



中国教育科研参考

2020年第13期
总第(479)期

中国高等教育学会编

2020年7月15日

目 录

我国“双一流”建设成效评价的若干思考.....	钟秉林 王新风 (02)
构建中国特色评价体系 推进世界一流大学建设.....	王战军 刘 静 (07)
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 (13)
“双一流”建设评价的发展性原则及其方法.....	孔晓明 周 川 (22)
“双一流”建设背景下高校评价体系之重构.....	胡仁东 (28)

编者的话:“双一流”建设是我国为推进高等教育事业跨越式发展、建设高等教育强国而实施的一项战略性工程。为了推进这一工程,中央和地方政府出台一系列文件、投入大批建设经费,“双一流”建设高校在政策和经费支持下纷纷加大学科建设力度、提升科学研究水平,创新学校管理理念。目前,“双一流”建设的评估验收和调整期限即将到来,这些高校的“双一流”建设成效如何、如何实施有效评价,成为高等教育管理部门和各大高校关注的热点问题。为了深入认识这一问题,本刊以“‘双一流’建设评价”为选题,集中选编若干文章,供读者参阅。

主编:王小梅 本期执行主编:王者鹤 责任编辑:刘军伟
地址:北京市海淀区学院路35号世宁大厦二层《中国高教研究》编辑部
邮编:100191 电话:(010)82289809
电子信箱:gaoyanbianjibu@163.com
网址:www.hie.edu.cn(中国高等教育学会——学术观点栏目)

我国“双一流”建设成效评价的若干思考

钟秉林 王新凤

中华人民共和国成立以来，在教育资源有限的背景下，我国政府一直采取高等教育重点建设政策。从“211工程”“985工程”到“双一流”建设，高等教育重点建设政策体现了效率优先的基本原则，为我国经济社会发展提供了持续充足的动力。2019年，我国高等教育毛入学率达到51.6%，高等教育进入普及化阶段，“双一流”建设是普及化阶段实现高等教育内涵式发展的重要举措。2020年，“双一流”建设高校将要接受中期建设成效评估，目前教育部正在组织专家紧锣密鼓地构建“双一流”建设成效评价指标体系。对“双一流”建设成效进行科学评价是促进“双一流”建设高质量、可持续发展的关键，有利于引导高校优化学科结构，促进高校与学科的内涵建设和质量提升。对“双一流”建设成效评价的关键是要树立以质量为核心的评价理念，明确关键评价内容之间的辩证关系、评价标准的多重要素考量，完善多元主体的协作体系。

一、以质量为核心的评价理念

20世纪80年代以来，西方高等教育质量评价领域发生了较大转向，受人本主义思潮、全面质量管理、第四代评估理论等理论思潮的影响，高等教育质量评价越来越重视学生中心、结果导向、持续改进、多元参与等理念。“双一流”建设成效评价是当前我国高等教育质量保障体系结构的重要组成部分，也是政府、企业和公众对高校进行问责的重要依据，亦应体现这样的评价理念。当然，与对其他类型高校建设成效的评价相比，“双一流”建设成效评价的侧重点有所不同，比如在结果导向上“双一流”建设成效评价更聚焦于拔尖创新型人才培养和原创性的科研成果，强调学术增量评价等。

（一）学生中心，聚焦一流人才培养

以学生为中心是人本主义心理学的主要思想，强调教育应最大限度地激发学习者的兴趣。以

学生为中心的理念伴随着人本主义思潮的发展而被教育界认可，世界各国在构建高等教育质量保障体系过程中也越来越重视学生的中心地位，强调将以学生为中心的学习和教学作为质量保障的主题，将学生作为核心利益群体吸纳到高等教育质量评价的过程中。在全球化、国际化背景下，拔尖创新人才的培养成为提升国家综合实力与国际竞争力的关键。一流人才培养也是“双一流”建设的重要内涵，与以往的高等教育重点建设政策相比，“双一流”建设遴选标准的制定及其指标体系的设计更加重视一流人才的培养。因此，“双一流”建设成效的评价也应更加重视以学生为中心，这具体体现在建设学校或学科是否将一流人才培养作为“双一流”建设的核心目标，人才培养模式与教育教学改革是否有利于培养世界一流的创新型、复合型人才，人才培养质量及社会声誉是否得到显著提高；是否吸纳学生群体作为“双一流”建设评价的核心利益群体之一，尊重学生的利益诉求，发挥学生在学校内部治理和教育教学质量评价中的作用与功能，等等。

（二）结果导向，强调增量评价

结果导向的评价理念来自于美国学者斯派蒂（William G. Spady）提出的结果导向的教育（Outcome-based Education, OBE）。21世纪以来，该评价理念的运用逐步从基础教育领域拓展到高等教育领域，强调高等教育教学和管理的效能，侧重于评价高等教育系统的产出。“双一流”建设的产出体现为一流的人才和一流的研究成果，即培养基础扎实、创新能力强和综合素质高的拔尖创新人才，产出符合国际标准、体现中国特色的原创性成果。“双一流”建设成效评价一方面应聚焦于学生学习结果的评价，将学生的就读经验与学习产出作为评价教学质量的重要标准，并注重一流研究成果的评价，尤其是基础学科领域面向国家重大战略

需求、解决关键核心技术“卡脖子”问题的成效；另一方面应注重对资源使用效率的评价，将一流人才与一流研究成果作为财政投入的重要依据，坚持以绩效为杠杆，重视各项建设目标与建设方案的符合度和达成度，重视学校内部质量监控与保障体系的有效性，重视资金投入使用效率的评价。一流大学建设成效的评价不应单纯地“数人头”，而应关注一流人才培养和高层次人才领衔的学科建设团队绩效产出的增量情况，包括承担国家重大项目的情况、服务国家战略作出突出贡献的情况等；一流学科的评价也应重视增量评价，不仅要考察科研成果、科研项目、学术人员、招生规模、科研经费等方面是否有增量，更应考察科研团队在研究问题的推进、研究层次的提高等方面是否有增量。当然，如何区分简单增量和学术增量，如何有效考核学术增量，还需要探索新的评估方法。

（三）持续改进，构建动态调整机制

持续改进的理念来自戴明（W. Edwards Deming）为代表的质量管理理论，是指致力于建立常态性的评价机制并不断改进，以建立持续改进的质量文化的一种评价理念。高等教育质量保障领域的持续改进强调建设质量、管理和教育服务水平的持续提高。“双一流”建设要重视滚动竞争、优胜劣汰的价值取向，构建“警示”机制和退出机制，对建设过程偏离建设方案、建设效果与总体建设方案不符的高校和学科进行动态调整，避免出现身份固化、“贴标签”和“只进不出”的现象，促进“双一流”建设工作的持续改进。“双一流”建设既是一项复杂的系统工程，又是一项长期的动态建设过程。建设进程没有终点，建设目标的实现不是靠行政发文“宣布”或学校自我“宣称”，也不取决于学校学生规模大小和学科布局结构是否综合，关键是激发学校的内生动力，集中精力抓好内涵建设，在建设过程中持续改进，在改进过程中不断提升，使人才培养质量、优势学科专业建设水平和办学声誉得到国际学界和国际社会的广泛认可。“双一流”建设评价体系的完善，有赖于以“双一流”建设成果为核心，构建有效的质量监测制度，加强建

设成效的跟踪评价和评价研究。

（四）多元参与，实现多方协同治理

多元参与的评价理念来自以古贝（Egon G. Guba）和林肯（Yvonna S. Lincoln）为代表的第四代评估理论。该理论强调共同构建、全面参与、多元价值，尊重利益相关者的价值构建。第四代评估理论对教育评估领域影响深远，高等教育质量评价越来越重视政府、雇主、学生、家长等多元利益群体的共同参与，从政府或者第三方机构的外部评价或者高校自身的评价逐步走向多元群体的广泛参与，进而实现高等教育质量保障的协同治理。“双一流”建设突出了这种多方协同、价值共建的价值导向，主要表现为在强化高校主体责任的前提下，发挥专家学者的指导咨询作用，重视第三方评估结果的合理运用等。对“双一流”建设成效的评价应在落实高校评价主体地位、履行政府监督问责职能的同时，积极引导专家学者、第三方评估组织等多方主体参与评价，探索构建多主体协同共治的“双一流”建设评价机制与体系，这既有助于优化高等教育治理体系、提升高等教育治理能力，也有助于完善共建共治共享的社会治理制度。

二、成效评价的多重关系辨析

“双一流”建设的目标是建成世界一流大学 and 世界一流学科，其评价的核心内容是一流大学与一流学科建设的成效，其产出体现为一流的人才和一流的研究成果。这需要明确一流大学与一流学科、学科建设与人才培养、科学研究与学科建设等关键评价内容之间的辩证关系。

（一）一流大学与一流学科

一流大学建设与一流学科建设二者相辅相成、相互促进、辩证统一。一流大学建设是一流学科建设的基础，学校整体办学水平的提升为一流学科建设提供物质与精神保障；一流学科建设是一流大学建设的支撑，一流学科建设的成效会引领学校整体办学水平的提升。也有观点认为，统筹推进世界一流大学和一流学科建设在政策支持上可以适当分离，即便一所大学在整体上达不到一流水平，也不妨碍部分学科达到一流水准，我国“双一流”建

设政策也体现了这一理念。因此，一流大学和一流学科建设成效的评价应该明确各自的侧重点：评价一流大学建设学校时应注重其学科综合实力和国际竞争力的全面提升，强调不同学科相互支撑、协同发展；评价一流学科建设学校时则应突出优势学科建设，从而带动学校整体水平提升。

（二）学科建设与人才培养

人才培养与学科建设应进行一体化建设，突出学科的育人功能。人才培养是学科建设的重要内涵，也是学科建设水平的重要体现；一流的人才培养可以有效促进学科建设，提升教师教学与科研水平，其中研究生是高校学科建设与科学研究的生力军。学科建设是人才培养的平台，为人才培养提供知识体系和环境保障；一流的学科建设可以反哺人才培养，培育精良师资队伍，优化课程体系和教学内容，营造良好育人氛围。从国内外大学或者学科排行榜来看，学生的培养质量和创新能力是评价大学和学科排名的重要指标，也是大学社会声誉的重要载体，而我国在人才培养的质量方面与主要发达国家还存在一定差距。因此，“双一流”建设必须在培养创新人才方面进行多样化探索，力争有所建树、有所突破。“双一流”建设高校应充分发挥人才培养的核心功能，加强科教融合，拓展学科育人功能，使学科成为创新型人才培养的重要基地。应该指出，如何将学科和科研资源转化为优质的教学资源，让学生直接从中受益，是“双一流”建设的难点，也是“双一流”建设成效评价需要重点关注的内容之一。

（三）科学研究与学科建设

学科建设与科学研究协同共生、相互促进、协调发展。学科建设决定了高校的科研方向与特色，为科学研究提供知识支撑、人力支持和条件保障；一流的学科建设有助于造就高水平的科研队伍，推动重大科研项目的开展和科技创新成果的产出。科学研究是知识创造与知识体系更新的重要途径，是学科建设的重要载体和支撑；一流的科学研究有助于完善学科体系、改造传统学科和孕育新兴交叉学科，有助于提高教师水平、创新人才培养模

式和丰富教学内容。当前，“双一流”建设高校的学科建设与科学研究要重视问题导向，服务于国家重大战略需求，主动回应经济社会发展和科技文化发展的现实诉求；要重视基本理论研究，在概念体系、理论体系和知识体系等方面不断完善，发挥理论指导作用；要重视创新研究方法，完善方法论体系，尤其注重将理论思辨与实证研究相结合，阐释复杂的现象，揭示隐蔽的规律；要重视学科交叉融合，促进不同类型层次、不同教育领域的“跨域”研究，促进教育与经济社会和文化、科技的“跨界”研究，促进不同学科理论与方法交叉融合的“跨学科”研究。“双一流”建设成效评价应该关注学科建设与科学研究相互支撑与促进的关系，设计科学合理的评价指标体系。

三、评价标准的多重要素考量

“双一流”建设成效的评价标准应有明确的导向性，尤其要处理好世界一流与中国特色的关系、“双一流”建设共性与学校及学科发展个性的关系，注重主观、定性评价与客观、定量评价指标的融合。

（一）中国特色与世界一流

“双一流”建设的目标是实现“中国特色”与“世界一流”，“双一流”建设的评价标准要凸显“中国特色”“世界一流”两个特征，正确认识二者之间辩证统一的关系。一方面，“双一流”建设必须坚持世界一流的国际标准，建设成效要获得国际学界和国际社会的认可：自然科学领域的学科建设要对接国际标准，跟踪和引领学科发展趋势；哲学社会科学领域的学科建设也要形成国际话语权和提升国际影响力。基于此，“双一流”建设高校在制定建设方案时应多数选择办学特色或学科领域相近的国际名校及相关学科作为对标学校及学科，在成效评价时亦应有所呼应。另一方面，“双一流”建设必须要扎根中国大地，建设好中国的世界一流大学和一流学科，坚守中国特色。立足中国本土实践，解决中国自然和社会领域的重大理论与现实问题，本身也是对世界作出的重要贡献。“双一流”建设高校要主动对接国家和区域重大发展战

略，指导经济社会发展实践，彰显办学特色和学科优势。

总之，“双一流”建设成效的评价标准及指标体系既要体现国际标准，亦要体现“双一流”建设高校在服务国家和区域重大战略需求方面的成效，将国际化指标与本土化指标有机地结合起来。

（二）共性要求与分类评价

“双一流”建设的评价标准既要参考国内外大学评价和学科评价公认的较为一致的评价标准，又要坚持尊重差异、分类评价的原则，对不同高校类型、不同学科类型区别对待。一方面，“双一流”建设评价标准应参考国内外公认的大学排行与学科排行。比如我国上海软科的世界大学学术排名（ARWU）、英国全球教育集团的世界大学学科排名（QS）、泰晤士高等教育的世界大学排名（THE）和《美国新闻与世界报道》发布的世界大学排名（U. S. News）等的评价标准及指标体系都具有较好的参考与借鉴价值。另一方面，“双一流”建设评价标准亦要充分考虑不同大学和学科类型的特点，以及这类大学和学科服务国家重大战略需求和行业领域需求所作出的突出贡献。应该指出，质量保障的重要原则之一是拒绝在单一指标或证据的基础上评价学生、项目，或者学校的表现状况。比如一些“双一流”建设高校将入选ESI前1%或者1‰作为重要的建设成效指标，但ESI的本质是索引服务，其学科分类与国内差别很大，主要侧重在理工科，而人文社科类的评价不占优势，主要以论文的被引频次为排名标准，这种单一的评价标准尤其在社会科学领域存在较大的局限性。因此，相关评价标准应尊重学科差异性，对不同学科可以通过调整指标内涵或者权重的方式进行分类评价，既要考虑不同学科体系下学科之间的差异，更要关注同一学科不同层次的差异。

（三）主观、定性评价与客观、定量评价

以往对大学和学科的评价活动更多重视客观、量化的评价标准，但是人们越来越意识到这一评价导向的局限性。量化评价以数量化的评价标准为主，便于客观结果的呈现与管理者的使用，但高

等教育丰富的价值判断难以单纯用抽象的数字来精确表达。正如有观点认为：“不是所有的大学教育的结果、不是所有发生在我们教室或演播室里的重要事情、不是所有的师生互动交流中珍贵的东西、不是所有能让学院增值的力量都可以或应该以数字和测量来加以评定的。”也正是基于以往评价理论存在的管理主义倾向、忽略价值的多元性以及过分强调调查的科学方式，古贝和林肯基于建构主义提出了第四代评估理论。第四代评估理论承认价值的多元性，认为评估不是对事物客观现象的判断，而是参与评估的利益群体沟通、协商，进而达成共同建构的过程。因此，第四代评估理论重视利益相关者的主观评价。当然，主观的评价指标如何设计和呈现、结果如何判断和使用，是需要深入思考和研究的问题，而且第四代评估理论也不可能解决所有的问题。有研究者认为，有望超越理性主义和建构主义走向复合的第五代评估。

因此，“双一流”建设的评价标准和指标体系设计要考虑定性定量评价相结合、主观与客观评价相结合。但“双一流”建设成效评价是否可以在这种复合的评价途径中趋于完善，尚有待探索；而从某种角度而言，这种探索也是对评估理论和评价体系构建的丰富与完善。

四、评价体系的多元主体协作

促进评价主体的多元化是当前大学发展评价的主要趋势之一，无论是以政府、高校还是以第三方评估机构为主导，都存在单一主体不可克服的矛盾。“双一流”建设的评价应在政府监督和指导之下，形成多主体协作的评价体系，重点是落实高校主体责任，发挥同行指导咨询作用，参考第三方评估结果，促进高等教育质量的提升。

（一）切实落实高校自主评估基础地位

高校在“双一流”建设中负有主体责任，也是“双一流”建设成效评价的主体。教育部明确了在“双一流”建设中期和末期，建设高校应根据建设方案对建设情况进行自评，并发布自评报告。当前“双一流”建设高校的中期自我评估应对照学校制定的一流大学或者一流学科建设方案，集中考察

建设效果与总体方案的符合度、建设方案主要目标的达成度、在第三方评价中的表现度、在服务国家和社会中的贡献度，不仅要看在论著发表及获奖情况，更要看在解决经济社会发展重大问题、服务国家和地方决策、指导社会实践等方面所作的贡献。因此，自然科学学科应集中于提高社会生产力，人文社会科学学科则应集中于解决社会实践问题。高校的自评报告是“双一流”建设评价的基础，无论是中期评估还是末期评估，专家评议都要以自评报告为主要依据。因此，高校应该致力于内涵式发展，注重学校整体办学水平和学科建设质量的提升，增强内部质量保障体系的有效性，提高自我评估、持续改进的能力和水平。

（二）重视发挥专家指导咨询作用

专家评价是大学评价的重要组成部分，“双一流”建设评价应重视同行专家和教育领域专家的指导和咨询作用。首先，在学校遴选与方案论证阶段，专家应审定遴选标准并基于数据和事实确定建设高校名单，集中审议“双一流”建设高校的建设方案并提出修改意见。其次，在中期检查阶段，专家应在高校自我诊断性评价的基础上进入部分建设高校进行实地考察和咨询指导，并根据中期建设成效评价指标体系对建设高校进行评价。最后，在建设末期，专家应依据教育部制定的“双一流”建设评价指标体系，在高校自评报告基础上对建设成效进行总体评价。有学者认为，我国同行评价机制存在同行评价不足、评价标准单一、评价方法趋于一致、评价社会化程度不高、行政力量干预过强等问题。“双一流”建设成效评价应重视这些问题，进行制度化的探索，比如完善同行专家与教育专家遴选机制、优化评价标准和指标体系、创新专家考察和评价方式、加强专家培训和评审监管等，以提高同行评价的针对性和有效性。

（三）合理参考第三方评估结果

“双一流”建设高校的遴选参考了国内外第三方评价内容，将其作为遴选指标体系中的指标，并赋予不同的权重。有研究者认为，这具有开创性

意义，这是“双一流”建设政策与我国以往重点高校建设政策相比的突出变化。“双一流”建设评价体系的构建，在评价主体上有赖于形成一个大学、政府、社会协同评价的体系，并重视第三方评价机构的作用。在评价标准的确立、指标体系的构建和建设成效的评价等方面，我国可以参考第三方评估机构构建的大学排名、学科排名等评价标准和得出的评价结果，但同时更应该注意的是，我国应该针对建设高校和学科的特点，赋予大学、政府、社会各类指标以不同的权重，综合评判建设成效，谨慎参考和科学借鉴国内外大学和学科排行榜的指标体系，完善中国特色的“双一流”建设评价指标体系，为我国高校的分类发展、争创一流提供科学工具和建设保障。“双一流”建设高校则应根据学校或者学科发展情况，对标不同评估机构的结果，吸取其合理元素，明晰提高建设成效的途径。

五、结语

“双一流”建设是一项新的高等教育重点建设策略，目前国内的相关研究逐年增加，但是关注点并不均衡。从初步的文献分析来看，研究学科发展、指标体系、排行榜、国际化的居多，但探讨体制改革、依法治教、人才培养、资源配置等问题的较少；在“双一流”建设宏观背景下研究高校发展问题的居多，但结合不同层次和类型高校实际深入开展院校研究的较少。教育研究者和高校管理者要进一步拓展研究视野，加强“双一流”建设与评价研究，面向问题、扎根实践，将学理分析与大学实际相结合，将批判性与建设性相结合，为世界一流大学与一流学科建设提供坚实的研究基础，推动“双一流”建设的健康可持续发展。

（原文刊载于《高校教育管理》2020年第4期）

（钟秉林，北京师范大学教育学部高等教育研究院名誉院长、教授、博导，国家教育咨询委员会委员，国务院学位委员会委员，北京100875；王新风，北京师范大学教育学部讲师、教育学博士，北京100875）

构建中国特色评价体系 推进世界一流大学建设

王战军 刘 静

“双一流”建设是党中央审时度势、慎重抉择的重大战略决策，将为支撑国家各项战略发展，促进经济转型升级，加快科技创新步伐，实现高等教育现代化，贡献重要的人力支持和智力保障。在“双一流”建设中，评价具有重要的导向作用，因此，“双一流”建设成效评价受到各方的高度关注。习近平总书记在9月10日全国教育大会中指出，要“扭转不科学的教育评价导向，坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾”。构建创新多元的世界一流大学建设成效评价体系，是适应新时代高等教育的新发展，扩大世界一流大学的建设成效，保障“双一流”建设战略顺利实施的有力举措。

一、融通中外，简约可行，创新现代高等教育评价理念

新中国成立以来，我国的高等教育经历了恢复重建期、规模扩张期、加速发展期和当前的内涵发展期四个阶段，建设重点逐步从推崇数量规模扩张，转变为聚焦结构优化与质量提升。世界一流大学的建设将对实现高等教育由外延式扩张向内涵式发展的转变起到示范引领作用，其建设评价将着眼于内涵式建设的成效。建设中国特色、世界一流的评价体系，要在参照国际质量评价标准的基础上，建立起具有中国智慧的成效评价体系。世界一流大学建设成效评价以促进我国建成高等教育强国为目标，以国家经济社会发展的重点战略和新时代高等教育的使命为导向，采取“融通中外，简约可行”的评价策略，坚持国际视角、中国特色和第三方评价的“三个视角”，强调中国特色与国际评价相结合、定性评价与定量评价相结合、动态评价与周期性评价相结合、长期建设目标达成情况与年度建设成效评价相结合的“四个结合”的评价方法，围绕“达成度、贡献度、支撑度、引领度、满意度”五个评价维度，监测我国世界一流大学的建设状态与建设成效。

（一）建设成效评价突出目标导向

“双一流”建设从“五位一体”总体布局和

“四个全面”战略布局出发，强调“面向国家重大战略需求，面向经济社会主战场，面向世界科技发展前沿”，“突出建设的质量效益、社会贡献度和国际影响力，突出学科交叉融合和协同创新，突出与产业发展、社会需求、科技前沿紧密衔接”。围绕“三面向”“三突出”的建设方向，我国世界一流大学建设旨在引导高水平大学积极发挥在国家重大战略需求、经济社会发展、世界科技发展中的贡献、支撑与引领作用，带动我国高等教育在人才培养、科学研究、社会服务、文化传承、国际交流中综合水平的总体提升，实现高等教育强国的宏伟建设目标。

世界一流大学建设成效评价将全面综合、科学客观、公平公正地评价“双一流”建设成效。实施成效评价，将对世界一流大学的建设起到导向作用。实施中国特色的世界一流大学建设成效评价，将引导高校立足中国实际，找准本校优势与特色，主动服务社会需求；实施中国特色的世界一流大学建设成效评价，将突出建设业绩，打破以往身份固化的情况，实行总量控制下的动态调整，对各建设高校具有良好的激励约束作用；实施中国特色的世界一流大学建设成效评价，将引导高校紧抓立德树人这一根本任务，突出高校建设的目标达成度、战略贡献度、行业支撑度、创新引领度和社会满意度，走出中国大学内涵式发展道路。

（二）建设成效评价贯彻两个策略

世界一流大学建设成效评价以服务战略、创新驱动、科学客观、世界一流为指导思想，全面贯彻“融通中外、简约可行”的评价策略。“融通中外”是对成效评价“中国特色、世界一流”的诠释，具体说来，是指评价建设成效的目光不能局限于国内，更要放眼全球，要求各建设高校在服务国家发展的同时，将视野扩大到促进全人类的进步上。“简约可行”是对评价操作层面的基本要求，评价体系不再追求评价指标的大而全，而是注重事关发展的关键环节和重大问题，在评价方面力求不给建设高校增添负担，引导建设高校不去刻意追逐

评价指标，不浪费人力物力财力去“拉项目”“戴帽子”“挂牌子”，倡导建设高校凝心聚力，静心搞建设，专心谋发展。

（三）建设成效评价坚持三个视角

世界一流大学建设成效评价反映了丰富多元的视角，突出了历史的眼光和全球的视野。首先，突出评价的国际视角。注重国际标准的引入与评审专家的国际化，参考多种具有影响力的国际标准，吸纳国际各领域专家评委参与一流大学建设评价。其次，凸显中国元素与中国特色。紧密围绕我国“双一流”建设的建设任务与改革任务，始终不忘服务中华民族伟大复兴的初心，在世界一流大学建设与评价中发出中国声音，书写中国方案。第三，坚持第三方评价视角。“鼓励第三方独立开展建设过程及建设成效的监测评价”，选择国内外具有影响力的多个第三方，或直接选用第三方评价中能够代表“双一流”建设目标的高水平评价维度和基础指标，或直接邀请多个第三方共同参与世界一流大学的建设评价。在这其中，要注意慎用营利性评价机构的评价结果。

（四）建设成效评价强调四个结合

世界一流大学建设成效评价利用综合评价模式，强调多种评价方法相结合。第一，注重国际评价与中国特色相结合。我国世界一流大学的建设将紧密结合国家的未来发展，衡量世界一流大学的建设成效，应着重考察高校建设发展是否准确把握我国高等教育的“四个服务面向”问题，突出“为人民服务，为中国共产党治国理政服务，为巩固和发展中国特色社会主义制度服务，为改革开放和社会主义现代化建设服务”的办学方向。建设成效评价指标的选取，要在积极总结我国历史上在重点大学建设评价、学科评估经验的同时，参考多种具有影响力的国际标准，警惕第三方评价普遍存在的数据易得、论文导向、英文期刊主导的问题，及多选用量化数据而忽略质性数据的情况，最终建立起具有中国特色的评价体系。第二，注重定量评价与定性评价相结合。定量评价指标具有强可比性和强可测性，在反映世界一流大学建设的共性问题时，能够标准、客观、精确、简便地评价建设高校情况。定性评价指标更着重过程中的表现，作为发展性评价，更能够反映世界一流大学建设的个性化特点，

便于灵活和主观能动地评价各建设高校。世界一流大学建设高校具有不同的建设起点，不能采取一刀切的定量评价方式，同时，一些评价维度难以分解为定量指标，需进行定性评价，要注意平衡两种评价方式，将二者科学地结合起来。第三，注重周期性评价与动态评价相结合。构建世界一流大学建设动态监测平台，确保评价的常态化，监控各建设高校建设举措的执行与建设目标的达成。在进行周期性评价的同时，将测量与干预相结合，对于在动态监测中出现问题的建设高校，及时进行建设指导，反馈建设问题，关注建设的未来发展与持续改进。第四，注重长期建设目标达成情况与年度建设成效评价相结合。在建设期末综合考核的基础上，加强对世界一流大学建设的年度建设成效评价，促进各建设高校将建设目标合理分解，保障总体目标的高效达成。

世界一流大学建设成效评价的意义深远，高度体现了对建设过程负责、对历史负责、对人民负责的责任意识。建设成效评价体系的设计，将有力促进“双一流”建设总体目标的实现，提升我国高等教育综合实力和国际竞争力，为实现“两个一百年”奋斗目标和中华民族伟大复兴的中国梦书写奋进之笔。

二、扎根中国大地，注重内涵式发展，构建世界一流大学建设成效评价指数

世界一流大学建设成效评价体系坚持目标导向、多元视角和多种方法的评价理念，其主体部分是世界一流大学建设成效指数。建设成效指数的构建，将有效摒弃以往评价指标体系零散、片面、繁冗的不足，直指“双一流”建设目标，具有框架层次清晰、结构分明的特点。建设成效指数各维度间相对独立，意义指向明确，将有利于突出中国品格、中国特色、中国魅力的建设导向，避免引导各建设高校建成“第二个哈佛、牛津、斯坦福、麻省理工、剑桥”。

（一）建设成效指数框架

世界一流大学建设成效评价的内容丰富多元，建设成效指数由达成度、贡献度、支撑度、引领度、满意度五个维度组成，将从五个方向综合评价世界一流大学的建设情况。

达成度考察一流大学建设方案目标的完成情

况。作为世界一流大学建设成效评价的第一个维度，衡量既定建设目标的达成情况体现了成效评价的目标导向。我国世界一流大学建设既要考虑中国实际，遵循教育规律，又要按照各省市、各建设高校的建设方案科学合理地逐步推进。达成度要分别考察各建设高校的建设成效与我国“双一流”建设总体目标对比的完成情况、与区域“双一流”建设目标对比的完成情况、与自身《建设方案》中建设目标对比的完成情况。达成度的评价一是考察各建设高校的建设成效在师资队伍建设、创新人才培养、科研水平提升、文化传承创新、推进成果转化等建设任务中的完成情况；二是考察各建设高校在坚持党的领导、内部治理改革、关键环节突破、社会参与成效、国际合作交流等改革任务中的完成情况。各建设高校要不断对标自身建设现状与国家、区域战略发展及本校建设的目标要求，侧重对核心与重点问题的突破，提高建设效率。

贡献度考察一流大学在服务社会主义现代化强国建设中的作用发挥情况。作为世界一流大学建设成效评价的第二个维度，衡量高校解决实际问题的能力和水平。实现科教兴国战略、人才强国战略、创新驱动发展战略等重大的国家战略，建设制造强国、科技强国、质量强国等社会主义现代化强国，迫切需要一流大学参与其中，各建设高校要面向国家与社会需求，“把论文写在祖国的大地上，把科技成果应用在实际现代化的伟大事业中”，提升对经济建设与国家发展的贡献率。因此，贡献度的评价一是考察建设高校在服务国家重大战略与区域经济社会发展中的贡献程度；二是考察建设高校有效服务行业企业发展的能力；三是考察各建设高校在智库建设、前沿理论、产业转型升级等方面所发挥作用的情况。各建设高校应发挥特色与优势，努力成为促进技术变革、加速创新发展的动力源，国家重大决策的智囊团，在研究战略需求、服务战略需求中做出卓绝贡献。

支撑度考察一流大学在支撑发展与创新中的作用情况。作为世界一流大学建设成效评价的第三个维度，体现了一流大学在支撑社会发展中所起到的无可替代的作用。加速我国核心竞争力和综合实力的提升，需要一流大学提供富有创新精神和实践能力的复合型人才、拔尖创新型人才和应用型人

才，开展能够解决重大实践问题的创新型研究，增强“关键核心技术创新能力”。因此，支撑度的评价一是考察各建设高校在国家重点战略领域、区域经济发展、相关行业等方面发展中的不可或缺作用；二是考察建设高校在内涵式发展中所取得的重大成果、零的突破或填补空白的能力；三是考察建设高校在大学内部治理问题与中国特色现代大学治理方面的卓越成效及其对我国现代大学制度建设的支撑作用。各建设高校要推进资源共享进度，与外界形成良好的协同合作机制，为建设社会主义现代化强国提供有力支撑。

引领度考察一流大学对学科前沿与文化思想的引领情况。作为世界一流大学建设成效评价的第四个维度，衡量对国内外相关领域的引领情况，体现了一流大学在国内外的影响与引领作用。世界一流大学建设要在学科建设、科学研究、文化引领中起到卓越的领跑作用，着力发展与国家安全和重大利益相关的学科、国家急需支撑产业转型升级和区域发展的学科，提高基础研究水平，加强战略性、全局性、前瞻性问题的研究，增强文化自信，形成推动社会进步、引领文明进程的一流大学文化。因此引领度的评价一是考察建设高校在学科建设中的引领程度；二是考察建设高校在世界科技发展中的引领程度；三是考察建设高校在促进先进思想和优秀文化传播中的引领程度。各建设高校要大力发展学科交叉融合与协同创新，力争做国际科技发展前沿中的领跑者，在服务区域发展的同时攀登学科领域高峰。

满意度考察一流大学建设成效在国内外的认可与满意情况。作为五个维度中的压轴维度，它是我国世界一流大学建设评价的最高标准，衡量国内外对一流大学建设成效的满意情况。满意度的评价一是考察各建设高校在政府、社会、公众中的满意情况，反映“双一流”建设期待与建设现实的匹配程度，展示各方对建设高校的综合表现印象。满意度的评价二是关注一流大学的国际认可与评价情况。通过有影响力的第三方评价，客观综合地评价各建设高校的国际排名情况，同时考察各建设高校对国际人才的吸引情况，观测建设高校在国际标准和规则制定中的话语权情况。

针对世界一流大学建设中的现实问题及复杂

情况，建设成效指数的设定是开放的，各维度反映的要素并非封闭不变的，随着建设情况的变化，可以进行动态增减，及时引导建设方向转移，反映建设状态变化。

需要关注的是，针对世界一流大学建设起点不同、区位优势不同、历史发展积淀不同等现实情况，要实事求是，扶优扶需扶特扶新，引导各建设高校走出自己的建设路子，实现“差别化发展”。针对东中西部经济社会发展不平衡的状况，要“优化高等教育结构，支持中西部建设有特色、高水平大学”，重点考察建设高校在优、需、特、新行业领域，以及在“一带一路”为代表的国家战略中的贡献与支撑情况，并在资金与资源分配上，实行差别对待。最终的目标在于引导各建设高校合理定位，差别发展，办出特色，努力形成支撑国家长远发展的世界一流大学体系。

（二）建设成效指数构建

建设成效指数既要体现中国特色，又要展示全新的世界一流大学建设成效的评价模式，其构建的逻辑起点是世界一流大学面向未来的时代意义和对标国际的全球质量标准，具有鲜明的目标导向与建设定位。建设成效指数的五个维度相互独立且较为全面地勾画了我国世界一流大学的建设目标。五个维度之间呈现由基本要求到最高要求的逻辑关系：“达成度”强调建设目标导向，是考核建设目标的完成程度；“贡献度”“支撑度”与“引领度”从三个角度分解了我国世界一流大学建设的“五大建设任务”和“五大改革任务”，是成效评价体系中极具中国特色的评价维度，体现了世界一

流大学的“四个服务面向”；“满意度”是世界一流大学建设的最高要求，体现了办让党放心、让人民满意的我国高等教育办学要求。

建设成效指数所需的定性评价，由覆盖政府有关部门、高校、科研机构、行业组织人员及国外相关领域人士组成的专家人员进行，定量评价则基于世界一流大学建设动态监测平台展开。平台的开发要坚持“‘互联网+’的理念，以大数据驱动为核心，充分利用现代信息技术和方法”，对一流大学建设提供全时段、全方位的监测，确保成效评价的动态化与及时性。

监测平台将实现建设状态模型的实时制作，快速、高效识别各建设高校的建设状态，对建设成效及时预警。利用“四个结合”的多元评价方法，利用动态监测平台能够测度一流大学建设成效指数结果，并根据数据结果建立雷达模型（见图1），形象、生动地反映各建设高校的建设成效状态并实现建设高校与其他高校建设成效的横向对比。在雷达模型中，模型的外缘反映了42所世界一流大学在五个维度次级指数中的最佳状态，模型从中间到边缘呈放射状，由低到高地呈现指数数据结果与建设成效状态。各高校的数据形成位置不同、形状各异的闭环，通过高校间的图形与位置对比能够表明高校的建设优势与短板。

雷达模型还能够对同一所高校建设的年度目标进行科学设定。通过对建设高校当前与以往建设成效的纵向对比情况，根据历年的建设进度，预测或预设下年度的建设成效目标，明确下一步建设的努力方向（见图2）。



图1 世界一流大学建设成效指数雷达模型图

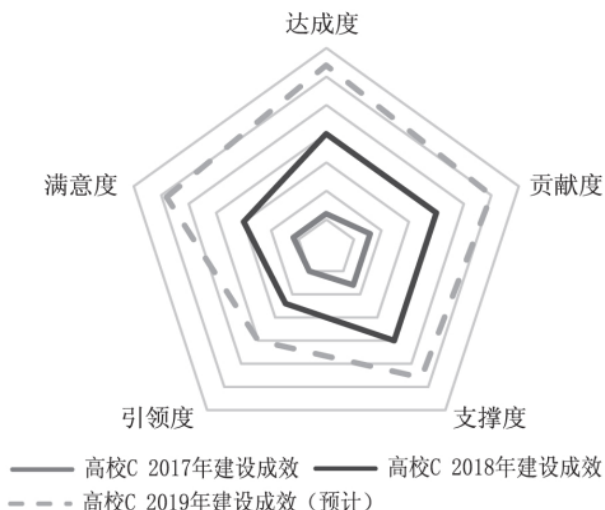


图2 世界一流大学建设成效年度目标设定模型图

三、不忘初心，牢记使命，加快建设中国特色的世界一流大学

（一）重视目标达成，提升我国高等教育的整体满意度

随着我国社会主要矛盾的变化，我国高等教育的主要矛盾也转变为人民群众对公平优质高等教育日益迫切的需求，与优质高等教育资源供给不平衡不充分的矛盾，这一矛盾涵盖了高等教育质量与公平两个方面的问题。对我国高等教育而言，“既要在促进教育公平，缩小发展差距”上“补短板”，“又要在提高教育质量，优化教育结构”上“做加法”，为广大人民提供更加多层次、多样化的优质高等教育。世界一流大学建设将在提高教育质量、优化教育结构方面起到领头雁、排头兵的引领作用，整体提升政府、社会、人民对我国高等教育的满意度。

1. 全面达成建设目标，提升国内满意度

提升国内满意度，要确保全面、按时达成建设目标。各建设高校的建设方案是我国“双一流”建设总体方案的分解，只有确保各建设高校高效达成建设目标，才能实现我国在本世界中叶建成高等教育强国的“双一流”建设总目标。各建设高校向社会公布的建设方案要点涵盖了学校的近期（2020年）、中期（2030年）、远期（21世纪中叶）建设目标与学校的学科建设总体规划、拟建设学科及其对带动学校整体建设的作用，细化的分阶段建设目标、建设内容、建设举措，便于进行自我考核评价和实施成效评价。要按照建设方案的任务要求，把控时间节点，确保建设目标完成，这是提升一流大学建设满意度的基础。提升国内满意度，要提高党和政府对一流大学建设的满意度。要切实加强党的领导，牢牢把握中国特色社会主义办学方向，践行中国特色的办学理念和办学模式。提升国内满意度，要提高社会对一流大学建设的满意度。要加强社会服务能力，不负社会期待，为社会发展提供有力的创新支撑和人才支持。提升国内满意度，要提高人民对一流大学建设的满意度。通过教学质量提升，更好回应人民群众“上好学”的期待。

提升国内满意度，还要提高广大师生对一流大学建设的满意度。要持续完善以章程为统领的高校内部治理体系改革，切实增进师生的认可度与归

属感。

2. 积极参与全球治理，提升国际认可度

高等教育的对外开放成为我国世界一流大学建设的强大助力。近年来我国高等教育对外开放的总体良好局面，提升了我国高等教育的国际话语权，使我国一流大学参与全球治理、迈向国际高等教育舞台中心成为现实。各建设高校要站在建设我国新型大国外交关系，构建人类命运共同体的战略高度上，强化全球化教育开放的顶层设计，打造国际化大学。实质性地推动高水平国际合作交流，加强与国际顶尖大学与学术机构开展各类合作，致力于成为“世界高等教育改革的参与者、推动者和引领者”。持续以开放的态度和更大的全球吸引力，进一步吸引全球优质师资与优秀生源，倡导师生以全球的视野参与国际事务，为专家学者参与国际学术组织创造平台和条件，培养学生参与国际组织的能力。值得一提的是，围绕“一带一路”开展的教育对外开放工作为我国教育迈向世界教育中心绘制了路线图，为我国高水平大学深度参与全球治理，提升国际地位和国际话语权打开了新局面，各建设高校要以此为契机，积极开展科技领域国际协同创新，增进全球竞争力，提升国际认可度。

（二）对标现实需求，增强对国家发展的贡献与支撑效用

高等教育对国家发展具有重要的贡献与支撑作用，纵观美国与日本在二战后的崛起，无不与高等教育的发展有着紧密的关联。高等教育的发展，为经济社会的发展贡献了人力资源与科技保障，高等教育质量的高低，决定了社会人力资本的质量，高等教育的结构调整，对经济社会转型升级的成败起到关键作用。世界一流大学建设将引领我国高等教育的发展，亦将成为社会经济发展的动力源泉。

1. 培养创新人才，为大国崛起提供优质人力资源

高校的立身之本在于立德树人。扎根中国大地，办中国特色的世界一流大学，首先需要回答“培养什么样的人、为谁培养人、怎样培养人”的基本问题。其次，高水平的师资是提升教学质量、培养一流人才、产出一流成果的关键，各建设高校要加强师资队伍建设，建成一支师德高、教学优、学术强的高水平师资队伍。围绕教学质量这一中

心，依靠一流的教学理念和一流的教师团队，培养学生的创新精神、实践能力和责任意识，造就高水平人才，为社会经济发展提供优质人力资源，并以人才培养带动各项建设工作开展。与此同时，培养好家国情怀这一立德树人的“灵魂”、净化好校风学风这一立德树人的“空气”，引导好大学生扣好人生的第一粒纽扣，培养出优秀的中国特色社会主义的建设者和接班人，服务民族复兴、大国崛起。

2. 回应国家需求，为社会发展现实问题的解决提供有效支撑

扎根中国大地，就要研究中国现状，解决区域问题，回应社会需求。各建设高校要将服务社会的功能与大学的人才培养和科学研究功能紧密结合起来，主动回应国家需求，高效服务社会，对接国家重大发展战略与区域经济发展现实问题。要站在解决经济发展现实问题的最前端，找准人才培养目标定位，合理布局研究力量，深化产教融合，成为驱动产业变革的马达；要加强对制约社会发展问题的研究与探索，有力发挥理论研究在辅助党和政府科学决策重大问题时的重要作用，成为服务国家决策的智库；要以国家需求和人民需要为导向，提升科研能力，逐步掌握重大关键性技术，成为实现重大颠覆性科技创新的突破者。

（三）引领未来发展，提升对人类发展的创新引领能力

党的十九大报告指出，创新是引领发展的第一动力，是建设现代化经济体系的战略支撑。我国世界一流大学要以面向未来，引领带动我国高等教育乃至整个社会的发展与变革为己任，在科技创新、文化创新、制度创新方面，不断超越，取得领跑位置。

1. 增强科技创新能力，产出一流科研成果

一流的创新能力和研究成果奠定了大学的学术声誉，在衡量一所大学的发展水平中起到重要的作用。各建设高校要积极思考国家建设、区域发展的需求，担负起新时代高校科技发展的新使命，积极开展前瞻性的基础研究和引领性的创新技术研究，加强与行业企业的密切合作，提升将知识转化为技术与生产力的能力，逐步实现科技工作从跟跑到领跑的过渡，为经济发展提速，为人类发展作出贡献。一流学科建设是一流大学建设的基础，是创

新成果孕育的平台。遵循学科发展规律，强化学科内涵建设，在优势学科高峰的支撑下，建设高水平学科群高原，形成学科高峰带动学科群发展，学科群滋养学科高峰建设的良好局面。一流大学建设高校要构建有利于科技创新的体制机制，通过科研团队和科研平台的搭建，从世界范围引进人才，打造高水平学术团队；通过积极转变科研评价机制，由过去对项目与经费数量的关注，转变为关注科研工作的实际贡献程度及育人效果；创建良好的科研环境，不断激发科研人员的创新活力，促进高水平科研成果的产出。

2. 传承创新中华文化，做人类文化传播引领者

良好的大学文化不仅是孕育人才的摇篮，也为一流成果产出提供肥沃的土壤。各建设高校要传承、创新、传播、引领中华民族悠久的传统文化，把中华文化融入大学文化的基因之中，成就具有中国风骨、中国气派的大学气质与大学性格。“坚持中国特色社会主义的道路自信、理论自信、制度自信、文化自信”，加强社会主义核心价值观教育，形成特色鲜明、引领社会文化方向的大学文化与大学精神。此外，在当今高等教育发展的复杂历史条件下，一流大学的大学文化要积极应对传统文化与国外文化的交融，在交融过程中，保持自身的文化特色，并不断创新，形成引领世界文化发展的能力。总之，各建设高校要继承全人类最先进的文化，发扬本民族传统文化的精髓，弘扬本校大学精神和校风文化，传承优秀的学术文化，为一流大学建设打好文化基础。

“山再高，往上攀，总能登顶；路再长，走下去，定能到达。”世界一流大学的建设，担负着国家的期望、民族的未来、人民的梦想。世界一流大学建设成效评价将为我国世界一流大学建设的梦想保驾护航，各建设高校要勇于追求，执着前行，砥砺奋进，为实现“双一流”建设目标贡献力量。

（原文刊载于《清华大学教育研究》2018年第12期）

（王战军，北京理工大学研究生教育研究中心主任、教授，北京 100081；刘静，北京理工大学人文与社会科学学院博士研究生，北京 100081）

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

“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评价的元评价，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有益启示与借鉴。

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”。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。”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国家战略行动。

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”。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。

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。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，世界大学排行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行榜30多个，其中THE、QS、U.S. News、ARWU最具影响力。从这4家世界大学排行榜来看，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，政府在世界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

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

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）

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%

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 世界大学排名指标体系包括：全球研究声誉

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%
		博士学生比	2.25%
		师均博士学位授予数	6%
研究	体量、收入与声誉	师均收入	2.25%
		声誉调查	18%
		师均科研收入	6%
被引频次	研究影响力	师均科研论文量	6%
国际视野	员工、学生和研究人员	归一化影响因子(FWCI)	30%
		国际国内学生比	2.5%
		国际国内教师比	2.5%
工业收入	创新-知识转移	合作者为外国人的论文比	2.5%
		师均来自工业界的科研收入	2.5%

(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)

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%

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)

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%
教师质量	获诺贝尔科学奖和菲尔兹奖的教师折合数	Award	20%
	各学科领域被引用次数最高的科学家数量	HiCi	20%
科研成果	在《Nature》和《Science》上发表论文的折合数	N&S	20%
	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%
师均表现	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。

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。

(三) 世界大学排行榜的积极作用与消极作用

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，完全背离了大学教育的使命与初心，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，这一认识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，这一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。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。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。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。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。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：世界大学排名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，许多大学排名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家世界大学排名主要存在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家世界大学排名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。

（原文刊载于《中国高教研究》2019年第9期）

（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）

“双一流”建设评价的发展性原则及其方法

孔晓明 周川

“双一流”建设高校和学科的遴选离不开评价，“双一流”建设高校和学科的年度考核也离不开评价，“双一流”建设项目期满最终的验收考核更离不开评价。如何进行“双一流”建设的评价，显然离不开“双一流”建设评价的指标，而评价指标制定的前提是评价原则，因此，确定“双一流”建设评价的原则便成为“双一流”建设评价工作中首先需要考虑与完成的最基本任务，也成为“双一流”建设工作的基础。

而在《统筹推进世界一流大学和一流学科建设总体方案》（简称《总体方案》）的基本原则中明确提出要“坚持以绩效为杠杆”，要“强化绩效”，建立健全绩效评价机制，积极采用第三方评价，提高科学性和公信力。在《统筹推进世界一流大学和一流学科建设实施办法（暂行）》（简称《暂行办法》）中提出要“动态管理”，制定科学合理的绩效评价办法，开展中期和期末评价。

随着“双一流”建设高校名单的公布，接下来的工作重心必将转向建设与评价；高校的重心工作是建设，而教育行政主管部门的重心工作则是组织第三方开展评价。因此，确定评价的原则与方法事关“双一流”建设的导向问题；特别是当前业已开展有关建设，“双一流”建设评价的原则问题就显得尤为重要，有必要先行研究与明确。

只有确定了评价的原则，才能有利于“双一流”建设高校和第三方评价机构依据评价原则确定评价标准；按照评价标准，“双一流”建设高校才能开展重点项目的建设 and 建设学科的良好发展。

一、“双一流”建设的评价现状

1. 评价的现状

目前关于“双一流”建设的评价原则与方法等政策性文件精神，主要体现在《暂行办法》相关文件中。在《暂行办法》中提出要制定科学合理的绩效评价办法，开展中期和期末评价；建设中期，建设高校根据建设方案对建设情况进行自评，发布

自评报告。专家委员会根据建设高校的建设方案和自评报告，参考有影响力的第三方评价，对建设成效进行评价，提出中期评价意见。根据中期评价结果，对实施有力、进展良好、成效明显的建设高校及建设学科，加大支持力度；对实施不力、进展缓慢、缺乏实效的建设高校及建设学科，提出警示并减小支持力度。

建设期末，建设高校根据建设方案对建设情况进行整体自评，对改革的实施情况、建设目标和任务完成情况、学科水平、资金管理使用情况等进行全面分析，发布整体自评报告。专家委员会根据建设高校的建设方案及整体自评报告，参考有影响力的第三方评价，对建设成效进行评价，提出评价意见。

2017年9月，教育部公布“双一流”建设高校名单，同时教育部、财政部、国家发展改革委有关负责人在“双一流”建设有关情况答记者问（简称“答记者问”）中提出，选用第三方评价的原则。

一是坚持中国特色，即坚持正确的政治方向、坚持服务国家需求、坚持传承中国文化，引导高校把“四个服务”作为建设的出发点和落脚点，不简单套用国外的标准排名，不把中国特色狭义理解等同于学科特色。

二是坚持综合性、多维度，利用各类评价，从各个侧面、多个维度对高校实际情况进行评价，不与单纯的人才指标、论文指标挂钩。

三是坚持定性评价与定量评价相结合，发挥综合效用。确定采用的国内国际评价为：（1）四类国内第三方评价：人才培养类，体现高校人才培养水平，涵盖本科生和研究生教育；学科水平类，体现高校学科水平；贡献奖励类，体现高校学科的社会贡献；政策导向类，综合考虑高校服务国家重大需求、重大发展战略布局以及政策连续性等因素。（2）高校认可度较高、客观性较强的国际第三方学科评价，体现各学科的实力及国际影响力。

2. 问题与不足

(1) 注重即时性现状，忽视历史性发展。无论是《总体方案》和《暂行办法》，还是在“答记者问”中，只提及“需要制定科学合理的绩效评价办法”；并且由于“双一流”大学建设项目刚刚拉开建设的序幕，针对建设的评价并未正式开始实施。因此，目前评价的问题主要体现在学校自我评价和第三方评价中。以某省份高水平大学建设评价为例，“高水平大学建设申报申请书”中就规定有自我评价。

在“高水平大学建设申报申请书”的建设目标中明确规定要求填报“到2020年学校综合办学水平在全国位次以及在学科建设、人才培养、传习团队、科学研究、社会服务、国家交流与合作等方面的预期建设成效，并提出具体的量化指标”。而这些量化的建设目标，正是未来验收评价的具体指标体系。

所以，如此填报的结果必然是建设评价时只会注重即时性的发展现状，最终必然会忽视历史性的发展过程。而第三方评价的问题则体现在各自的评价指标之中。无论国外的泰晤士、US News、ESI，还是国内的ARWU、武书连榜等，这些第三方评价的评价指标所采集的数据也是关于某一历史时间点的数据，关注的是即时性数据与成绩，无法察觉与反映出历史性的发展过程，更缺少现状与历史的对比以及增值性的发展分析。

因此，无论是从大学自我评价层面而言，还是从第三方评价的指标来看，“双一流”建设的评价在原则与方法上过分注重即时性的发展现状与目标，忽视了大学长期的历史性发展过程，而这些却又正是世界一流大学发展所必需的历史沉淀与文化积累。

由此造成的现状就是大学在“双一流”建设中忽视对现有教师的长远发展规划，只重视能带来即时性指标提升效果的短期行为，如疯狂“挖人”现象的不断上演。这种急功近利做法的直接后果就是区域发展的不平衡，动辄千万甚至上亿的引进费用也造成了高校内部资源分配的不公以及高校财政赤字的扩大。这与现阶段党中央和国务院所提出的

要解决的主要矛盾在原则上、方向上是相背离的，也是与社会主义公平性不相符合的。

(2) 注重绝对数量指标，忽视相对发展性指标。同样道理，在“双一流”建设的评价中，现有各个相关文件中所体现的问题还表现为只注重绝对数量指标，忽视了相对性的发展过程。这一问题也是主要体现在高校的自我评价和第三方评价中。以某省份的“双一流”建设为例，在“高水平大学建设申报申请书”中的预期目标和重点任务没有一项不是具体的指标和数量要求。如此规定与要求，其结局必然是承担“双一流”建设的学校和承担评价工作的第三方机构都只会注重各项指标的数量，特别是通过建设所呈现出来的绝对发展的数量，忽视了相对性的增加量和增加的幅度，更不会考虑到增加量与经费投入的比例与效益。

在第三方评价中同样如此，其评价指标所采集的数据大多也是源于各建设高校公开发布和自我填报的数据，这些数据也是仅仅关于各项指标所达到的数值和增量，亦是绝对性的发展数据，自然也就无从得到相对性的增加量、增幅和效益的提升，其结果显然也是忽视相对性的增量、增幅及效益的比较。忽视了增量和增幅，也就忽视了比较，自然也就忽视了历史性的发展。

所以，无论是从大学自我评价层面而言，还是从第三方评价的指标来看，目前“双一流”建设的评价指标体系的问题还表现为仅仅注重绝对性数量，忽视相对性的增长量和增加值。

(3) 自我评价指标体系过于繁琐，第三方评价形式单一。在《总体方案》中，规定的建设任务包括建设一流师资队伍、培养拔尖创新人才、提升科学研究水平、传承创新优秀文化、着力推进成果转化等五个方面，每个方面又包含若干要点。《暂行办法》规定，建设期末，建设高校根据建设方案对建设情况进行整体自评，对改革的实施情况、建设目标和任务完成情况、学科水平、资金管理使用情况等进行全面分析，发布整体自评报告。

因此，高校建设末期的自评报告必然是囊括师资队伍、创新人才培养、科研水平、文化传承、成果转化的方方面面，其指标体系至少包括一级指

标、二级指标、三级指标，可谓极其繁琐。国内第三方评价涵盖四个大类，既要体现高校人才培养水平，需要涵盖本科生和研究生教育；又要体现高校学科水平；还要体现高校学科的社会贡献；最后要兼顾高校服务国家重大需求与战略。这些复杂的评价要求必然造成复杂的评价指标体系。除此之外，还有国际第三方评价。

在《暂行办法》所提出的要制定科学合理的绩效评价办法，开展中期和期末评价。中期和期末评价均是建设高校根据建设方案对建设情况进行自评，发布自评报告；专家委员会根据建设高校的建设方案和自评报告，参考有影响力的第三方评价，对建设成效进行评价，提出评价意见。

由此可见，第三方评价对于“双一流”建设评价意义重大。但是事实上，目前的第三方评价，无论是国内还是国外的第三方评价，全部以大学排行榜或者学科排行榜的形式出现，只是评价指标赋分与权重的区别所造成的分值与排名的不同而已。而这种排名方法本身就是一种单一性的体现。尽管第三方评价有其好处，但是由于其采集数据的渠道和指标体系自身的缘由，其排名的信度与效度还是需要进一步探究的。正是如此，所以在“答记者问”时才提出要“参考有影响力的第三方评价”。因此，就目前的第三方评价现状来说，其关于大学和学科的排名在评价形式上是单一的。

二、发展性评价原则与方法

1. 发展性评价

1981年美国教育评价标准联合委员会对教育评价进行了综合性界定，他们认为“教育评价是对教育目标和它的优缺点与价值判断的系统调查，为教育决策提供依据的过程”[10]。国内较早开展发展性评价研究始于20世纪末，王汉澜先生在《教育评价学》（1995年）一书中指出，发展性评价是从评价的目的与功能出发对评价所做的界定，指以促进学生、教师和学校发展为目的的评价。并且还指出，发展性评价应该是多种评价方式的集合，是多样化的评价体系。发展性评价的出发点是为评价对象确定个体化的发展性目标，不断收集评价对象发展过程中的信息，并根据具体的情况做出诊断性的

评价，提出具体的、有针对性的改进意见。发展性评价更重视评价对象未来的发展，而不是刻意地为其下一个结论，确定一个等级，因而在教育评价的实践中脱颖而出，成为目前教育界最科学合理的评价方式之一。

其后，刘川等人在《发展性评价的实践与思考》一文中认为：“发展性评价就是根据一定发展性目标，运用发展性的评价技术和方法，对学生素质发展的进程进行评价解释，使学生在发展性教育评价活动中，不断地认识自我，发展自我，完善自我，使之不断积淀、发展、优化其自我素质结构，促进学生在德、智、体诸方面素质得到生动活泼和和谐的发展。”

不难发现，王汉澜先生在《教育评价学》一书中早已提出发展性评价是以促进学生、教师和学校发展为目的；但是早期的期刊论文对于发展性评价研究则仅仅是关注学生评价；其后，期刊论文的发展性评价研究才拓展到教学评价、教师评价、课程评价，最后拓展到学校评价领域。

如前所述，早期的发展性评价是以充分发挥评价对学生学习与发展的促进作用为根本出发点，以融合教学与评价为基础和核心，以教师运用评价工具不断开展行动研究和反思，从而改进其教学和课程设计中或途径，并最终促进学生、教师教学以及课程三方面共同发展的评价。这样的观点直至目前依然盛行不息，如叶爱英等也提出发展性评价倡导教育评价要以发展为本，要注重专业发展和个性发展，要相互合作和相互信任，评价也应是开放式的。

2. 评价的原则与方法

发展性评价应是从评价的目的与功能出发对评价所做的界定，是以促进学生、教师和学校发展为目的的评价。而本文所论述的发展性评价原则及方法与前述的发展性评价既有联系，又有区别。

一方面，我们所提出的“双一流”建设评价的发展性原则具体是指在世界一流大学和一流学科建设的评价中需要考虑其发展程度和增加幅度，亦即“双一流”建设高校的评价中需要更多考虑的是通过“双一流”建设和发展而取得的增长数量，而

不是通过建设与发展所达到的绝对的发展数量；同时还需要考虑并对比为了达到这些数量所进行的相关投入，包括人力和财力等相关投入，亦即人员和经费。从这一层面上说，我们在本文所提出的评价的发展性原则与发展性评价原则还是有着非常明显的差异与区别的。

另一方面，在“双一流”建设评价的发展性原则中所强调的是“双一流”高校通过建设实现的增长量，注重的是“双一流”高校通过建设而实现的各项指标的增幅及其效益；其本身也是“双一流”建设中所取得的客观现实的发展。在这一层面上说，“双一流”建设评价的发展性原则与“双一流”高校建设的发展性评价原则还是有着一定的内在联系，并且可以认为是具有高度的关联性，而非截然不同、相互对立。

由此不难发现，“双一流”高校建设评价的发展性原则相对于发展性评价而言，其范畴更广、内涵更大、层次更高。因为在“双一流”高校建设评价的发展性原则中不仅包括评价本身的方法，更需要关注与促进“双一流”建设高校自身的发展；同时还要用发展的眼光和方法来贯穿整个评价活动、评价过程与评价方法。

我们完全可以说，这种评价的发展性原则既是一种评价原则，也是一种评价方法，更是一种评价的策略。相对其他评价方法而言，更加注重的是高校自身的成长性发展，而不是绝对的发展高度和数量。同时，还要参考投入的成本量以及所取得效益。从这一层面上说，这种综合考量投入与增长量的评价方法也正符合《暂行办法》所提出的制定科学合理的绩效评价办法的要求。

具体地说，世界一流大学和一流学科建设评价的发展性原则可以具象成下列相关特性的评价方法：一是增值性，即“双一流”建设评价中需要重点考量建设中各项指标增加的数量，而不是绝对的数量。绝对的数量也不是不重要，可以用以横向比较；而相对的增加量则可以更好反映自身的发展程度。如果绝对的数量全国第一，但是增加量却不增反减，那么在“双一流”建设评价中就是不合格、不达标的指标。这就是“双一流”建设评价的发展

性原则的具体体现。

二是效益性，即“双一流”建设中，不仅需要考量增加的数量，同时还要考量成本的投入，主要是经费的投入。举例说明，A、B两所高水平大学，经费投入A大学是B大学的两倍；而产生的可以量化的成果，A大学达不到B大学的两倍，则可以认为A大学的评价得分低于B大学的评价得分。这即为“双一流”建设评价的发展性原则的效益性特征。

三是绩效性，即通过对于“双一流”建设取得优良效益高校的进一步奖励和补助，加大资助力度，促进其进一步发展，取得更大的效益和进步。绩效性既是基于效益性而生产的特性，也是有区别于效益性的特性，同时还符合《暂行办法》中所提出的制定科学合理的绩效评价办法的要求。四是定量与定性相结合的特性，即“双一流”建设评价中不仅需要定量的评价，也需要定性的评价。这也是“双一流”建设评价的发展性原则与方法所必须具备的。

三、发展性原则倚重的评价方法

1. 增值性评价

增值性评价也不是一种新的评价方法，早在20世纪80年代，美国人就提出了增值评估的概念。美国最成熟的增值评估法是由田纳西大学的统计学家威廉·桑德斯博士创造的，从20世纪90年代开始便在田纳西州使用，借助田纳西增值评估系统评估学校系统、学校、教师促进学生的学业成就增长的有效性。

而最早将增值评价应用于学校效能的评价是以“科尔曼报告”为起点，逐渐在世界范围内发展起来。英国于1992年开始在部分地区进行学校增值评价实践的试点工作，2002年在全英格兰和威尔士推行学校增值评价模式。

在此，关于世界一流大学和一流学科建设的增值性评价是指建设高校自评指标体系中需要包含有对比建设前的各项相对应的指标的增加值，也就是说，要有与建设前的对比，以显示增加值的数量和比率。

显然，这种增值性评价也是一种发展性评价，不过关注的重点既有对主体发展的关注，更有

对于被评价主体发展量的关注。因此，相比单纯的发展性评价，增值性评价显得更加的具体和全面。

2. 效益性评价

教育效益性评价最早出现于20世纪80年代，伴随着经济建设的大潮，教育经济学领域提出了教育效益评价的概念。高校效益评价，亦称之为高校投资效益评价，具体是指投入与产出的比率，即高校年度成果与年度经费投入的比率。

单位经费投入（通常以万元为单位）、年度成果数越多，说明高校效益性越高；反之，说明效益性越低。在“双一流”建设的评价中，也可以采用效益性评价法。特别是在横向比较与考核各高校的“双一流”建设绩效时，这种效益评价法显得更为实用。

否则，只看绝对的数量或是绝对的增加值，都会造成“马太效应”，由此造成事实上的不公平和评价中的欠科学性。因此，效益性不仅注重“双一流”高校建设所取得的成绩，还需要考虑到人力、物力和财力的投入。

显然，效益性评价方法正是“双一流”建设评价的发展性原则所需要和强调的具体方法，相对于粗犷式的发展评价，效益性评价显得更为科学，更符合国家经济发展方式转型和社会发展的需求。所以，效益性评价方法也属于双一流高校建设评价的发展性原则所包含的具体评价方法。

3. 绩效性评价

绩效是业绩和效率的统称，包括活动过程的效率和活动的结果两层含义。可见，绩效不仅强调成绩，同时也注重效率；不仅重视活动的过程，也还注重活动的结果。由此可见，绩效评价就是对组织或个人的活动过程及其结果进行考核和评估的过程。在绩效评价中，活动过程中重点关注的是效率，对于活动结果，注重的则是业绩；既是过程与结果的并重，也是业绩与效率的统一。在活动过程及其结果中，无论是效率，还是业绩，都离不开发展，更离不开增长。没有发展，就没有增长，更没有业绩和效率可言。

因此，绩效性评价正是评价的发展性原则的重要体现和应用。正如《暂行办法》中提出要“动

态管理，制定科学合理的绩效评价办法”，而在“双一流”建设评价的发展性原则中同样注重发展的过程评价，也注重对于建设结果的评价。因此，绩效性评价也是“双一流”建设评价的发展性原则所注重的评价方法之一。

4. 定量与定性结合的评价

正如王文澜先生所说，任何单一的评价都不可能全面衡量个体的发展水平和发展过程。同样，定量评价方法或定性评价方法，任何一种都无法完成“双一流”建设的全面评价。特别是对于当前评价中定量评价充斥整个评价指标体系的现实，就更需要定性评价方法的介入。定量评价关注的是发展变量，而定性评价则是关注发展的质量和原因分析。无论是发展的数量，还是发展的质量，都是评价的发展性原则所具有的内在属性和必须具备的评价方法。

因此，“双一流”建设评价的发展性原则必然包含定量评价和定性评价，必然是定量评价和定性评价的统一与结合。庆幸的是，定量评价与定性评价相结合的评价已经被教育主管部门所认可与重视。在“答记者问”中，有关部门负责人对于遴选国内外第三方评价的原则时就提出要“坚持定性评价与定量评价相结合，发挥综合效用”。当然，这种定量评价与定性评价在高校自评和专家评价的过程中，也需要得到贯彻落实。

四、“双一流”建设评价的发展性路径

1. 评价指标体系注重体现中国特色社会主义思想

无论是在《总体方案》中，还是在《暂行办法》中都提出“双一流”建设要高举中国特色社会主义旗帜，坚持社会主义办学方向。因此，我国高校“双一流”建设具有明显的中国特色社会主义思想的特征。既然“双一流”建设的指导思想具有鲜明的中国特色社会主义特点，那么“双一流”建设的评价指标体系自然不能缺少中国特色社会主义的特点。只有具有中国特色社会主义特性的“双一流”评价指标体系，才能建设具有中国特色社会主义特色的“双一流”高校。

在当前比较认可的ESI学科排名指标中，22个

学科中只有一个是社会科学，在此排名中，我国人文社会科学最强的北京大学位列179名，复旦大学排名332，南京大学位列600名，其他国内高校的人文社科排名就不容乐观了。这与习近平提出的“四个自信”理念也是不相符合的。其结果一方面导致不少“双一流”高校急功近利，大量裁撤社会科学门类，专注于所谓的ESI学科发展，努力提升学科排名。

姑且不论跨学科、综合发展对于大学、对于知识生产的重要性，就是从学生全面成才的角度来说也是不应该的行为。大量裁撤人文社会学科的结果也会忽视社会主义特性的研究，特别是忽视“双一流”建设中的中国特色社会主义特性。因此，有必要在“双一流”建设评价指标体系中融入中国特色社会主义的指标内容。

2. 评价主体上多元主体、多方参与

“从教育的内涵看，任何单一的评价都不可能全面衡量个体的发展水平和发展过程。发展性评价应该是多种评价方式的集合，是多样化的评价体系。”因此，对“双一流”建设这一庞大工程的评价也应该采取多样化的评价。多样化的评价方法，必然要求多元主体的参与，特别是被评价主体的参与。陈玉琨教授也指出，教育评价离不开对于教育价值的理解，而教育的价值与教育的需求密切相关，教育需求决定了教育价值的具体内涵；教育的需求主要分为个人的教育需求和社会的教育需求，因此，教育评价的主体必然是包括个人和社会各方在内的多元主体。

同理，对“双一流”建设这一复杂的高等教育项目评价也必然要求多元主体的参与，特别是被评价主体的参与。学生，作为“双一流”建设高校教育质量考查的主体，理应直接参与“双一流”的评价活动；同理，与学生直接相关的家长、用人单位，也应该被纳入“双一流”建设的评价之中。只有这样，“双一流”建设评价的出发点才能真正做到根据评价对象确定个性化的发展性目标，不断收集评价对象发展过程中的信息，并根据具体的情况作出诊断性的评价，提出具体的、有针对性的改进意见。

3. 评价方法上注重形成性评价与总结性评价相结合

早期的教育注重总结性评价，而现代教育重视评价的形成性功能，即形成性评价。总结性评价和形成性评价的评价目的和期望不同，总结性评价是外部导向的，是为了区分被评价对象的等级；而形成性评价更多关注的是发展的程度。这一点也是与“双一流”建设评价的发展性原则所强调的发展程度、增加幅度不谋而合。

尽管总结性评价与形成性评价的评价目的和导向不同，关注的重点有所不同；但是两者的关系是辩证统一、密不可分的，总结性评价必须以形成性评价为基础，形成性评价的提升和总结必然离不开总结性评价，总结性评价还是形成性评价社会责任的体现。所以，形成性评价和总结性评价的结合在“双一流”建设的评价中显得尤为重要和必要。

4. 评价内涵上注重内部质量文化建设

“双一流”建设评价的发展性原则，其根本目的是为了更科学、更合理地实施“双一流”建设的评价活动，从而在下一轮“双一流”建设的遴选和建设过程中，促进“双一流”建设高校取得更大的进步，早日实现建设成为世界一流大学和一流学科的目标。高校建设世界一流大学和一流学科的过程，也是高校不断提高自身教育质量的过程。

因此，高校注重质量发展与开展“双一流”建设在本质上是一致的。有研究指出，人才培养和评价是高校内部质量保障的重要环节，也是实现“双一流”建设高校核心使命的重要组成部分。

所以，“双一流”建设的评价离不开高校的内部质量建设；而高校的内部质量建设从高校自身来说，则必须形成内部质量文化，才能自觉开展质量建设，从而提升“双一流”建设高校的内部质量，早日实现建设成世界一流大学和一流学科的宏伟目标。

（原文刊载于《江苏高教》2019年第12期）

（孔晓明，苏州大学教育科学研究院博士生，江苏苏州 215123；周川，苏州大学教育科学研究院院长、教授、博士生导师，江苏苏州 215123）

“双一流”建设背景下高校评价体系之重构

胡仁东

一、背景与现实

2010年,《国家中长期教育改革和发展规划纲要(2010—2020年)》提出开展特色评价、绩效评价和独立评价的要求,即克服高校同质化倾向,“形成各自的办学理念和风格,在不同层次、不同领域办出特色”“改进管理模式,引入竞争机制,实行绩效评估,进行动态管理”“鼓励专门机构和社会中介机构对高等学校学科、专业、课程等水平和质量进行评估”。特色评价、绩效评价、独立评价的三维度导向为高校评价体系改革提供了基本思路。

2015年,国务院发布的《统筹推进世界一流大学和一流学科建设总体方案》强调构建完善中国特色的世界一流大学和一流学科评价体系来激发高校办学动力与活力。2017年,教育部等部门印发《统筹推进世界一流大学和一流学科建设实施办法(暂行)》,确定了“双一流”建设高校的认定标准,既要坚持中国特色,又要借鉴国际评价经验。同年,中办、国办印发《关于深化教育体制机制改革的意见》,提出建立“立体式”教育质量监测评估体系,健全第三方评价机制。政府相关政策文件的出台为实践者结合实际设计高校评价体系奠定了重要基础。

“双一流”建设背景下,高校评价体系的建立与完善处于政府实施“管办评”分离的制度要求、社会各界对于高等教育质量的利益诉求、高校自身发展的价值追求“三合一”的现实环境之中。探索高等教育利益相关者利益诉求的契合点、政府提高高等教育质量的发力点、高校追求自身发展的着力点,是高校评价体系构建的根本目的。这在实质上要解决的是高校人才培养与社会需求之间的契合度、科学研究在世界范围内的领先度、社会服务在社会发展中的贡献度、文化传承创新的影响度等关键性问题。在高质量教育发展的新时代要求之下,重新评估我国高校评价

体系存在的问题并提出新的构建思路具有重要的现实意义。

二、基于文献的问题审视

高校评价体系的^{价值何在? 2006年第2届大学排名国际专家组会议提出“高等教育机构排名的柏林原则”: 排名既为不同利益相关者提供选择需要,也可为政府、认证和评价机构有关评估工作提供补充。评价活动首先是源于社会公众对于大学了解的需要,源于社会监督的需要,进而成为教育管理部门的管理依据。万冬根认为,高校评价体系的^{意义在于为高考考生提供了解高校实力水平的的依据、为高校发展提供一个标尺。这些判断和认识,总体上反映出高校评价体系满足“公共利益需求”的价值取向,但我国高校评价体系中存在着评价主体单一、评价标准缺乏弹性以及评价指标的重学术倾向等问题;教书育人的评价突出得不够。综合相关研究文献,我国高校评价体系的主要问题表现为客体被主体忽略、目的被方法代替、内涵被外延置换等。}}

1. 客体被主体忽略

评价首先是一种价值判断,并在此基础上构建人的价值观。在评价活动中,评价者与被评价者之间的互动、沟通与交流应当被置于重要地位。评价的主客体之间的价值认同是评价活动的前提和基础。对于高等教育而言,评价活动中主客体之间互动的缺失主要表现为以下两点。

首先,评价样式无法反映人的价值诉求。有研究者认为,评价样式之不足主要体现为:一是对高等教育本质和规律认识与把握不足;二是评价标准无法得到各利益相关者的认同。更为重要的是,学生作为重要主体,应参加到评价过程中,“但实际操作中大部分却被隔离在评价体系之外”。在现实中,高校教学督导对教师的评价具有硬约束作用(许多高校做出了“一票否决”的规定),而学生对教师的评价却因受到多方面

的制约并未产生多大影响。

其次，强行政评价难以彰显高校个性。在我国传统教育管理体制中，评价成为教育行政机构实施管理控制的工具。因此，出现了行政主导的一体评审、评价标准缺乏分类指导、突击应付而缺乏长效机制等问题。如果高校仅仅迎合行政主管部门的评估，势必导致对高等教育发展多样化要求的忽略，还会抑制高校的个性发展。典型的如本科教学工作水平评估，尽管有所改进，但“评价标准化”依然存在，对高校特色的反映并不充分。

2. 目的被方法代替

方法具有工具性特点，它以实现目的为基本准则；为了实现某种目的，工具的选择以适用、有效为前提。张红伟等认为，我国高等教育评估存在“工具理性驱逐价值理性”“鉴定功能排斥其他功能”“评估手段置换评估目的”三大陷阱，需要“重视高等教育评估的特殊性”“正视不同评估的效用”“回归评估是手段而不是目的”。方法与目的的倒置主要表现为鉴定性评价的限制、量化评价方法的限度和形成性评价的局限等方面。

一是鉴定性评价的限制。有研究者认为，我国高校评价传统上是以测量、描述为主的鉴定性评价，它通过不同的观测点或指标的确认与判断，抽离了高校所依存的社会情境。从具体情境中提取出来的指标，经过评价者的现场考察并用一种“规范”语言进行描述形成的判断，本身已经远离事实，因为任何评价者都处于“信息不完备”之中，也就难逃西蒙的“有限理性”之眼，因此，其鉴定结果无法达至评价目的。

二是量化评价的限度。量化评价方法是以结果为导向的评价，即按照事先预设的教育计划、方案展开活动，事后对其最终结果进行评价，以评估其总体效益，并做出“优劣”之分、价值“高低”之判断。这种评价方法无疑方便管理者的管理，但高等教育的丰富价值难以用抽象的数字来精确表达。所以，量化评价只能在局部范围内使用，如果将其用于高等教育的整体性评价，必将误导高等教育活动本身。

三是形成性评价的局限。形成性评价的目的是增进高等教育价值，主要通过教育活动相关信息的搜集、整理与分析来发现问题并提出建议，调整事先预定的计划、方案或项目等。在形成性评价中，对评价对象打分或进行优劣分等排队并非必须，有时甚至可能是背离评价目的的。因此，人们更加重视表现性评价，因为它更能反映实践的真实状态，并以此弥补形成性评价运用不当造成的损失。

鉴定性评价、量化评价和形成性评价是人们在实践中探索形成的评价方法，它们对于改进高等教育活动、增进高等教育质量具有一定的作用，并在实践中得到广泛应用。但它们均有自己的使用范围和条件，如果不准确把握这些评价方法的优势与不足，方法与目的的倒置将会使大学组织等相关主体的行为选择发生偏离。人们对于高校评价的质疑，源于管理者、评价者的“排队评价”“数值评价”和“优劣评价”之外在目的，由此，高校会通过夸大事实、包装结果、游说政府、与评价方合谋等方式争取“好的评价结果”，这与促进高等教育高质量发展的评价初心相去甚远。

3. 内涵被外延置换

对于大学的评价，充分发挥大学履行组织职能并彰显其理性价值、文化个性显然是其重要功能。但看似合理的量化评价，事实上“不符合完整的大学理性和中庸之道”。大学组织在科学主义评价的裹挟下，其文化存在被遮蔽、象征价值被悬置的问题。

一是大学组织的文化存在被遮蔽。在“重科研轻教学的大学分类定位问题、评价的逻辑起点及资料来源的科学性问题、指标体系及权重设定仅能反映数量而难以反映质量”的影响下，大学热衷于排名、升格、专业扩张、硕博点增设等，误解科学评价的内涵，引发人们对大学组织文化缺失的担忧。关于大学发展合理性、有效性和责任性的评价，有研究者认为，好的大学评价既要符合大学未来发展的需要，更应尊重不同文化背景中的大学历史与内涵，充分彰显大学组织的文化存在之属性。作为一种文化存在的大学组织，

文化元素是其内生“基因”。政府、社会与家庭对大学的重视和崇敬，是因为大学组织具有社会精神灯塔指引的重要价值。

二是大学组织的象征价值被悬置。大学组织的象征价值往往以一种抽象的符号资本来体现，根据布尔迪厄的认识，符号资本可以被理解为得到社会认可的，能生产、再生产和长期积累的荣誉、声名、精神、特殊性或神圣性等以符号化方式存在的稀缺性资源。对于高校而言，符号资本具有象征性，它能提升一所高校的认可度和可信度。在布尔迪厄的四大资本（经济资本、社会资本、文化资本和符号资本）中，最具转换能力的资本就是符号资本。在以追求数字优势为目标的评价过程中，人们忽略了“可感受、可理解、可欣赏”等维度，使得大学所累积的声誉、历史、文化和精神等符号资本被悬置。符号资本在高校评价中的缺位会给人们的判断带来偏差。求学者选择一所高校就读，更多倾心于其象征价值。

从上述问题的检视来看，我国高校评价体系缺陷之根源在于以下几点。第一，评价范式落后。高校评价体系涉及理论基础、基本要素、价值取向、现实背景等诸方面，一个有效的评价体系应是在全景式、多维度综合考量基础上，形成事实依据、价值判断与工具选择的有机整体。因此，转换评价范式是我国高校评价体系重建之要。第二，理论准备不足。高校评价体系相关理论源于实践又高于实践，其价值和意义在于理性指导实践，它要解决“谁评价、评价什么、如何评价”的合理性、合目的性和合规律性问题；在多元主体、多种类型、多维价值诉求的背景下，理论体系的建立健全是我国高校评价体系重建之基。第三，现实观照欠缺。高校评价体系落实在操作层面，关键是在价值判断和事实判断的基础上，形成不同利益相关者满意的评价结果，从而为他们做出合理选择提供指导，以追求高等教育活动的增值效应。我国高校评价中的“计件”量化指标“一刀切”明显难以显示其客观性。因此，观照现实是我国高校评价体系重建之需。

基于上述分析和高校组织价值与功能的特殊性，我们认为，高校评价体系是建立在公共价值

理念之上，以增值论、利益相关者理论、博弈论、治理理论等为基础，制定评价规则、建立评价指标、规范评价流程，“管办评”各主体有效参与对高校组织的教学、科研、社会服务、文化传承创新等进行历时与共时、价值与事实、定性与定量的综合考量，促进高校内涵式发展的有机整体。

三、高校评价体系重构设想

高校评价体系如何重构？见仁见智。在评价理念上，阎光才认为，“双一流”建设成效在评价理念上至少应坚持服务国家实效与贡献。黄宝印等人提出要构建学科评价体系的中國标准、建立评估方法的中國模式和评估结果发布的中國方式。在评价指标上，史静寰认为可通过设计“基于高等教育综合功能的多维、多层评价指标体系”，并从院校“人才培养指数”“学术生产指数”“社会贡献指数”“院校发展指数”等维度改进高校评价体系。罗燕提出“实现高校分类评价”“提炼关键性综合指标”“在评价模式上逐步实现鉴定性评价与建构性评价模式的融合”的“双一流”高校评价改进思路。在评价方法上，黄岚认为，教育评价的价值取向趋于相互借鉴以不断完善评价路径；投入、产出、效益相结合的综合评估思想逐渐显现；以关键指标评价大学发展特色，应结合教育诉求和文化背景进行。

基于上述认识，我们认为，“双一流”建设背景下我国高校评价体系构建应在宏观的范式转换、中观的理论重构、微观的实践关注三个层面着力。

1. 范式转换

一是建立基于公共价值理论的评价视角。首先，以公共价值治理理论作为指导，将公共价值的实现作为最高目标。政府无论采取何种形式获取（举办、购买等）高等教育，都应当在效率与公平之间寻求平衡点，而不是以效率作为唯一准则和终极目标。政府作为公共价值实现的代理人，伍德罗·威尔逊(Woodrow Wilson)从公共行政学的视角认为，我们应从政府是否通过适当和成功的方式并以尽可能高的效率和尽可能少的成本来履行其代理人职责，是评价其工作质量的

依据。“适当与成功”“效率与成本”本身包含公平与效率兼顾的价值取向，政府的首要职责是保证公共价值的实现。其次，将新公共管理的治理理念、制度、机制融入高校评价体系之中。管理范式转向治理范式，更加注重多元性、开放性、回应性，在“管办评”三方主体之间建立起对话沟通的平等互动机制，使“评”成为管理主体和办学主体各自规范行为的基本依据。

二是确立“管办评”各主体的角色定位。第一，政府应充当公共价值保护者。公共治理蕴含公共利益至上的基本理念，政府应以价值理性与工具理性、公平政府与效益效率的平衡为治理基础，理顺体制机制，确保公共利益最大化。第二，高校应充分发挥其公益性组织的功能。高校应以立德树人、追求学术自由、建好社会精神灯塔为价值旨归，办有特色、有品位的一流大学、一流学科。第三，评价机构应助推公共价值的实现。作为助力“双一流”建设的第三方评价者，应定位于非营利机构而不是营利机构，其使命在于根据政府政策导向，引导、监督高校推进“双一流”建设。

三是形成基于公共利益的评价方法。第一，政府弱化对高校评价方式的干预将会降低高校对政府的依赖，阻止高校通过“寻租”的方式获得资源，同时也将促使高校利用评价结果走特色发展之路。第二，高校本身提供的是一种公共服务，“合身”的评价方法不仅能激发高校办学的积极性，也会使教育的公益性得到充分体现。第三，评价机构作为非营利组织，在评价方法设计上须坚持遵守公共价值，避开与被评价者“合谋”所造成“评价失灵”，回归评价的本质。第四，其他利益相关者参与高校评价具有离散性，而且倾向于认可“权威”发布的信息，尽管事实上存在信息不对称现象，但也应理性判断各种评价方法和结果，履行其维护公共利益之责。

四是培育公共精神的评价导向机制。首先，现代公共治理范式重在考察公共管理者行为与公共治理目标的一致性，但现代社会信息的共享性导致公私领域之间界限的模糊，加剧了公共行政人员的公共角色与其公民角色的冲突，与此相关

的潜在利益和责任容易干扰客观决策。所以，应当引导公共行政人员公共角色与公民角色的分离。其次，公共生活中的公共治理等同于政府行政的看法往往会限制社会公众参与公共事务的机会；在公共精神的引领下，公共领域中的人更能发挥批判和监督功能，进而倡导公利而非私利。基于上述两点认识，高校评价体系的构建需要分离相关公共管理的“公”与“私”的角色，并倡导公利优先。

宏观层面范式转换的目的在于通过评价视角、“管办评”主体角色、评价方法、评价导向的重新确定，改变传统高校评价体系中通过评价来“排队”、显示“数值”长处、判断“优劣”等工具性价值取向，确立高校评价的公共价值理念，形成公共利益优先的评价共识。政府应在范式转换中起主要作用；其他各相关利益方也需要转变观念，将群体利益、个体利益设定在合理范围内。

2. 理论重构

一是基本概念的新时代阐释。高校评价体系的基本概念包括高校评价本质、理念、价值与标准等。高校评价本质是高校评价体系重构的逻辑起点，我们需要进一步厘清其基本要素及要素之间的内在逻辑关系。基于治理视角的价值协商观，以互动、交流与沟通为基本方式的高校评价理念的建立有助于解决高校评价涉及的相关利益群体的多元诉求。高校本身是具有公共价值的机构，以“公共利益”为根本的高校评价价值取向是重新厘清评价中各种关系的前提。“双一流”建设须有标准，但标准容易陷入用一把尺子测量不同目标定位高校的矛盾之中，所以，如何使不同机构和群体共同参与标准制定十分重要。

二是理论分析框架的重新建立。“双一流”背景下高校“管办评”分离制度的根本目的是办人民满意的高等教育，但就目前现状来看，分离机制并未彰显其优势。在“管的主体”上，政府各部门管理不到位或管控越位时有发生；在“办的主体”上，办学自主权没有真正落实到位；在“评的主体”上，第三方评价者的目的、动机明显受市场利益驱动，在一定程度上其评价的结果

不能完全令人满意。“管办评”三方处于隔离状态，难以形成有效的合作机制。从历史与现实的角度，我们认为，以权力博弈论、公共价值治理理论、利益相关者理论、场域论为基础，构建“双一流”背景下高校评价体系的理论分析框架将有助于对评价本身的客观认识。

三是生成机理的历时共时判断。由于高校不具备办学资源的自身再生能力，资源依附性特点使其在很大程度上会通过“合理”的方式获取资源。一般来说，办学中竞争性资源主要通过评价来获取。所以，无论外部评价主体是谁，或者无论是哪类评价，高校都将会以评价指标和评价结果作为办学行为选择的重要参考。从生成与发展的角度看，高校评价体系有一个从无到有、从局部到整体、从简单到复杂的过程，我们应从长的历史镜头和现实视域中发现其生成机理，以便于引导当下高校评价体系的公共价值取向并预测其未来变化趋势。

中观层面理论重构的价值在于通过理性反思高校评价体系的既有问题，同时吸收中外评价先进经验，确认评价活动的理论基础，建立基于利益相关者利益平衡的价值协商评价观、高校增值的发展评价观。这就需要政府合理引导、高校认同、社会评价价值观的转变，从而形成和谐共生的高校评价生态环境。

3. 实践关注

一是体系样态的梳理。高校评价本质上是一种管理手段，各国都有自己的高校评价体系，虽然其成熟度、认同度、可行性等方面均存在差异，但也有值得借鉴之处。首先，从世界范围内来看，我们需要对行之有效的高校评价体系予以系统考察，运用比较研究法、历史分析法、个案分析法予以总结和提炼，发现高校评价体系特点，从而为构建具有促进“双一流”建设的中国特色高校评价体系提供参考与借鉴。其次，由于高等教育发展历史的不同、社会制度与环境的差异，世界各国高校评价体系各有千秋。这就需要从现实的维度考察发达国家高校评价体系的典型样态，把握其结构组成、制度环境、实施条件等，这有助于我国高校评价体系的健全与完善。

二是体系图谱的勾勒。评价体系改革的目的在于使“双一流”建设有成效，但由于第三方评价者本身就是高等教育发展到一定阶段的衍生物，可以说它是市场条件下生成的一类机构，难免会发生作为办学者的与第三方评价者共谋从而使双方获利的行为。我们应根据高校评价的特点，借鉴不同国家或地区高校评价体系，细化高校评价体系的划分维度及其内容，并根据不同高校评价体系样态形成“高校评价体系图谱”，便于我们清晰认识到在什么条件下选择什么样的高校评价体系，防止评价被异化。

三是建构与运用本土模型。高校评价体系模型的本土建构着力点在于高校评价体系相关的权力主体之间的博弈、公共价值治理主体的行为选择、相关利益群体的利益诉求。首先，探索模型变量。评价体系应以理论基础、基本要素、现实背景等为基础，并由此形成合理有效的框架，便于从宏观上把握高校评价的各个方面，从微观上引导高校正确认识评价。其次，增强制度供给能力。从政府、高校两个方面完善评价体系的制度设计，厘清各自的权责利，促使评价方服务于政府对高等教育的决策与监管、高校发展。最后，提高运行成效。评价方通过科学合理的指标体系设计，形成值得采信的评价结果，这里涉及评价方法与技术、评价工具与策略。

微观层面实践关注的现实意义在于以多样性、丰富性、复杂性的具体实践为出发点，从长时段的历史变迁和现时段的实践样态、价值判断和事实判断、客观的定性描述与科学的定量计算予以综合考虑，建立适合我国高校“双一流”建设目标的高校评价体系图谱；坚持对被评价者的压力与能力平衡、激励与问责并行原则。政府一方面应加强顶层设计、完善相关规则体系和评价流程；另一方面应倡导局部实践探索并将成功经验上升到制度层面，在方法和技术上形成中国特色的高校评价体系。

（原文刊载于《国家教育行政学院学报》2019年第5期）

（胡仁东，江苏师范大学教育科学学院教授、博导，江苏徐州 221116）