“钱名交易”的潜规则。③大学排名结果的消极影响,具体包括排名的指挥棒易造成高

校同质化、加剧高校发展的马太效应。

当前中国“双一流”建设中存在的一个突出问题是“洋标准”泛滥,部分大学对标4家世界大学排行榜办学。这些排名机构也从中国“双一流”建设中发现了“商机”,纷纷调整服务面向,以中国为主要经营市场,大有全面控制中国大学评价话语权之势。“洋标准”的泛滥严重误导了中国“双一流”建设。

1. 冲击了中国大学的办学道路。大学排名在西方原本是一个生意,但在中国成为了许多“双一流”建设大学办学的指挥棒。部分“双一流”大学,不按照教育规律办学,不按照学校制定的发展战略办学,而是按照各种境外大学排行榜办学,完全背离了大学的使命与初心,偏离了正常的办学道路。

2. 冲击了中国高等教育评价体系。改革开放40年来,中国高等教育督导实现了由督政、督学向评估监测转变。在这个过程中,我们逐步建立起一套有中国特色的高等教育评价体系。然而,随着“洋标准”的流行,我们将“管办评分离”狭隘地理解为“政府管教育、大学办教育、社会评教育”,认为政府没有评价教育的权力,只能让“第三方机构”来从事大学评价,严重干扰了政府的问责性评估。

3. 冲击了中国的学科分类体系。学科分类体系具有地域性和国别性。中国的学科分类逻辑不同于美国的学科分类逻辑。美国ESI学科体系将学科分为22个,美国Web of Science将学科细分为250个小类。我国教育部的学科体系将学科分为13个门类110个一级学科。作为高等教育领域的国家标准,学科分类体系属于国家政策范畴,具有极高的权威性。当前,部分“双一流”大学不遵照中国自己的学科分类体系办学,而是依据国际排行榜确立的学科分类标准办学,动摇了中国学科分类的国家标准。

4. 冲击了中国特色的高等教育体系。一个国家的高等教育是与一个国家的文化模式相适应的。习总书记强调,中国有独特的历史、独特的文化、独特的国情,决定我们只能走中国特色的教育现代化之路,扎根中国大地办大学。中国传统高等教育强调“立德树人”和大学的政治功能。这种传统与西方的3A原则(学术自由、学术自治、学术中立)存在很大的不同。在“洋标准”的导向下,中国传统的大学概念被贬得一文不值,中国大学失去“文化自信”与教育自信。

应该说,世界大学排行榜对中国“双一流”建设动态监测与成效评价仅仅具有些微的参考价值。这是因为两者存在3个方面的不同。①评价定位不同。世界大学排行榜具有商业目的,服务于特定顾客群体,主

要用于大学品牌运作和学生挑选学校;“双一流”建设动态监测和成效评价服务于政府行政问责和高校“双一流”建设。②评价性质不同。世界大学排行榜属于第三方评价;“双一流”建设动态监测和成效评价并不是要形成一个新的第三方评估或大学排行榜,只是为政府提供一个评估和遴选“双一流”建设高校的方案。③评价原则不同。世界大学排行榜强调客观性评价、终结性评价、产出评价及共性评价。“双一流”建设动态监测与成效评价坚持主观评价与客观评价相结合、形成性评价与终结性评价相结合、大学评价与学科评价相结合、投入评价与产出评价相结合、共性评价与分类评价相结合。

三、世界大学排行榜的元评估

“双一流”建设需要国际可比指标,世界大学排行榜的确有可能为中国“双一流”建设与评估提供某种启示与借鉴,但不能照单全收。为此,我们可以引入元评估的思想对各种世界大学排行榜进行二次评估,以解析出有用的评估标准。所谓元评估(Meta-evaluation),就是对评估本身的评估,旨在检验评估中可能出现的各种偏差,运用统计和其他的方法来估计产生的偏差对评估结论的影响。对世界大学排行榜进行元评估,就是按照一定的理论和价值标准对世界大学排行榜的评价技术与方法、数据与结论进行评价与研究,以获得可靠的质量证据。

(一)元评估的基本原理

大学评价中的元评估旨在寻找大学办学的质量证据。以人才培养质量为例,国内外大学评价关于人才培养质量证据的假设很多,如资源优势、高选拔性、学习投入度、高就业率、雇主满意。其中,资源优势和高选拔性属于投入性指标,这种假设把质量的定义集中在成本和选拔性上,暗示教育质量供应是有限的,这个观点构成了世界大学排名和评估的一个哲学基础^[56]。学习投入度属于过程性指标,这个假设把质量定义在学生整个学习过程中,这种观点已经成为大学建立内部质量保障体系的理论基础。学习投入度包括两个方面的内容:一是指学生对学术活动与非学术活动的投入程度,二是指学校吸引学生投入学术活动与非学术活动的力度和能力。根据这种假设,判断人才培养质量高低就看学生学习投入的高低,看学校吸引学生投入学习所做的努力,看学校提供的育人环境是否为学生的学习活动提供有力支持。高就业率和雇主满意属于产出性指标,这种假设强调结果导向的质量评价。应该说,这些假设都从某一个特定的角度部分揭示了人才培养的质量,但都不够全面。这说明用什

么样的事实或指标作为恰当的、操作性强的人才培养质量证据上,应该放弃单一证据说。

现代大学是一个复杂的开放系统,衡量大学办学质量的证据需要有多种数据点。事实上,衡量大学的质量标准有三种可能的选择。①判断的标准,即大学绩效的判断依赖于一个预先制定的判断标准,设计指标体系进行大学的评估与排名就是判断的标准;②比较的标准,即大学绩效的判断依赖于与别的组织的表现相比来评价表现的好坏,寻找同型大学进行对标分析就是比较的标准;③专家的标准,即大学绩效是按照专家的观点和价值观来评判表现的状况,国外大学普遍采用的学科国际评估就是专家的标准^{[5][12]}。由于任何单一证据都会遭到批评,因此应该采取有效的多证据原则,这是大学评价开展元评估的理论基础。

“双一流”建设采用元评估包括两个关键步骤。①研究者提供最佳证据。首先研究者收集所有的第三方评估,然后运用严谨、系统、客观的方法来获得与大学质量相关的、可靠的、有效的数据。这个阶段是一个技术过程,研究者的专业性主要体现在评判“最佳”证据,这需要研究者建立专门的世界大学排名数据库,并根据各种大学排名方法的科学程度进行证据分级,一般来说,可重复、可验证的客观证据高于主观证据。②决策者基于证据进行决策。在“双一流”高校遴选和动态调整上,不能只基于决策者个人的经验和信念,必须基于科学的证据,开展基于证据的决策。这个阶段是一个政治过程,所有利益相关者的意见和关切必须得到公平的展现。决策者应该对各种证据持开放态度,各利益相关者都可以参与质量证据的提供,学校自评结果可以纳入证据的范围。

(二)“双一流”建设元评估的两个维度

开展元评估需要确立评估的标准和框架,根据循证决策理论,“双一流”建设的元评估包括技术性维度和合法性维度。

1. 技术性维度。所谓技术性维度,就是大学的评价的全球共识,这是元评价的技术性标准。大学评价的国际趋势表现在4个方面。①结果导向原则(柏林原则)。大学评价必须坚持以产出指标为主。评价指标应尽可能选择产出而非投入指标。尽管投入指标是大学实现卓越的基础,但其只能反映大学的一般情况,产出指标能提供大学声望和质量更为精确的评价^[6]。这是大学评价的“柏林原则”。②协商评估原则。一切评估都是主观的。大学评价不是一个客观过程,而是一个社会政治过程,有价值观问题。因此,大学评价必须坚持协商性评估,在评估体系的建构中,在质量证

据的收集过程中,要重视所有利益相关者的主张、焦虑和争议,倾听各方意见,通过公开的讨论和辩论形成共识。③多元评估原则。大学评价必须坚持多证据原则,拒绝在单一的指标或证据基础上评价大学的表现,而应采用有效的多证据。从产出的角度看,大学的产出应该包括人才、科研、社会服务以及社会影响。因此,大学办学质量证据的选取需要有多种数据采集点,至少应该包括人才培养、科学研究、社会服务和大学声誉四个方面。④可行性原则。大学评价数据必须坚持可采集、易采集原则。数据的权威性是第一位的,指标体系可以片面,但数据必须干净、没有被污染。

2. 合法性维度。所谓合法性维度,就是大学评价的中国特色,这是元评价的政治标准。中国特色的大学评价体现在4个方面。①突出内涵发展。党的十九大报告明确提出,加快一流大学和一流学科建设,实现高等教育内涵式发展。习近平总书记多次强调,规模扩张并不意味着质量和效益增长,走内涵式发展道路是中国高等教育发展的必由之路。因此,我们要以内涵发展统领大学评价体系的建构,坚决破除唯规模、唯数量、唯论文、唯投入的顽瘴痼疾,从根本上解决大学评价指挥棒问题^[7]。大学办得好坏,不在于规模大小,关键是要办出特色,形成符合自身特点的办学理念和风格。②突出立德树人。2018年5月2日,习近平总书记在北京大学考察时提出,要把立德树人的成效作为检验学校一切工作的根本标准,要把立德树人内化到大学建设和管理各领域、各方面、各环节,做到以树人为核心,以立德为根本^[8]。这为“双一流”建设高校的建设与评估指明了方向。世界大学排名将中国大学导向了“重科研,轻教学”的办学之路,背离了大学的根本任务在于立德树人的初心。以立德树人的成效作为检验学校一切工作的根本标准,才能推进中国大学办学回归常识、回归本分、回归初心、回归梦想,真正把办学的重心放在人才培养上。QS、THE、U.S. News 大学排行榜中的生师比、雇主满意度调查、教师质量、声誉调查可以作为立德树人的指标。此外,生源质量、小班课程的比例、教学成果奖、人才培养的优秀度也可以作为立德树人的指标。③突出分类评估。“双一流”建设与评价必须坚持“先分类后分层”原则,拒绝“一把尺子量天下”。不同类型的学校,评价标准应该不一样,这样才能引导学校准确定位、办出特色。基于知识生产、传播的广度和深度(学科逻辑)和大学的服务面向(社会需求逻辑)两个维度可以形成二维四象限的分类体系,即根据主干学科(一级学科博士点)的覆盖面,可以将大学分为综合性大学和多

科性大学;根据大学的资源配置和服务面向,可以将大学分为部属大学和地方大学。如是,“双一流”大学可以分成四种类型:部属综合性大学、部属行业特色大学、地方综合性大学、地方行业划转大学。^④突出服务国家战略。“双一流”建设实施办法明确提出“三个面向”:面向世界科技前沿、面向国家重大需求、面向经济建设主战场。“三个面向”建基于中国大学的办学逻辑。如果说美国大学的办学逻辑主要是学术逻辑(3A原则),中国大学的办学逻辑首先是社会需求逻辑,然后才是学术逻辑。正因为如此,中国大学比西方大学更强调服务国家、服务社会。“双一流”建设与评估要更加关注大学是否是国家尖端科学研究和技术发展的主要力量、是否是创造知识和产生新思想的主要源泉。具体说来,在评估中,要重视大学专利授予和专利转让,要重视大学的智库成果。

(三)4家世界大学排行榜的元评估

从技术性维度来看,4家世界大学排名体系共涉及17个一级指标,平均拥有4个以上,这表明学术评价机构趋向于多维度寻找大学质量证据;就一级指标的内容而言,都强调研究成果的文献计量、学术声誉和师资队伍,这表明世界大学排名侧重科研与学术发表,而且主要依靠爱思唯尔(Elsevier)的Scopus数据库和汤森路透(Thomson)的Web of Science数据库。如果我们再进一步分析其二级指标,很容易发现:世界大学排行榜只有投入指标和产出指标,缺乏过程指标。由于大学产出很难测量,许多大学排行榜要么倾向以投入指标代替产出指标,要么以主观指标代替客观指标。以投入指标代替产出指标违背了大学评估的“柏林原则”,以主观指标代替客观指标,影响了证据的可靠性,因为可重复、可验证的客观指标优于主观指标。

从合法性维度来看,4家世界大学排行榜主要存在4方面的问题:①重规模,轻质量。重视大学的规模和数量指标,轻视大学质量和内涵指标,不利于推进高等教育内涵式发展。②重投入,轻产出。由于人才培养质量难以测量,世界大学排名机构流行以生源质量代替培养质量。此外,重输入与输出指标,轻视培养过程指标,过程指标缺失不利于落实“三全育人”。③重科研,轻教学。世界大学排名以科研为导向,科研成果权重远远高于教学与社会服务权重。这种“唯科研、唯论文”的导向与“要把立德树人的成效作为检验学

校一切工作的根本标准”指导思想相悖。④重共性评价,轻分类评价。世界大学排名机构将所有的大学视为同质大学,一把尺子量天下。分类评估的缺失不利于引导大学准确定位,办出特色。

总而言之,4家世界大学排行榜在技术性标准上存在缺憾,在合法性标准上严重不足,因此,不能直接将其作为“双一流”建设动态监测和成效评价的依据。当然,这并不意味着我们不能从中解析出一些有用的二级指标。通过世界大学排行榜的元评估,我们发现至少有两个方面的二级指标可以为我所用:一是学生与学习方面的指标,主要包括师生比、博士学生比、国际学生占比、雇主声誉等。二是教师与科研方面的指标,主要包括发文量(总发文、师均发文、顶尖期刊发文)、被引频次(总被引、师均被引)、前1%和前10%高被引文献(数量、比例)、归一化影响因子(归一化引用影响力)、合作者为外国人的论文比、国际教师占比、师均科研收入、师均来自工业界的科研收入、学术声誉等。

(周光礼,中国人民大学教育学院副院长、教授,北京 100872;蔡三发,同济大学发展规划部部长、教授,上海 200092;徐贤春,浙江大学规划处副处长,浙江杭州 310013;王顶明,西北师范大学副校长,甘肃兰州 730070)

参考文献

- [1] 伯顿·克拉克.高等教育新论——多学科的研究[M].王承绪,徐辉,郑继伟,等译.杭州:浙江教育出版社,1988.
- [2] 习近平总书记在全国高校思想政治工作会议上的重要讲话[N].人民日报,2016-12-09.
- [3] 周光礼.世界一流大学建设的“东亚模式”:政府行为及其局限性[J].中国高校科技,2019(4).
- [4] 周光礼.超越问责逻辑,建立大学内部质量评估体系[J].大学教育科学,2012(4).
- [5] BOGUE E G, HALL K B.高等教育中的质量与问责[M].毛亚庆,刘冷馨,译.北京:北京师范大学出版社,2008.
- [6] 周光礼,武建鑫.什么是学术评价的全球标准——基于四个全球大学排行榜的实证分析[J].中国高教研究,2016(4).
- [7] 习近平出席全国教育大会并发表重要讲话[EB/OL].(2018-09-12)[2019-07-13].<http://www.zzyedu.cn/cxb/79/eb/c452a31211/page.htm>.
- [8] 习近平.在北京大学师生座谈会上的讲话[EB/OL].(2018-05-03)[2019-07-13].http://www.xinhuanet.com/politics/2018-05/03/c_1122774230.htm.

(下转第34页)

On Institutional Guarantee of the First-class Disciplines' Formation and Development from the Perspective of Embedded Discipline Governance

HE Xiaofang

(*Dalian University of Technology, Dalian 116024*)

Abstract: The construction of world-class universities and first-class disciplines is not only an important opportunity for universities to achieve leapfrog development, but also a major test of university governance system and ability. To build a world-class discipline, university must pay more attention to the regulations system suitable for the generation and development of first-class discipline, and must transform the macro policy discourse of “discipline construction” into the micro practical logic of “discipline governance”. Discipline governance is in university governance system in an “embedded” way. This perspective challenges the current university governance system from three aspects: the establishment of discipline ecosystem, the optimization and coordinated development of discipline structure, and the straightening out of the relationship between discipline structure and college governance.

Key words: discipline governance; embeddedness; first-class disciplines; university system

(上接第 28 页)

Construction and Evaluation of World-class Universities: International Experience and Chinese Exploration

ZHOU Guangli¹ CAI Sanfa² XU Xianchun³ WANG Dingming⁴

(1. *Renmin University of China, Beijing 100872;*

2. *Tongji University, Shanghai 200092;*

3. *Zhejiang University, Hangzhou 320013;*

4. *Northwest Normal University, Lanzhou 730070*)

Abstract: The goal of “Double First-Class” Construction is to build world-class universities, which is inseparable from international standards and international references. In the world’s two largest systems of higher education, the Anglo-Saxon system has traditionally followed the logic of the market, embraced liberal ideals and regarded world-class universities as products of free competition. The Roman system follows administrative logic, pursues the idea of nationalism, and emphasizes that the creation of world-class universities is designated by the government. Higher education in east Asia has a strong European continental color. In line with this, the Anglo-Saxon system carried out independent improvement assessment, while the Roman system carried out government-led accountability assessment. The construction of “Double First-Class” requires international comparable indicators, and useful evaluation criteria can be analyzed through the meta-assessment of world university rankings. The meta-evaluation of “Double First-Class” Construction includes the technical dimension and the legal dimension. Through meta-evaluation of QS, THE, U.S.News and ARWU, we find that the world university ranking list is deficient in technical standards and seriously deficient in legal standards. Therefore, it cannot be directly used as the basis for dynamic monitoring and effectiveness evaluation of “Double First-Class” Construction.

Key words: world-class university; “Double First-Class” Construction; university ranking; meta-assessment