



江苏理工学院
JIANGSU UNIVERSITY OF TECHNOLOGY

高教视点

2020年第4期

教育评价改革

2020年12月

目 录

校园之声

“新职师”带领职教师资建设走入新阶段.....	1
加大力度建设高水平技术师范大学.....	4
解决校企合作难题要回答好几道题.....	6

政策文件

中共中央 国务院印发《深化新时代教育评价改革总体方案》.....	8
----------------------------------	---

高层传音

以评价改革牵引教育领域综合改革.....	13
吹响新时代教育评价改革的号角.....	15

名家视角

深化新时代教育评价改革研究（笔谈）.....	17
新时代深化教育评价改革的根本遵循.....	24
建设中国特色世界一流大学的重要抓手.....	26
教育评价改革，关键在落实.....	28

媒体报道

深化改革，立好教育评价指挥棒.....	30
打好教育领域综合改革“关键一役”.....	31
构建合实际、高水平的教育评价体系.....	33
指引新时代教育评价改革的纲领性文件.....	34

他山之石

分类评价：指标设计、操作程式和结果应用——以上海高校分类评价为例.....	35
遵循与超越：基于《柏林原则》的高校分类评价透视——以上海为例.....	41
以首善标准建构首都高质量教育体系.....	49
江苏：积极参与学科评估，努力推动内涵发展.....	51
浙江：评价改革从区域层面破题.....	53

研究动态

教育评价：应然性与实然性的博弈及超越.....	55
教育科研评价如何走出困局.....	62
高等教育评价中的价值冲突与融合.....	72
创新评估范式 引导大学评价.....	85

【编者按】思想是行动的指导，要实行教育的改革和发展，首先要求教育思想和观念的转变。高等教育研究不仅为高校提供教育教学改革的理论支持、实验探索和咨询，也为地方政府教育发展提供咨询服务。高校必须依靠高等教育科学研究，实现教育模式的建设和加速教育教学改革的进程，进而实现学校特色发展。我校部分教师、党政干部在高等教育研究领域默默耕耘，在党报、党刊、高水平学术期刊等媒体平台发表了高质量教育评论和高水平论文。本期特摘在《中国教育报》刊登的教育评论，以期反映我校教师高等教育研究成果。后期将继续选登其他评论及论文，敬请关注。

“新职师”带领职教师资建设走入新阶段

朱林生

应新时代、新职教的新要求，职教教师队伍建设步入新阶段，应有新作为。为此，我们提出推进新时代职教教师教育改革，建设“新职师”。新职师是职教教师对职业教育“类型观”的主动回应，是探索中国特色、职教特点的教师教育目标、理念、方法与路径的一次尝试，也是职业教育配合新工科、新商科等新的人才培养要求，从而确保职教教师教育与其他类型教育同步发展、同步提高而作出的主动变革。

新职师之一——新时代的职教教师

与传统职教教师不同，新时代的职教教师应该具有以下特征：

新时代职教教师是服务型教师。新时代职教教师需要主动服务社会与产业，直接把自己的才智贡献给社区与企业。新时代，职教教师将成为推动社会与产业进步的有生力量，而不再是手插口袋的旁观者。教师主动直接的服务经历，将为培养合格的职业人提供有益的经验 and 环境。

新时代职教教师是跨界型教师。新时代职教教师为了高质量教学，将来往穿梭于课堂、实训中心、企业、社区，成为多种职业身份和角色的“复合体”。在新时代，双师型教师、产教融合型教师的概念将逐渐失去市场，因为这些都将作为职教教师必备的内涵。

新时代职教教师是创新型教师。新时代职教教师将更多地创造性丰富课程内容、创新性参与生产工艺研究与工作流程研究、引导学生开展创造创新活动。在新时代，由于全社会对创新的巨大需求，职教教师的教育教学和社会服务活动将重新充满创新性、艺术性，而不仅仅是单纯教学技术的再现。

新时代职教教师是国际化教师。新时代职教教师将主动回应“坚定文化自信”“走进世界舞台中央”的号召，以国际视野重新审视学生和日常教学，推动中国职业教育在世界职教体系中的重新定位。

新职师之一——新时代的职教教师教育

目前，职教教师教育存在以下一些问题：

一是职前教育与在职培训存在脱节。职教教师教育由职前的职业技术师范教育和在职的教师培训组成，但由于大多数职教教师来自非职业技术师范院校，职前的职业技术师范教育规模非常有限，造成了大多数教师在入职前没有接受过严格的师范教育，师范素养成为许多教师的职业生涯缺项。另外，由于本科与职业教育专业目录不同，许多教师求学阶段的专业

与毕业后从教的专业存在较大差异，专业能力存在弱项。

二是岗位成长与在职培训存在脱节。教师的岗位成长是“全纳型”要求，而在在职培训一般都以集中学习的方式按专题进行，是“点位型”要求，导致在职培训难以对教师的岗位成长形成针对性帮助，使在职培训出现了“听起来很有道理，做起来帮助不大”的尴尬。

三是师范再教育是职教教师教育的空白。对于非师范专业的职教教师而言，通过在职学习习得一些教育学知识和师范规范，但这些碎片化的知识和规范难以有效促进教育智慧的形成，因此需要在教师成长的不同阶段对他们进行师范再教育，这在我国几乎是一个空白。

新时代的职教教师教育应该是职前师范教育、岗位教师成长、在职教师培训、在职师范再教育的“四位一体”循环教育活动，职教教师就是在一次次的循环中迭代成长。在这个四位一体结构中，存在两个显性和一个隐性关键点，其中职业技术师范体系的构建与分布式在职培训机制是两个显性关键点，前者决定了职教教师教育的科学性，后者决定了职教教师教育的实效性；产教融合是隐性关键点，在每个循环的所有环节，产教融合都在发生作用，不同阶段的循环有不同形式、不同层次的产教融合。

新职师之一——新时代的职业技术师范教育

自上世纪 80 年代成为一种独立的师范教育形态起，职业技术师范教育一直在各种质疑、踌躇中前行，规模一度萎缩，功能一度消散，这种状态与新时代的要求明显相悖。在新时代，职业技术师范教育将成为融合教育的新范式。

职业技术师范教育是责任交融的共同体——以职业技术师范院校为中心节点，政府部门、其他师资培训机构、中高职院校、企业等为平等培养培训主体。

职业技术师范教育是阶段、层次融通的新培养体系——职前职后贯通，本硕层次教育一体化设计、分段培养。

职业技术师范教育是质速融合的新培养过程——加速职教教师专业成长，提升职教教师教学、实践、研究等全面能力。

新职师之一——新时代的职业技术师范院校

在新时代，职业技术师范院校（含综合性大学中附设的职业技术师范院系）将以协同、融合为基本理念推动职业技术师范教育事业的发展。

新时代职业技术师范院校将有新的培养规格与学制模式。职业技术师范院校的培养规格将是毕业生经过三年岗位成长与在职培训全面达到新时代优秀青年职教教师的标准。为达到这一目标，职业技术师范院校将采取“4+2+3”的学制模式，即：四年制本科职教师范生具备独立开展理论教学与实践教学能力，能够在专业建设与课程开发中承担重要任务；两年制教育硕士具备高水平开展理实一体教学能力、一定的专业管理与课程管理能力和一定的技术研发能力；经过三年岗位成长与在职培训毕业生成为全面成熟的优秀青年教师。

新时代职业技术师范院校将有新的办学模式与教学模式。新的合作模式包括校地协同模式——职业技术师范院校与特定的行政区域合作，共同探索职教教师教育整体改革的方式方法；校校企联盟模式，职业技术师范院校与中高职院校、长期合作的企业或产业园区共同开展“三三三”式（三个导师、三个阶段、三个场所）师资培养；泛专业培养模式，突破目前职业技术师范教育专业划分与中高职教育脱节的瓶颈，开展师范教育+专业教育+跨专业必选三结合式培养，力图培养由教育学知识、厚基础专业知识和广博跨专业知识组成的职技师

范生知识结构体系和跨专业解决问题的能力。

新时代职业技术师范院校将创造全新的证书教育模式。职业技术师范院校将成为职业技术师范教育课程的开发者与输出者，鼓励其他本科高校加盟证书教育，使更多的院校、更多的学生参与到“1+X”证书教育中来，指数级放大职业技术师范教育资源，又好又快又多地培养职教师范生，满足不断增长的职教教师需求。

新时代职业技术师范院校将有新的质量标准。将以培养具有高尚师德、国际视野、产教融合服务能力、智能化时代专业知识体系、紧跟科技与产业发展的实践能力、帮助学生终身成长的教育教学能力的新职教教师为目标，建设有中国职教特色的职教教师培养质量标准体系和质量保障体系，全面更新教育教学评价机制。

新职师的四个方面是多维共生的，其中作为职教教师的新职师是基础，作为职教教师教育的新职师是核心，作为职技师范教育和职技师范院校的新职师是重点抓手。打造“新职师”既可以指培养新职教教师，也可以指创新职教教师教育模式，还可以指建设新型职业技术师范院校。

——来源：《中国教育报》，2019年10月22日

加大力度建设高水平技术师范大学

庄西真

经过改革开放 40 多年的努力，我国职业教育教师队伍无论是数量还是质量都有了长足进步，支撑了过去 40 多年职业教育的快速发展。进入新时代，我国职业教育进入了体系建设和全面提高教育质量的新征程，对职业教育教师素质和能力提出了新挑战。

多元复杂职业情境对职业院校教师知能结构构成新挑战。职业院校学生形成岗位工作能力，不仅需要学习知识，还要学习如何操作，掌握两者间的联系，以应对多元复杂的职业情境。很多职业院校教师源于普通教育体系，难以将理论知识转化为工作所需要的实践知识，难以培养学生的动手操作能力。

技术技能人才成长规律对职业院校教师专业教学方法构成新挑战。技能人才有着独特的成长规律，其中最重要的是从新手成长为专家，在理论学习与工作学习中实现两者的融合与促进。职业院校教师只有遵循技能人才成长规律开展教育教学活动，才能提高教育教学质量。

持续有效培养高素质劳动者对职业院校教师职业认同感构成新挑战。职业院校教师要在有效的时间内把学生培养成各行各业中高素质劳动者，要求从业者不仅自身素质要硬，更要有职业认同感，形成关于教师职业的正确价值观。

职教教师队伍建设面临的挑战与职教教师培养密切相关，细究起来，我国职教教师培养存在以下几个突出问题：一是独立完整特色鲜明的职教教师培养培训体系尚未形成；二是缺少起带动辐射作用的中国特色高水平职技师范院校；三是缺少高水平的职教教师培养教材和职教教师培养模式。因此，有必要从以下几个方面入手，改进职教教师培养工作。

一要从国家经济社会事业发展的大局出发重视职教教师培养。人口是影响中国未来经济社会高质量发展的关键变量，能否把庞大的人口变成支撑中国经济高质量发展的高素质劳动者和技能人才，主要靠职业教育。职业教育能否完成这个任务，关键看教师队伍素质。教师队伍素质的高低，决定于教师培养的质量。过去若干年的现实以及教师职业的特点告诉我们，科学的、长期系统的培养是造就一支爱岗敬业、素质优良教师队伍的不二法门。抓住职教教师培养这个环节，就是抓住了职教教师队伍建设的“牛鼻子”。

二要优化布局结构，加快建立独立完整的现代化职教教师培养培训体系。根据我国不同地区间经济发展水平及经济发展对职业教育的规模、层次要求，来优化布局职教教师培养资源。如在经济发达地区的长三角、珠三角和京津冀地区各建设一所重点技术师范大学，在中西部地区建设一至两所技术师范大学。在已有职教教师培养架构基础上，建立由技术师范大学、技术师范学院、普通高校内设二级学院组成的独立完整的职教教师培养体系。

三要强化标准引领，以职教教师资格证书制度改革引领专业发展。构建以师德、教学能力、教研能力、应用科研能力等为主体的职业院校教师能力标准和评价体系。改革职业教育教师资格证书制度，完善资格等级及获取方式，构建职前职后一体化、校企双主体的职教教师培养体系及配套制度。

四要根据“分类发展、分类管理”原则，精准施策助推“新职师”建设。致力于培养职教教师的技术师范大学为我国职业教育和经济社会发展作出了很大贡献。现在我国经济进入高质量发展阶段，对职业教育教师培养提出了新要求，“老职师”建设与改革遇到了新问题，

政府相关部门（教育部）要在招生、就业、学位授予、经费等方面出台一揽子政策，支持这些院校建设“新职师”。

五要加大建设高水平技术师范大学力度，探索职教教师培养的中国模式。职业教育与普通教育是两种不同教育类型，具有同等重要地位。职技师范与普通师范是两种不同教师教育类型，具有同等重要地位。当务之急，要在全国遴选几所具有深厚职教教师培养积淀和职业教育科学研究基础的学校，给予优惠政策，支持其尽快成长为具有中国特色、展示中国成就、达到世界水平的技术师范大学。

六要加强对职教教师成长规律的深度研究。高水平职教师资的培养不仅是一项系统性的常规工作，更是一项复杂度更高、极具挑战性的创新工程。如果没有对职教教师成长规律的深度研究，很难对职教教师培养带来针对性的指导与帮助。当前阶段，尤其需要加强关于职教教师的岗位特征、知能结构以及专业成长阶段性特征的研究，真正拿出经得起理论和实践考验的优秀成果，从而为科学培养职教教师奠定扎实基础。

——来源：《中国教育报》，2020年11月17日

解决校企合作难题要回答好几道题

李德方

深化“产教融合、校企合作”也是党和政府给职业教育提出的时代命题。笔者以为，在实践中要有效解答这道难题，需要回答好彼此关联的几个问题。

首先需要回答的是，“产教融合、校企合作”是否有统一的模式？

所谓“产教融合、校企合作”，是指在产业部门和教育部门之间、实际生产经营过程和教育教学活动过程之间建立密切联系，发挥学校和企业双方优势，把以课堂传授间接知识为主的教育环境与直接获取实际经验能力为主的生产环境有机结合起来，进而达到培养高质量人才的目的。

“产教融合、校企合作”不仅因校而异、因专业而异，而且因人而异。

由于“产教融合、校企合作”是人才培养的一种手段，对于主要培养工程技术型人才的应用型本科高校与培养技术技能型人才的高职院校和中等职业学校，因为各自的培养目标并不完全相同，因而在人才培养过程中对企业资源的利用和介入程度也不尽一致，所以要因校而异。目前我国职业院校不同专业对“产教融合、校企合作”的需求也不尽相同，有的甚至差异很大。如机械制造专业和果蔬花卉生产技术专业，前者是和工厂企业合作，后者则是要到田间地头，与农户合作、向农民请教，两者不仅合作对象不同，合作方式、合作时间和合作内容也大相径庭。

相同层次、相同专业的不同学习者，对“产教融合、校企合作”的需求也不一致。如以升学为目标追求的学生，和以直接就业为目标的学生，两者在学习内容标准、学习时间分配和学习方式选择上也会不同。

事实上，“产教融合、校企合作”的复杂性、多样性和跨界性等特点，正是导致“产教融合、校企合作”长期难以深化的原因之一。加上我国幅员辽阔、民族多样和社会经济发展上的客观差异，很难用大一统的办法实施并推行“产教融合、校企合作”，也不会有一个“放之四海而皆准”的统一模式。

其次需要回答的是，没有统一模式的“产教融合、校企合作”，在实践中如何有效实施？

对于学校来说，在实践中有三种情形可供选择：一是所有的人才培养环节，包括由企业承担的教学事宜都是学校自己做，如新加坡的“教学工厂”就是这样的范本；二是如果做不到自给自足，那就努力练好内功，用自身的行动和服务赢得企业的青睐与支持，不难发现，凡是校企合作做得好的，无一不是教师服务社会有能力、学校人才培养质量有优势、解决企业难题有妙招的职业院校；三是如果做不到前两点，也要尽力履行好自己应尽的职责和义务，使校企合作效率最大化。

校企合作只有在双方各自发挥最大优势的情况下才是最有效率的，也是最理想的。职业院校首先要用足自身资源，做足自身功课，仅就自己无力作为之事才谋求与相关企业合作。换言之，只有职业学校承担了学校本身应该承担的通用知识和技能培养的任务后，尚需企业完成后续的专用知识与技术技能培养的情况下，才应选择校企合作。也就是说，“产教融合、校企合作”要本着实事求是的精神，合作之前首先要研究分析各个专业人才培养的具体环节，各个环节要达到的目的标准以及包括实施场所等在内的人才培养实施条件等要素，在此基础

上，带着明确的标准和方案寻找合作对象，否则就会反复上演实习学生沦为工厂廉价劳动力的闹剧。

再其次需要回答的是，合作企业在为学校解决人才培养所需时产生的成本该由谁来支付？

这些成本包括因合作培养人才而消耗的材料成本、使用设备设施的折旧成本、指导学生的人工成本、支付给学生的实习补贴和福利保险等。由于合作成本是在职业院校人才培养目标达成所必需的环节中产生的，理所当然要由培养这些人才的主体来支付。也就是说，全额拨款的公办职业院校由拨款机构承担，如市属学校由市财政承担，省属院校由省财政承担，其他类型学校由各自的举办者承担。过去长期把校企合作不能有效开展的板子打在企业身上是极不合理的，“产教融合、校企合作”问题的实质是校企双方在合作过程中产生的成本及其分担问题，解决这一问题的办法就是要回归常识、尊重规律、实事求是。

解决了“产教融合、校企合作”的培养成本，企业依然不热心怎么办？国务院办公厅《关于深化产教融合的若干意见》给出了明确答案，“加快推进教育‘放管服’改革，注重发挥市场机制配置非基本公共教育资源作用”。也就是说，如果学校举办者在支付上述培养成本的基础上，根据市场运行规律，给予合作企业一定比例的利润空间，使参与企业不仅不会亏本，还有一定的利润，使“产教融合、校企合作”成为合作企业新的创利渠道，这样可能还会催生出一批专职从事这项工作的新型企业。事实上，德国校企合作搞得很好，并不能完全归因于德国的传统习惯和文化氛围，研究结果表明，通过校企合作，德国的企业不仅能获得信誉这种无形资产，而且能获得一定的经济回报。同样，如果我国按照实事求是的原则，采取成本补偿，同时辅之以宣传、表彰和激励等措施，那么未来中国的“产教融合、校企合作”将不再是“剃头挑子一头热”，相反会形成学校热心、企业热情的“双热”局面，这道始终困扰我国职业教育发展的时代难题就会迎刃而解。

——来源：《中国教育报》，2019年3月19日

中共中央 国务院印发《深化新时代教育评价改革总体方案》

新华社北京10月13日电 近日，中共中央、国务院印发了《深化新时代教育评价改革总体方案》，并发出通知，要求各地区各部门结合实际认真贯彻落实。

《深化新时代教育评价改革总体方案》全文如下。

教育评价事关教育发展方向，有什么样的评价指挥棒，就有什么样的办学导向。为深入贯彻落实习近平总书记关于教育的重要论述和全国教育大会精神，完善立德树人体制机制，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，提高教育治理能力和水平，加快推进教育现代化、建设教育强国、办好人民满意的教育，现制定如下方案。

一、总体要求

（一）指导思想。以习近平新时代中国特色社会主义思想为指导，全面贯彻党的十九大和十九届二中、三中、四中全会精神，全面贯彻党的教育方针，坚持社会主义办学方向，落实立德树人根本任务，遵循教育规律，系统推进教育评价改革，发展素质教育，引导全党全社会树立科学的教育发展观、人才成长观、选人用人观，推动构建服务全民终身学习的教育体系，努力培养担当民族复兴大任的时代新人，培养德智体美劳全面发展的社会主义建设者和接班人。

（二）主要原则。坚持立德树人，牢记为党育人、为国育才使命，充分发挥教育评价的指挥棒作用，引导确立科学的育人目标，确保教育正确发展方向。坚持问题导向，从党中央关心、群众关切、社会关注的问题入手，破立并举，推进教育评价关键领域改革取得实质性突破。坚持科学有效，改进结果评价，强化过程评价，探索增值评价，健全综合评价，充分利用信息技术，提高教育评价的科学性、专业性、客观性。坚持统筹兼顾，针对不同主体和不同学段、不同类型教育特点，分类设计、稳步推进，增强改革的系统性、整体性、协同性。坚持中国特色，扎根中国、融通中外，立足时代、面向未来，坚定不移走中国特色社会主义教育发展道路。

（三）改革目标。经过5至10年努力，各级党委和政府科学履行职责水平明显提高，各级各类学校立德树人落实机制更加完善，引导教师潜心育人的评价制度更加健全，促进学生全面发展的评价办法更加多元，社会选人用人方式更加科学。到2035年，基本形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系。

二、重点任务

（一）改革党委和政府教育工作评价，推进科学履行职责

1.完善党对教育工作全面领导的体制机制。各级党委要认真落实领导责任，建立健全党委统一领导、党政齐抓共管、部门各负其责的教育领导体制，履行好把方向、管大局、作决策、保落实的职责，把思想政治工作作为学校各项工作的生命线紧紧抓在手上，贯穿学校教育管理全过程，牢固树立科学的教育发展理念，坚决克服短视行为、功利化倾向。各级党委和政府要完善定期研究教育工作机制，建立健全党政主要负责同志深入教育一线调研、为师生上思政课、联系学校和年终述职必述教育工作等制度。

2.完善政府履行教育职责评价。对省级政府主要考核全面贯彻党的教育方针和党中央关于教育工作的决策部署、落实教育优先发展战略、解决人民群众普遍关心的教育突出问题等情况，既评估最终结果，也考核努力程度及进步发展。各地根据国家层面确立的评价内容和指

标，结合实际进行细化，作为对下一级政府履行教育职责评价的依据。

3.坚决纠正片面追求升学率倾向。各级党委和政府要坚持正确政绩观，不得下达升学指标或以中高考升学率考核下一级党委和政府、教育部门、学校和教师，不得将升学率与学校工程项目、经费分配、评优评先等挂钩，不得通过任何形式以中高考成绩为标准奖励教师和学生，严禁公布、宣传、炒作中高考“状元”和升学率。对教育生态问题突出、造成严重社会影响的，依规依法问责追责。

（二）改革学校评价，推进落实立德树人根本任务

4.坚持把立德树人成效作为根本标准。加快完善各级各类学校评价标准，将落实党的全面领导、坚持正确办学方向、加强和改进学校党的建设以及党建带团建队建、做好思想政治工作和意识形态工作、依法治校办学、维护安全稳定作为评价学校及其领导人员、管理人员的重要内容，健全学校内部质量保障制度，坚决克服重智育轻德育、重分数轻素质等片面办学行为，促进学生身心健康、全面发展。

5.完善幼儿园评价。重点评价幼儿园科学保教、规范办园、安全卫生、队伍建设、克服小学化倾向等情况。国家制定幼儿园保教质量评估指南，各省（自治区、直辖市）完善幼儿园质量评估标准，将各类幼儿园纳入质量评估范畴，定期向社会公布评估结果。

6.改进中小学校评价。义务教育学校重点评价促进学生全面发展、保障学生平等权益、引领教师专业发展、提升教育教学水平、营造和谐育人环境、建设现代学校制度以及学业负担、社会满意度等情况。国家制定义务教育学校办学质量评价标准，完善义务教育质量监测制度，加强监测结果运用，促进义务教育优质均衡发展。普通高中主要评价学生全面发展的培养情况。国家制定普通高中办学质量评价标准，突出实施学生综合素质评价、开展学生发展指导、优化教学资源配置、有序推进选课走班、规范招生办学行为等内容。

7.健全职业学校评价。重点评价职业学校（含技工院校，下同）德技并修、产教融合、校企合作、育训结合、学生获取职业资格或职业技能等级证书、毕业生就业质量、“双师型”教师（含技工院校“一体化”教师，下同）队伍建设等情况，扩大行业企业参与评价，引导培养高素质劳动者和技术技能人才。深化职普融通，探索具有中国特色的高层次学徒制，完善与职业教育发展相适应的学位授予标准和评价机制。加大职业培训、服务区域和行业的评价权重，将承担职业培训情况作为核定职业学校教师绩效工资总量的重要依据，推动健全终身职业技能培训制度。

8.改进高等学校评价。推进高校分类评价，引导不同类型高校科学定位，办出特色和水平。改进本科教育教学评估，突出思想政治教育、教授为本科生上课、生师比、生均课程门数、优势特色专业、学位论文（毕业设计）指导、学生管理与服务、学生参加社会实践、毕业生发展、用人单位满意度等。改进学科评估，强化人才培养中心地位，淡化论文收录数、引用率、奖项数等数量指标，突出学科特色、质量和贡献，纠正片面以学术头衔评价学术水平的做法，教师成果严格按署名单位认定、不随人走。探索建立应用型本科评价标准，突出培养相应专业能力和实践能力。制定“双一流”建设成效评价办法，突出培养一流人才、产出一流成果、主动服务国家需求，引导高校争创世界一流。改进师范院校评价，把办好师范教育作为第一职责，将培养合格教师作为主要考核指标。改进高校经费使用绩效评价，引导高校加大对教育教学、基础研究的支持力度。改进高校国际交流合作评价，促进提升校际交流、来华留学、合作办学、海外人才引进等工作质量。探索开展高校服务全民终身学习情况评价，促进学习型社会建设。

（三）改革教师评价，推进践行教书育人使命

9.坚持把师德师风作为第一标准。坚决克服重科研轻教学、重教书轻育人等现象，把师德表现作为教师资格定期注册、业绩考核、职称评聘、评优奖励首要要求，强化教师思想政治素质考察，推动师德师风建设常态化、长效化。健全教师荣誉制度，发挥典型示范引领作用。全面落实新时代幼儿园、中小学、高校教师职业行为准则，建立师德失范行为通报警示制度。对出现严重师德师风问题的教师，探索实施教育全行业禁入制度。

10.突出教育教学实绩。把认真履行教育教学职责作为评价教师的基本要求，引导教师上好每一节课、关爱每一个学生。幼儿园教师评价突出保教实践，把以游戏为基本活动促进儿童主动学习和全面发展的能力作为关键指标，纳入学前教育专业人才培养标准、幼儿教师职后培训重要内容。探索建立中小学教师教学述评制度，任课教师每学期须对每个学生进行学业述评，述评情况纳入教师考核内容。完善中小学教师绩效考核办法，绩效工资分配向班主任倾斜，向教学一线和教育教学效果突出的教师倾斜。健全“双师型”教师认定、聘用、考核等评价标准，突出实践技能水平和专业教学能力。规范高校教师聘用和职称评聘条件设置，不得将国（境）外学习经历作为限制性条件。把参与教研活动，编写教材、案例，指导学生毕业设计、就业、创新创业、社会实践、社团活动、竞赛展演等计入工作量。落实教授上课制度，高校应明确教授承担本（专）科生教学最低课时要求，确保教学质量，对未达到要求的给予年度或聘期考核不合格处理。支持建设高质量教学研究类学术期刊，鼓励高校学报向教学研究倾斜。完善教材质量监控和评价机制，实施教材建设国家奖励制度，每四年评选一次，对作出突出贡献的教师按规定进行表彰奖励。完善国家教学成果奖评选制度，优化获奖种类和入选名额分配。

11.强化一线学生工作。各级各类学校要明确领导干部和教师参与学生工作的具体要求。落实中小学教师家访制度，将家校联系情况纳入教师考核。高校领导班子成员年度述职要把上思政课、联系学生情况作为重要内容。完善学校党政管理干部选拔任用机制，原则上应有思政课教师、辅导员或班主任等学生工作经历。高校青年教师晋升高一级职称，至少须有一年担任辅导员、班主任等学生工作经历。

12.改进高校教师科研评价。突出质量导向，重点评价学术贡献、社会贡献以及支撑人才培养情况，不得将论文数、项目数、课题经费等科研量化指标与绩效工资分配、奖励挂钩。根据不同学科、不同岗位特点，坚持分类评价，推行代表性成果评价，探索长周期评价，完善同行专家评议机制，注重个人评价与团队评价相结合。探索国防科技等特殊领域教师科研专门评价办法。对取得重大理论创新成果、前沿技术突破、解决重大工程技术难题、在经济社会事业发展中作出重大贡献的，申报高级职称时论文可不作限制性要求。

13.推进人才称号回归学术性、荣誉性。切实精简人才“帽子”，优化整合涉教育领域各类人才计划。不得把人才称号作为承担科研项目、职称评聘、评优评奖、学位点申报的限制性条件，有关申报书不得设置填写人才称号栏目。依据实际贡献合理确定人才薪酬，不得将人才称号与物质利益简单挂钩。鼓励中西部、东北地区高校“长江学者”等人才称号入选者与学校签订长期服务合同，为实施国家和区域发展战略贡献力量。

（四）改革学生评价，促进德智体美劳全面发展

14.树立科学成才观念。坚持以德为先、能力为重、全面发展，坚持面向人人、因材施教、知行合一，坚决改变用分数给学生贴标签的做法，创新德智体美劳过程性评价办法，完善综合素质评价体系，切实引导学生坚定理想信念、厚植爱国主义情怀、加强品德修养、增长知

识见识、培养奋斗精神、增强综合素质。

15.完善德育评价。根据学生不同阶段身心特点，科学设计各级各类教育德育目标要求，引导学生养成良好思想道德、心理素质和行为习惯，传承红色基因，增强“四个自信”，立志听党话、跟党走，立志扎根人民、奉献国家。通过信息化等手段，探索学生、家长、教师以及社区等参与评价的有效方式，客观记录学生品行日常表现和突出表现，特别是践行社会主义核心价值观情况，将其作为学生综合素质评价的重要内容。

16.强化体育评价。建立日常参与、体质监测和专项运动技能测试相结合的考查机制，将达到国家学生体质健康标准要求作为教育教学考核的重要内容，引导学生养成良好锻炼习惯和健康生活方式，锤炼坚强意志，培养合作精神。中小学校要客观记录学生日常体育参与情况和体质健康监测结果，定期向家长反馈。改进中考体育测试内容、方式和计分办法，形成激励学生加强体育锻炼的有效机制。加强大学生体育评价，探索在高等教育所有阶段开设体育课程。

17.改进美育评价。把中小学生学习音乐、美术、书法等艺术类课程以及参与学校组织的艺术实践活动情况纳入学业要求，促进学生形成艺术爱好、增强艺术素养，全面提升学生感受美、表现美、鉴赏美、创造美的能力。探索将艺术类科目纳入中考改革试点。推动高校将公共艺术课程与艺术实践纳入人才培养方案，实行学分制管理，学生修满规定学分方能毕业。

18.加强劳动教育评价。实施大中小学劳动教育指导纲要，明确不同学段、不同年级劳动教育的目标要求，引导学生崇尚劳动、尊重劳动。探索建立劳动清单制度，明确学生参加劳动的具体内容和要求，让学生在实践中养成劳动习惯，学会劳动、学会勤俭。加强过程性评价，将参与劳动教育课程学习和实践情况纳入学生综合素质档案。

19.严格学业标准。完善各级各类学校学生学业要求，严把出口关。对初、高中毕业班学生，学校须合理安排中高考结束后至暑假前的教育活动。完善过程性考核与结果性考核有机结合的学业考评制度，加强课堂参与和课堂纪律考查，引导学生树立良好学风。探索学士学位论文（毕业设计）抽检试点工作，完善博士、硕士学位论文抽检工作，严肃处理各类学术不端行为。完善实习（实训）考核办法，确保学生足额、真实参加实习（实训）。

20.深化考试招生制度改革。稳步推进中高考改革，构建引导学生德智体美劳全面发展的考试内容体系，改变相对固化的试题形式，增强试题开放性，减少死记硬背和“机械刷题”现象。加快完善初、高中学生综合素质档案建设和使用办法，逐步转变简单以考试成绩为唯一标准的招生模式。完善高等职业教育“文化素质+职业技能”考试招生办法。深化研究生考试招生改革，加强科研创新能力和实践能力考查。各级各类学校不得通过设置奖金等方式违规争抢生源。探索建立学分银行制度，推动多种形式学习成果的认定、积累和转换，实现不同类型教育、学历与非学历教育、校内与校外教育之间互通衔接，畅通终身学习和人才成长渠道。

（五）改革用人评价，共同营造教育发展良好环境

21.树立正确用人导向。党政机关、事业单位、国有企业要带头扭转“唯名校”、“唯学历”的用人导向，建立以品德和能力为导向、以岗位需求为目标的人才使用机制，改变人才“高消费”状况，形成不拘一格降人才的良好局面。

22.促进人岗相适。各级公务员招录、事业单位和国有企业招聘要按照岗位需求合理制定招考条件、确定学历层次，在招聘公告和实际操作中不得将毕业院校、国（境）外学习经历、学习方式作为限制性条件。职业学校毕业生在落户、就业、参加机关企事业单位招聘、职称评聘、职务职级晋升等方面，与普通学校毕业生同等对待。用人单位要科学合理确定岗位职

责，坚持以岗定薪、按劳取酬、优劳优酬，建立重实绩、重贡献的激励机制。

三、组织实施

（一）落实改革责任。各级党委和政府要加强组织领导，把深化教育评价改革列入重要议事日程，根据本方案要求，结合实际明确落实举措。各级党委教育工作领导小组要加强统筹协调、宣传引导和督促落实。中央和国家机关有关部门要结合职责，及时制定配套制度。各级各类学校要狠抓落实，切实破除“五唯”顽瘴痼疾。国家和各省（自治区、直辖市）选择有条件的地方、学校和单位进行试点，发挥示范带动作用。教育督导要将推进教育评价改革情况作为重要内容，对违反相关规定的予以督促纠正，依法依规对相关责任人员严肃处理。

（二）加强专业化建设。构建政府、学校、社会等多元参与的评价体系，建立健全教育督导部门统一负责的教育评估监测机制，发挥专业机构和社会组织作用。严格控制教育评价活动数量和频次，减少多头评价、重复评价，切实减轻基层和学校负担。各地要创新基础教育教研工作指导方式，严格控制以考试方式抽检评测学校和学生。创新评价工具，利用人工智能、大数据等现代信息技术，探索开展学生各年级学习情况全过程纵向评价、德智体美劳全要素横向评价。完善评价结果运用，综合发挥导向、鉴定、诊断、调控和改进作用。加强教师教育评价能力建设，支持有条件的高校设立教育评价、教育测量等相关学科专业，培养教育评价专门人才。加强国家教育考试工作队伍建设，完善教师参与命题和考务工作的激励机制。积极开展教育评价国际合作，参与联合国2030年可持续发展议程教育目标实施监测评估，彰显中国理念，贡献中国方案。

（三）营造良好氛围。党政机关、事业单位、国有企业要履职尽责，带动全社会形成科学的选人用人理念。新闻媒体要加大对科学教育理念和改革政策的宣传解读力度，合理引导预期，增进社会共识。构建覆盖城乡的家庭教育指导服务体系，引导广大家长树立正确的教育观和成才观。各地要及时总结、宣传、推广教育评价改革的成功经验和典型案例，扩大辐射面，提高影响力。

以评价改革牵引教育领域综合改革

刘自成

党的十九届五中全会是党在“两个一百年”历史交汇期、“两个大局”演进关键期、“两大奇迹”迭代升级期召开的一次具有全局性意义、开局性特点和格局性影响的历史性会议。全会审议通过的《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》，“改革”是高频词、关键词。教育评价是教育改革的龙头，事关教育改革的导向。贯彻好全会精神，要认真落实习近平总书记关于抓好深化新时代教育评价改革总体方案落实落地的重要指示，坚持以评价改革为牵引，全面深化教育领域综合改革，实现改革与育人深度融合、高效联动，推动学习贯彻五中全会精神往深里走、往心里走、往实里走。

一是以加强党对教育工作的全面领导为方向。中国共产党是引领我们正确前行的“指南针”，也是我们攻坚克难的“主心骨”，更是我们应对风险挑战的“定盘星”。《建议》明确把坚持党的全面领导作为“十四五”时期经济社会发展必须遵循的首要原则。中共中央、国务院印发的《深化新时代教育评价改革总体方案》，列专节对改革党委和政府教育工作评价进行部署，着力推动各级党委和政府科学履行职责。要推动完善教育领导体制，着力健全党委统一领导、党政齐抓共管、部门各负其责的教育领导体制，推动各级党委认真履行好把方向、管大局、作决策、保落实的领导职责，强化党委教育工作领导小组职能作用，加快构建党的领导纵到底、横到边、全覆盖的工作格局。要推动加强思想政治工作，坚持以习近平新时代中国特色社会主义思想铸魂育人，推动各级党委和政府、各级教育主管部门、学校党组织把思政工作作为生命线紧紧抓在手上，贯穿教育管理、学校办学、教师教学全过程，引导广大师生听党话、跟党走，扎根人民、奉献国家。要推动完善党中央重大教育决策部署落实机制，推动各级党委和政府及时传达学习习近平总书记关于教育的重要论述和指示批示，定期研究教育工作、协调解决重要教育问题，坚持和践行科学的教育发展理念，把旗帜鲜明讲政治转化为坚决贯彻落实党中央决策部署的实际行动。

二是以推进落实立德树人根本任务为主线。立德树人关乎党的事业后继有人，关乎国家前途命运。《建议》明确提出要全面贯彻党的教育方针，坚持立德树人，加强师德师风建设，培养德智体美劳全面发展的社会主义建设者和接班人。《总体方案》着眼全面贯彻党的教育方针，以落实立德树人根本任务为主线，统筹设计各项重点任务和改革举措。要着眼学校这一主阵地，全面推进学校评价改革，坚持把立德树人成效作为评价学校的根本标准，以人才培养为中心，改进幼儿园、中小学、职业学校和高校评价机制和方式方法，健全学校内部质量保障制度，坚决克服重智育轻德育、重分数轻素质的片面办学行为，促进学生身心健康、全面发展。要聚焦教师这一引路人，全面推进教师评价改革，坚持把师德师风作为评价教师的第一标准，把认真履行教育教学职责作为评价教师的基本要求，强化教师思想政治素质考察，突出教育教学实绩，改进高校教师科研评价，坚决克服重科研轻教学、重教书轻育人等现象，引导广大教师认真践行教书育人使命、提升教书育人能力素质。要围绕学生这一需求侧，全面推进学生评价改革，坚持以德为先、能力为重、全面发展，坚持面向人人、因材施教、知行合一，坚决克服分数至上、用分数给学生贴标签的做法，创新德智体美劳过程性评价办法，完善综合素质评价体系，培养学生爱国情怀、社会责任感、创新精神、实践能力。

三是以坚持系统观念的方法论为遵循。系统观念是马克思主义的思想观点，也是我们党领导革命、建设和改革的重要方法。《建议》把坚持系统观念作为“十四五”时期经济社会

发展必须遵循的重要原则，明确要加强前瞻性思考、全局性谋划、战略性布局、整体性推进，实现发展质量、结构、规模、速度、效益、安全相统一。这为我们办好教育这一关系千秋万代和千家万户的事业，推进教育改革这一涉及方方面面、需要抓纲带目的工作，提供了具有基础性的思想和方法。针对教育评价这一历史性、实践性、世界性难题，《总体方案》坚持从系统观念出发加以谋划和解决，全面协调推动各领域各环节各方面评价改革。要统筹推进党委和政府、学校、教师、学生及用人单位评价改革，紧盯突出问题，强化改革的战略性、前瞻性、针对性，推进解决单兵难以突进的难题，增强政策的系统集成和改革的综合效能。要一体改进结果评价、强化过程评价、探索增值评价、健全综合评价，整体提升教育评价的客观性、服务师生成长的发展性、对教育教学改进的指导性、评价落地实施的科学性和公平性。要全面贯穿教、育、考、招、用各个环节，着力打通教育评价改革中的断头路、易堵点，消除不同环节改革的温差、色差，确保教育内外各环节改革同频共振、同向发力。要全面覆盖基础教育、职业教育、高等教育各领域和大中小幼各学段，统筹推进育人方式、办学模式、管理体制、保障机制改革，建设高质量教育体系，有力服务国家新发展格局。

【刘自成，教育部综合改革司党支部书记、司长。】

——来源：《中国教育报》，2020年11月26日

吹响新时代教育评价改革的号角

杜玉波

深化教育评价改革，关乎教育发展方向，关乎党的教育方针的全面贯彻，关乎教育发展良好环境。近日，中共中央、国务院印发《深化新时代教育评价改革总体方案》（以下简称《总体方案》），这是深入贯彻落实习近平总书记关于教育的重要论述和全国教育大会精神的重要举措，是新中国第一个关于教育评价系统改革的文件。《总体方案》的出台实施，对于全面贯彻党的教育方针，完善立德树人体制机制，破除“五唯”顽瘴痼疾，引导全党全社会树立科学的教育发展观、人才成长观、选人用人观具有重大意义。

具有突出的时代性、系统性和针对性

教育评价改革是一项世界性、历史性、实践性难题，涉及多重因素、不同主体，牵一发而动全身。《总体方案》针对教育评价的突出问题作出精准设计，从党中央关心、群众关切、社会关注的问题入手，着力破除不科学、不合理的教育评价导向和做法，对学校不再简单以升学率来评价，对教师不再简单以考分排名来评价，对学生不再简单以考试成绩来评价，党政机关、企事业单位对用人不再简单“唯名校”“唯学历”是举。

针对党委和政府、学校、教师、学生、社会五类关键主体，分类设计，充分考虑不同教育领域、不同学段、不同对象的特点，提出相关改革思路、措施和实施路径。紧扣举措落实落地作出统筹设计，坚持长远结合、稳步推进，提出“两步走”的改革目标，同时强调各级党委和政府要加强组织领导、各级各类学校要狠抓落实、加强专业化建设、营造良好氛围等，努力做到上下结合、左右结合、内外结合，确保改革形成合力。《总体方案》这些政策设计体现了攻坚克难的决心，做到了目标导向与问题导向相结合，具有突出的时代性、系统性和针对性。

把落实立德树人根本任务作为主线

培养社会主义建设者和接班人是教育工作的根本任务，也是教育现代化的方向目标。《总体方案》把落实立德树人根本任务、培养德智体美劳全面发展的社会主义建设者和接班人作为主线，贯穿于教育评价改革各项任务始终。

在学校评价上，坚持把立德树人成效作为根本标准，将落实党的全面领导、坚持正确办学方向、加强和改进学校党的建设以及党建带团建队建、做好思想政治工作和意识形态工作等作为重要评价内容，坚决克服重智育轻德育、重分数轻素质等片面办学行为。在教师评价上，坚持把师德师风作为第一标准，把师德表现作为教师资格定期注册、业绩考核等的首要要求，推动师德师风建设常态化、长效化。在学生评价上，坚持以德为先、能力为重、全面发展，坚持面向人人、因材施教、知行合一，完善德育评价，强化体育评价，改进美育评价，加强劳动教育评价，严格学业标准。这些举措都进一步明确和强化了全面贯彻党的教育方针的任务要求，有利于促进德智体美劳全面培养的教育体系和更高水平的人才培养体系加快形成。

形成破立结合、全面发力的举措

“唯分数、唯升学、唯文凭、唯论文、唯帽子”的顽瘴痼疾严重违背教育规律、人才成长规律，是教育评价改革的重点和难点。《总体方案》紧扣破除“五唯”，妥善处理“破”与“立”的辩证关系，形成一整套破立结合、全面发力的改革举措。

比如，破除“唯升学”，坚决纠正片面追求升学率倾向；破除“唯帽子”，不得把人才称号作为承担科研项目、职称评聘、评优评奖、学位点申报的限制性条件，体现的是改进结果评价，扭转功利导向。完善过程性考核与结果性考核有机结合的学业考评制度，改进高校教师科研评价，探索长周期评价等，体现的是强化过程评价，尊重客观规律。完善政府履行教育职责评价，在对省级政府评价上，既评估最终结果，也考核努力程度及进步发展；在学生评价上，要求探索开展学生各年级学习情况全过程纵向评价，体现的是探索增值评价，挖掘发展潜能。构建引导学生德智体美劳全面发展的考试内容体系，逐步转变简单以考试成绩为唯一标准的招生模式等，体现的是健全综合评价，力避片面倾向。4个评价有机统一，构成一个整体，形成全新的评价理念和方法，系统构建起科学的、符合时代要求的教育评价制度和机制。

新时代教育评价改革的号角已经吹响。这是在习近平新时代中国特色社会主义思想指引下，立足中国实际解决中国教育问题的一次重大制度突破和体制创新，充分彰显了中国智慧，展现了中国自信。相信《总体方案》的出台实施，必将显著提升我国教育治理能力和水平，推动发展具有中国特色、世界水平的现代教育迈出有力步伐。

【杜玉波，全国人大常委会委员、教科文卫委副主任委员、中国高等教育学会会长。】

——来源：《人民日报》，2020年10月22日

深化新时代教育评价改革研究（笔谈）

加强高等教育评价方式的改革探索

马陆亭

“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”，是健全立德树人落实机制的重要内容。教育评价是教育发展的指挥棒，是教育进入全面普及时期能否健康前行、科学发展的重要前提，在新时代具有特殊意义。既是思想，更是方法。

1. “五唯”具有发展阶段特征。“五唯”的形成起于数量、成于刚性、错在单一。应该看到，分数、学历、论文等可量化指标有其合理性的一面，如表面公平、操作简单，否则也不会累积成顽瘴痼疾难题。在精英教育时期构不成社会问题，而且让人信服，大众化、普及化阶段问题逐渐凸显，是典型的阶段性问题。

2. 破“五唯”就是要实现从“一”到“多”的“立”。“五唯”错在“唯”字上，错在偏激、过分。破“五唯”改革其实就是以量化指标为基础，找非量化点突破，以此逐步建立起适应新时代的教育评价体系。即数量指标是基础，“不唯”是扩展，定量与定性相结合。

3. 抓住核心内涵进行评价。“本质内涵、多元评价”是趋势，内涵是什么就评什么，“多”的地方不再刚性。内涵式发展即按本质属性发展，如每所高校都要实现教学、科研、社会服务三大职能，但侧重不同。

4. 架构高等学校体系。党的十九届四中全会做出的《关于坚持和完善中国特色社会主义制度 推进国家治理体系和治理能力现代化若干重大问题的决定》为什么在教育领域提出的是“构建服务全民终身学习的教育体系”？我们可以将我国高校分为三大类：理论型、实践型和职业型。当然，这只是一个粗线条的划分，各种类型高校之间有可能会产生“光谱地带”，但方向导向还是需要明确的。教育模式的改革方向是理论型高校要明显加强对学生的思维能力训练，实践型高校要加强综合应用能力训练，职业型高校要侧重动手能力操作能力训练。各自的教学安排即课堂理论、综合训练、实践操作的学时比例不同。以此作为三类高校的分野，而不是唯层次。教学、科研、社会服务三位一体共同遵循。

5. 探索增值评价。增值评价是对评价对象的进步幅度进行测评的一种评价方法，改变了过去学校评价重投入轻过程、重选拔轻培养的倾向，转向更加注重教育过程，探索实现以评促教、以评促学的评价改革。

6. 推动人工智能技术应用。最重要的是智能自吹响新时代教育评价改革的号角适应教育的应用和区块链技术的应用。智能自适应教育可使每位学生拥有自己独有的学习路径，能有效解决教育规模化与个性化的矛盾，实现规模化的个性教育；区块链技术具有透明化、数据不可篡改等特征，将对各级各类考试、各种学历学位证书的记录方式产生重大影响。

一流大学更要培养一流人才

王小梅

目前，正逢首批入选国家“双一流”建设高校中期检查刚刚结束。在周期总结报告中，各校对建设成绩进行了系统全面的梳理，专家的评议意见也给予了充分肯定。如有的高校的

周期总结报告指出，“双一流”建设以来，学校的一流学科建设持续领跑、一流人才培养体系加快形成、一流师资队伍建设成绩斐然、一流科研取得突破性进展、一流国家高端智库建设树立典范、一流大学文化固本铸魂、一流的国际交流与合作深度拓展、一流的制度体系提供有力支撑。这里的一流人才培养体系是“加快形成”，与其他的“持续领跑”“成绩斐然”“树立典范”“突破性进展”相比，是不是分量不够重？底气不够足？当然，我们不排除人才培养是一个隐性、滞后的过程，即使投入很大也很难看到立竿见影的效果。也有专家认为，一流人才是在社会实践和科研工作实践中历练成长的，但这不能排除学校教育的基础性作用。

党的十八大以来，习近平总书记围绕“培养社会主义建设者和接班人”做出一系列重要论述，深刻回答了“培养什么人、怎样培养人、为谁培养人”这一根本性问题。2019年3月，习近平总书记在学校思想政治理论课教师座谈会上进一步强调，我们党立志于中华民族千秋伟业，必须培养一代又一代拥护中国共产党领导和我国社会主义制度、立志为中国特色社会主义事业奋斗终身的有用人才。这为中国特色社会主义高等教育事业发展指明了方向，为新时代我国高等教育的改革发展提供了根本遵循。

“双一流”建设高校是国之重器，是国家高等教育核心竞争力的代表，承担着一流人才培养的时代重任。“只有培养出一流人才的高校才能成为世界一流大学。”“双一流”建设高校更应抓住“立德树人”这个根本任务，更加倾心用力培养一流人才，培养有理想、有本领、有担当，堪当时代大任的青年学生。

“古今中外，每个国家都是按照自己的政治要求来培养人的，世界一流大学都是在服务自己国家发展中成长起来的。”当前，我们国家对高等教育的需求、对科学知识和优秀拔尖人才的需求，比以往任何时候都更加迫切。大学发展到今天，承载的使命和责任越来越多。从人才培养、科学研究、社会服务，延伸到文化传承创新和国际交流合作，但究其根本，立德树人才是大学的初心和教育的本质。“双一流”建设高校要始终把立德树人作为检验学校一切工作的根本标准，把培养一流人才作为重要的价值取向。在推进“双一流”建设的过程中要始终坚持社会主义办学方向，牢牢抓住全面提高人才培养能力这个核心点，谋划和把握“双一流”建设大方向，并以此来带动其他工作。

“双一流”建设高校与其他高校相比有更浓郁的科研氛围、更强的学科实力。一方面我们要引导教师将自己的学术研究旨趣与国家的重大战略需求、社会发展的重大民生问题紧密结合；另一方面，更加鼓励和引导教师将自己的科研优势快速转化为教育教学内容，转化为培育高层次一流人才的优势，要全面提升一流人才培养能力，构建高水平人才培养体系。近日教育部、国家发展改革委、财政部联合印发的《关于加快新时代研究生教育发展的意见》明确提出：以国家重大战略、关键领域和社会重大需求为重点，增设一批硕士、博士专业学位类别。聚焦数理化、文史哲等基础学科，以强化原始创新能力为导向，实施高层次人才培养专项。统筹一流学科、一流师资、一流平台等资源，以超常规方式加快培养一批紧缺人才，为国家解决“卡脖子”问题和推进科技创新作出贡献。作为我国高等教育最高水平的代表，“双一流”建设高校要把培养国家战略紧缺人才当作责任与使命。要积极鼓励和支持学生参与一流学科领域的科研项目。为培养世界级科技大师、领军人才、拔尖人才打下坚实的基础。

在“双一流”评价指标体系中要更加强化“立德树人”的评价导向。可以说，培养、拥有一流人才是一所大学成为世界一流大学的重要条件。综观当今世界公认的一流大学，无不把培养引领未来社会发展的一流人才作为核心办学理念。在“双一流”评价中我们要探索把对一流大学、一流学科建设的评价与为国家培养急需的高层次一流人才的要求紧密结合起来。

要深入研究办学规律、人才成长规律，在人才培养上要“风物长宜放眼量”，要有定力、定心，力戒浮躁、不折腾，脚踏实地、守望教育、静待花开。

人才是实现民族振兴、赢得国际竞争主动的战略资源。“双一流”建设高校要坚持为党育人、为国育才，落实立德树人根本任务，围绕“培养一流人才、孕育一流成果、做出一流贡献”的目标，创新人才培养模式、构建高水平人才培养体系，努力为国家发展和社会进步提供可靠的、源源不断的一流人才。

新时代我国教育评价改革的若干新要求

刘复兴

新时代我们要实现教育现代化与建设教育强国，教育改革创新面临着许多重大的挑战和新的历史任务。教育评价改革在教育改革创新中具有关键的、前瞻性、引领的地位，事关教育发展方向。新时代我国的教育评价改革必须从中国特色社会主义新时代发展的总体要求出发，以习近平总书记关于教育的重要论述为指导思想，以教育评价的全面性、整体性、系统性改革为方法论原则，把握好教育评价改革的若干新要求。

一是教育评价改革要树立新标准。党的十八大以来，特别是党的十九大以后，习近平总书记作出了一系列关于教育的重要论述。在全国教育大会上提出了教育“九个坚持”新理念新思想新观点，确立了新时代我国教育改革的指导思想；在2016年全国高校思政会上鲜明地提出了教育的“四为”方针，明确了新时代我国教育改革创新的基本方针。党和国家陆续提出了一系列人才培养的新标准、新要求，如坚持立德树人，培养创新型人才，培养学生的必备品格与关键能力，培养党的民族复兴大任的时代新人，实施劳动教育等。新时代党和国家人才培养的新标准新要求是教育改革创新的新任务。教育评价改革必须遵循教育“九个坚持”，贯彻“四为”方针，把立德树人作为根本任务，把党和国家一系列人才培养的新要求作为新标准。

二是教育评价改革要面向新问题。近代以来，为了富国强兵，救亡图存，我们开始逐步学习国外特别是西方的科学技术。西学东渐成为过去近三百年来我国知识传播的基本特点。特别是改革开放以来，这个进程大大加快了，范围大大拓宽了，并且某种意义上成为改革开放的重要智力支撑。今天我们面对全球百年未有之大变局，全球化进程遭受逆转，贸易战、科技战、外交战等成为常态。科技竞争与文明冲突日益成为国际关系的主流。主要依靠西学东渐已不可能支撑我们实现第二个一百年的奋斗目标。在创新驱动发展的时代，我国必须更多地依靠自己的本土创新。教育是创新的沃土，是培养创新人才的主要途径。如何大力促进培养创新人才？如何鼓励和扶持中国本土科技创新？如何实现中国本土创新与西学东渐之间的平衡，是新时代教育评价改革必须要面向的新问题。

三是教育评价改革要完成新任务。一般而言，人类社会的知识体系是由自然科学与人文社会科学两个主要的部分组成的。自然科学承担着认识自然、改造自然的任务，人文社会科学承担着认识人与社会、改造人与社会的重任。自然科学往往是价值无涉的，人文社会科学则承担着建构一个国家民族价值体系与传承民族文化传统的根本任务，是一类与自然科学相比很不同的知识体系，是与文明、思想、价值、文化、意识形态等要素息息相关的，也是因文明而不同、因民族而不同甚至是因国家而不同的。在新时代，我们要建设中国特色社会主义现代化强国，就要具备领先世界的中国特色知识体系，不仅要有领先世界的自然科学技术，还要具有引领世界的人文社会科学体系。教育，主要是大学特别是研究型大学肩负着生产与

创造知识体系的重任。如何解决人文社会科学评价的难题，以马克思主义为指导、继承中国传统文化与发扬社会主义核心价值观，引领建设中国特色知识体系特别是建设中国特色人文社会科学知识体系，是新时代教育评价改革的新任务。

四是教育评价改革要确立新主体。传统上，我国的教育评价十分注重教育的结果评价、学科教学的评价和科学研究的评价，而且许多的评价活动是以内部评价为主。教育评价的主体与内容相对是比较狭窄的、单一的。这种倾向也是导致出现“五唯”现象的原因之一。新时代，我们要建立科学的、符合时代要求的教育评价制度和机制，就必须拓宽教育评价的主体和内容。要重视外部评价、社会评价与学生评价，加强受教育者在教育评价中的地位；要特别突出人才培养评价的重要性，切实建立健全落实立德树人体制机制；要改进结果评价，强化过程评价；探索增值评价，健全综合评价，系统研究与建立软性指标的评价机制，补足在德育、体育、美育和劳动教育评价方面的短板；要注重发展性评价和诊断性评价，把评价作为促进学生、教师和学校发展的重要手段。

五是教育评价改革要利用新技术。新时代是一个未来已来、将至已至的时代，以5G、人工智能、互联网、物联网、大数据、区块链和量子科学等众多前沿科技创新叠加发展为标志的第四次工业革命正在来临。技术变革正在决定与改变着社会结构，不断变革的社会结构又要求教育结构的变革。第四次工业革命背景下的教育结构是以人工智能、互联网、物联网、区块链为物质和技术基础，嵌入万物互联的社会结构，是万物互联的组成部分。教育的形态由工业时代的教育转变为信息化教育、算法教育、互联网教育、物联网+教育、人工智能+教育、区块链+教育；教育制度体系由金字塔状的以正式制度、刚性制度、等级制度为主要特征的制度体系转变为立体的甚至是一种边际不断扩展的球状的多回路、网络式由正式制度和非正式制度共同构成的弹性制度体系。新时代教育评价必须前瞻性地适应这种教育结构体系的变革，充分利用5G、人工智能、互联网、物联网、大数据、区块链等前沿技术构建科学的高效的评价制度。

六是教育评价改革要明确新路径。面对全球百年未有之大变局以及全球化进程遭受逆转的现状，我们绝不能关起门来搞教育评价改革，应该进一步解放思想，实事求是，以更加开放的姿态推进我国教育的改革创新。进一步加强教育评价领域的国际交流与合作。在持续加强与欧美发达国家交流合作的基础上，特别要加强“一带一路”方向国家的交流与合作。注重构建“N-1”多边合作交流机制。在后疫情时代有效恢复与开展与疫情控制良好的国家与地区的合作与交流。

大学评价必须超越指标

周光礼

“双一流”建设存在一个悖论。一方面，作为国家“千亿工程”，“双一流”建设需要进行社会问责与项目评估，需要寻找国际可比指标。因为脱离国际标准另搞一套，是永远不可能建成世界一流大学的；另一方面，作为一种文化现象，一个国家的大学是与一个国家的文化模式相适应的。从逻辑上来看，国家与国家之间的大学不具有可比性。我们如何衡量一所大学的办学成绩，如何确定中国大学在世界的坐标，如何评判“双一流”建设成效，教育评价这把尺子至关重要。长久以来，中国大学始终被缺乏合理、准确、系统的评价标准所困扰，各种国内外的商业机构高举量化评价的大旗，以文献计量方法建构各种“洋标准”“洋指标”“洋名次”，给中国大学乃至全球大学产生巨大的社会和舆论压力。实际上，一切教

育评价都是主观的。在此，我们有必要反思一下世界大学排名的特点和弊端。

1. 重规模，轻质量。世界大学排行榜的一个突出特点是重视大学规模和数量指标，轻视大学内涵和质量指标。无论是 ESI、U. S. News 还是 THE、QS，无一例外都重视大学组织的规模和整体水平，对大学的办学质量和效率关注不够。这些世界大学排名均强调发表论著的数量以及大学学科齐全，这实质上是一种“以数量取代质量”的不良倾向。作为发达国家话语权的载体，世界大学排名推动了西方大学概念的全球扩散。这些大学排名奉行一套西方研究型大学的逻辑：世界一流大学必须拥有世界一流的科学研究，世界一流的科学研究需要一支世界一流的教师队伍，建设世界一流的教师队伍必须坚持 3A 原则（学术自由、学术自治、学术中立）。这种大学逻辑在世界大学排名的导向下，迅速向全球扩散，成为广大发展中国家建设世界一流大学的“黄金标准”。其实，这只不过是西方精英大学的理念和意识形态。中国传统的大学概念与这种大学逻辑是不同的。中国传统的大学概念是“大学之道，在明明德、在亲民、在止于至善”。这里至少包含两层含义，一是重视立德树人的人文主义教育，二是强调教育与政治交叉融合。过于强调这种“国际标准”容易导致全球大学趋同化，不利于“扎根中国大地办大学”，不利于实现立德树人根本目标。

2. 重投入，轻产出。由于大学办学质量难以测量，世界大学排名机构流行用投入指标代替产出指标，最荒谬的是用生源质量代替毕业质量，用高考分数代替人才培养质量。其背后的逻辑是：投入多，有质量；投入少，没质量。在这种逻辑的影响下，中国大学评价高度关注生师比、生均图书资料、生均校园面积、生均建筑面积、生均办学经费等，其结果是把中国大学导向“成本最大化”的办学路径。其实，成本最大并不意味着办学质量最高，铺张浪费办不成世界一流大学。值得指出的是，投入指标只能反映大学开展教学和科研活动的资源供应水平，充其量只能决定大学组织的正常运转，只是大学实现卓越的条件。根据大学评价的“柏林原则”，大学评价必须坚持产出导向的原则。也就是说，大学评价指标应尽可能选择产出指标而非投入指标。因为投入指标只能反映大学的一般情况，产出指标才能提供大学声望和质量更为精确的评价。基于此，“深化新时代教育评价改革总体方案”强调要优化结果评价、探索增值评价。

3. 重科研，轻教学。现有的世界大学排名基本上是以科研为导向，科研成果的权重是最大的。其中，ESI 和 U. S. News 只有学术发表一个指标，科研权重占比 100%；ARWU（世界大学学术排行榜）有 5 个指标，与人才培养有关的指标只有一个（占比 10%），科研权重占比 90%；THE 和 QS，科研权重占比超过。这背后的原因有两个：一是科研成果较人才培养更为显性，容易从量化的数据当中进行排名；二是科研成果的数据获取与处理较为容易，而人才培养指标缺乏国际通用的衡量单位。在这些世界大学排名的导向下，中国“双一流”大学热衷追逐学术 GDP，出现严重的“重科研、轻教学”现象。当前中国“双一流”大学的一个突出问题是，科研与教学不能有效整合共同支撑人才培养，严重背离大学立德树人的根本任务。2018 年 5 月 2 日，习近平总书记在北京大学发表重要讲话，明确提出“要把立德树人的成效作为检验学校一切工作的根本标准”，“只有培养出一流人才的大学才是世界一流大学”。

4. 重自然科学，轻人文社科。现有的世界大学排名几乎都是基于科睿唯安的 Web of Science 和爱思唯尔的 Scopus 数据库。其中，ESI 根据 SCI、SSCI 论文的分布，将学科分为 22 个。在这 22 个学科中，自然科学占 20.5 个学科，人文社科仅占 1.5 个学科。在自然科学中，医学和生命科学学科约占一半以上。以其数据的客观性受到中国一流大学的普遍追捧，以至于在中国高教界衍生出一系列似是而非的“办学理念”，如“没有医学院的大学不是世界一流大学”，“ESI 是评价世界一流大学的可靠标准”。在这种错误导向下，中国大学乃至全球大学中出现了一个令人忧思的现象，即大学开始裁撤人文社会学科，因为人文社科对

大学排名贡献不大。在我国 140 所“双一流”建设高校中，有 40 所大学表示近期要新建医学院。大学人文社会学科的萎缩带来两个严重的后果，一个是大学的道德领导能力大大弱化，大学已经不能站在道德高度引领社会、批判社会；另一个是大学立德树人能力大大弱化，目前全球大学面临的一个共同问题是德育普遍边缘化，大学的组织目标、工作方式、评价体系均不支撑大学德育。

5. 近期“双一流”建设高校自评结果表明，中国已经不缺乏指标意义上的世界一流大学。根据国际可比指标，许多大学惊讶地发现自己的数据居然在很多方面已经达到甚至超过国外对标大学，欣喜之余纷纷宣布自己已经是世界一流大学。但社会对此并不认可，这是因为指标意义上的世界一流大学不是真正的世界一流大学，真正的世界一流大学是文化意义上的。所谓文化意义上的世界一流大学，就是一所大学的理念和制度能够给其他国家、其他文明带来实质性的启示，为其他国家的大学所认同、所模仿。这就要求我们的大学评价必须超越数量指标，注重大学的精神与灵魂、文化与内涵。

新时期教育评价改革的现实诉求：基于《总体方案》的解读

施晓光

《深化新时期教育评价改革总体方案》政策寓意深远、诉求明确。《总体方案》堪称新时期落实全国教育大会精神，深化教育评价体制改革的最重要的指导性文件，是对习近平总书记在“全国教育大会”上提出“破五唯”顽疾，“从根本上解决教育评价指挥棒问题”指示的具体落实，对我国未来教育事业改革与发展，尤其是编制好“十四五”规划，具有重要的理论指导价值和极强的实践引领作用。

1. 《总体方案》充分肯定了教育评价的判断、引领、指导、管理和改进功能，将教育评价改革置于“事关教育发展方向”前所未有的战略高度加以明确。2018 年，习近平总书记在“全国教育大会”上系统阐述了教育“九个坚持”的重要论述，创造性地提出新时代我国教育改革与发展的新理念、新基准和新路径。此次，审议通过《总体方案》的目的就是要确保我国教育改革与发展的方向不偏离“九个坚持”的轨道，使之同时满足中国特色社会主义现代化建设的社会外在需求与学生个体成长的内在需要两种不同的现实诉求，最终实现“社会本位”与“个人本位”两种教育价值的有机结合。

2. 《总体方案》充分地吸收了中外教育评价理论和实践研究的最新成果，提出教育评价要“遵循教育评价规律（内在逻辑），……以及不同类型教育特点”的指导思想和行动准则。这表明：①整个《总体方案》制定的科学依据是充分的，符合“有价值的事物不一定能被测量（评价），而被测量（评价）的事物不一定是有价值的”原理；②其实践基础是丰富的，充分考虑了教育系统的复杂性、学校教育的特殊性、学生个体发展的差异性，力争通过建立科学教育评价体制和机制，实现“管教评”分离，达成“以评促建”，满足教育不同“利益相关者”的不同诉求，实现教育公平的质量性发展。

3. 《总体方案》问题意识突出，着眼解决实际问题，具有极强的对象指向性。目前，中外教育评价理论流派林立，划分方法形式多样。但本次《总体方案》并没有严格地受限于“教科书”上某一理论和方法的局限，而是立足解决教育及评价领域存在的具体问题，对象目标明确，有针对性地提出了“改进结果评价”“强化过程评价”“探索增值评价”和“健全综合评价”24 字工作方针。其目的就是采用“刮骨疗法”，止损“五唯”顽疾。在历史上，“学历文凭”“学术论文”“测验考试”“选拔升学”以及“荣誉称号”等要素指标都在

我国教育管理中起到过不可替代的作用。因此,把这些要素作为考核指标本身并无大错,问题出在这些要素指标被慢慢异化成“唯一”,背离其设立和使用的初衷。近20余年,“五唯”问题愈加严重,积重难返,成为教育改革与发展道路上的“绊脚石”和“拦路虎”。尤其是在“中考”和“高考”考核中,整个社会、学校和家庭多年来早已习惯“唯分数论优劣”“唯升学论质量”。“智育至上”名副其实;追求排名大行其道;英雄学霸备受推崇;在教师管理、职称评定、业绩考核等方面,“唯文凭论水平”“唯论文论成果”“唯帽子论英雄”的现象屡禁不止,考核过度量化靡然成风……。这些问题对教育事业良性发展造成难以估量损害。应该说,“五唯”标准助长了人们的功利主义思想,是扭曲教育观和人才观在教育领域中的反映。“五唯”不破,改革成功无望;顽疾不除,卓越理想难成。当然,破“五唯”并不是全盘否定指标性考核和量化考核的作用,摒弃“五唯”,变教育评价标准的“唯一性”为“多元化”才是目的,乃为教育评价改革任务之重中之重。

《总体方案》中提出的“四种评价”模式具有极强的创新性。四种评价模式由四组表示程度不同的“动词”加以限定或修饰,体现了《总体方案》的“亮点”和创新。他们既规定了我国教育评价体系和制度的评价内容和范围,也明确了未来教育评价的工作重点和任务;既有对过去教育评价工作成果的肯定与反思,也有对未来教育评价政策实施效果的展望与期冀;极大地丰富了我国教育评价理论研究,为形成具有中国特色教育评价体系理论框架奠定了基础。①《总体方案》提出“改进”结果评价,实际是部分地肯定了这种评价方式,但同时也不掩盖其存在的局限。众所周知,结果评价是一种以结果为导向的评价模式,作用在于测量教育目标和任务在多大程度上可以实现或被完成。从技术角度看,结果评价本身没有问题,但如果结果评价目标和任务的预设的不合理、不科学,那就必然导致结果的“减值”,甚至“负值”。如果像传统教育评价那样将获得知识的多少作为检验学生学业发展的主要目标和依据,就必然会导致学校教育出现“重智育”“轻德育”“弱体育”“虚美育”“无劳育”的应试教育局面和学生培养中出现“高分低能”“只专不红”严重片面发展的缺陷。②提出“强化”过程评价,既肯定了这种评价方式的重要性和可行性,同时也暗喻其实践价值和执行程度远没有达到理想预期。③提出“探索”增值评价,虽然“增值评价”是国际上非常流行的一种评价模式,且逐渐在我国兴起,但增值评价作为舶来品,毕竟是一种新生事物,本土化移植的方式和成效仍然需要时间检验和不断探究。④提出“健全”综合评价,与单一性评价方式相比,综合评价的优势不言而喻,但仍存在制度缺陷和文化障碍,需不断加以完善。

总之,《总体方案》是一项意义重大,影响深远的好政策。今后的教育评价工作只有很好地学习和贯彻《总体方案》精神,践行“四种评价”方式,实现《总体方案》中提出的“建立科学的、符合时代要求的教育评价制度和机制”的改革目标才曙光在前,翘首可待。

【马陆亭,国家教育发展研究中心副主任、研究员,中国高等教育学会第四届学术委员会委员;王小梅,中国高等教育学会副秘书长,《中国高教研究》主编、编审;刘复兴,中国人民大学教育学院院长、教授;周光礼,中国人民大学评价研究中心执行主任、教授;施晓光,北京大学教育学院教授。】

——来源:《中国高教研究》,2020年第11期

新时代深化教育评价改革的根本遵循

翟博

立德树人是教育的根本任务，是贯彻落实党的教育方针的根本出发点和落脚点。科学的教育评价体系的建立和完善，是落实教育立德树人根本任务、落实党的教育方针的根本保障。没有或缺乏科学的教育评价制度和体系，教育工作的开展、教育改革的发展就会偏离方向，甚至迷失方向，更为严重的是还可能与教育的目标、方向、任务背道而驰，甚至危害教育的改革发展，危害学生的健康成长，影响党的教育方针的贯彻落实。

习近平总书记在全国教育大会上的重要讲话中指出，要深化教育体制改革，健全立德树人落实机制，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，从根本上解决教育评价指挥棒问题。习近平总书记的重要讲话，为我们深化教育体制改革、健全立德树人落实机制、建立科学的教育评价制度和机制指明了方向。中共中央、国务院最近印发的《深化新时代教育评价改革总体方案》（以下简称《总体方案》），指明了教育评价的方向，是落实全国教育大会精神任务、指导教育评价改革的纲领性文件，是新时代深化教育评价改革的根本遵循，对于我们全面贯彻党的教育方针，坚持社会主义办学方向，落实立德树人的根本任务，全面深化教育改革，系统推进教育评价改革，努力培养担当民族复兴大任的时代新人，培养德智体美劳全面发展的社会主义建设者和接班人，对于引导全党全社会树立科学的教育发展观、人才成长观、选人用人观，破除“五唯”顽瘴痼疾，扭转不科学的教育评价导向，建立科学的、符合时代要求的教育评价制度和机制，具有重要的指导意义。

教育评价事关党的教育方针全面贯彻落实，事关教育发展方向。“五唯”是当前教育评价问题的集中体现，反映了我国教育评价中不科学的评价导向。“唯分数”评价学生，将分数作为评价学生的唯一标准，分数第一，忽视了学生思想品德、身心健康、能力素质等发展成长的重要因素，不利于学生德智体美劳全面发展，损害了学生的身心健康，脱离了教育的育人本质。“唯升学”评价学校，片面追求升学率，把升学率作为评价学校的唯一标准、教育政绩的具体体现，违背了教育规律，不利于营造健康的教育生态。“唯文凭”评价人才，忽视了人的品性、能力和素质，不利于鼓励学生多样化成长、成才。“唯论文”评价教师，把论文作为教师职称评定、发展晋升的核心指标，忽视了教师教书育人的本质。“唯帽子”评价学科，把“帽子”作为评价人才和学科的重要指标，忽视了学科建设的本质、职责、使命和作用，不利于推进人才称号和学科建设回归学术性的本质。

学习贯彻落实好《总体方案》精神，必须破立并举，提高改革实效，把握和做好以下工作：

第一，全面贯彻党的教育方针，落实教育立德树人根本任务。全面贯彻党的教育方针，坚持社会主义办学方向，落实立德树人根本任务，核心是培养德智体美劳全面发展的人才。贯彻落实《总体方案》精神，必须完善立德树人体制机制，立足我国基本国情，遵循教育规律，坚持改革创新，以凝聚人心、完善人格、开发人力、培育人才、造福人民为工作目标，培养德智体美劳全面发展的社会主义建设者和接班人，加快推进教育现代化、建设教育强国、办好人民满意的教育。

第二，转变教育思想理念，树立科学的教育评价观。科学的教育评价观，必须坚持以育人为本，以人的全面发展、人的健康成长为根本，立足教育实践、面向未来发展，建立科学、

有效的教育评价机制。引导确立科学的育人目标，确保教育正确发展方向，改变单一化、工具化、功利化的教育评价方式，改变传统教育评价过分注重选拔与甄别的倾向，使教育评价回归教育的本真，更多地体现人的全面发展、人的健康成长、人才的培育培养，从注重“知识评价”向“能力素养评价”拓展、从注重“分数结果评价”向“全过程评价”转变、从注重“单向度评价”向“多向度多元综合评价”提升，推动多维度、全过程、立体式、综合性评价。

第三，深入推进教育评价改革，杜绝不科学的教育评价导向。对于教育评价改革，《总体方案》列出了重点任务：改进结果评价，强化过程评价，探索增值评价，健全综合评价。这也是深化教育体制改革以来首次出现这样的提法。传统教育评价以结果为唯一依据，强调结果在整个评价中的关键作用，这就导致了偏重分数、轻视过程的严重后果。这正是《总体方案》提出改进结果评价、强化过程评价的出发点。增值评价关注学生进步的增幅，强调的是学生较之原有学业的增值分，拓展了教育教学评价的渠道。综合评价是针对学生、教师或学校建立测评指标体系，对评价对象进行系统、全面和完整的综合评判。遵循教育规律，就是要针对不同主体和不同学段、不同类型教育特点，改进结果评价，强化过程评价，探索增值评价，健全综合评价。

第四，树立科学的人才评价观，创新人才培养模式。树立科学成才观念，扭转“唯分数、唯升学”的单一评价倾向，以激发多出人才、出好人才为评价导向。面向每一个孩子，建立多元人才评价标准，完善综合素质评价，注重过程评价，综合运用考试测验、表现性评价等多样化评价方式。创新人才培养方式，引导学生全面发展、全人格成长、全要素健康成长。创新人才培养模式，充分尊重学生个性，张扬学生特色特长。完善高校科研评价体系，鼓励重大原创性科研成果。

第五，建立科学的教育评价体系、科学的教育评价标准。坚持把立德树人成效作为根本标准。加快完善各级各类学校评价标准，着力破除唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾，建立科学的、符合时代要求的教育评价制度和机制。在学生评价方面，扭转“唯分数、唯升学”的单一评价倾向。在教师评价方面，扭转“唯文凭、唯论文、唯帽子”的单一评价倾向。建立科学的教育评价标准和评价指标体系，深化教师评价制度改革，完善职称评价标准，建立符合中小学教师岗位特点的考核评价指标体系，明确不得简单用升学率、学生考试成绩等评价教师，加快形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系。

【翟博，中国教育报刊社党委书记、社长。】

——来源：《中国教育报》，2020年10月29日

建设中国特色世界一流大学的重要抓手

靳诺

教育评价作为教育发展的指挥棒，贯穿学校评价、教师评价、学生评价、人才就业等关键领域，事关教育发展方向。中央全面深化改革委员会审议通过，近日，中共中央、国务院印发的《深化新时代教育评价改革总体方案》，分析了现有教育评价体系存在的突出问题，阐明了教育评价体系改革的指导思想、主要原则和目标任务，对新时代要构建什么样的教育评价体系进行了顶层设计，作出明确回答。《总体方案》是破除“五唯”，建立科学的、符合时代要求的教育评价制度和机制的行动纲领和施工蓝图，也是加快建设中国特色、世界一流大学的重要抓手。学习《总体方案》，有这样几点值得特别关注：

第一，把握教育的政治属性，在教育评价上落实党对教育的全面领导。中国共产党的领导是中国特色社会主义最本质的特征，是中国特色社会主义制度的最大优势。加强党的全面领导是我国教育事业健康发展的根本保证。《总体方案》特别强调要全面贯彻党的教育方针，坚持社会主义办学方向。习近平总书记在全国教育大会上提出教育是国之大计、党之大计。党怎么管教育，怎么抓教育？教育评价是非常重要的抓手。《总体方案》全面贯彻落实习近平新时代中国特色社会主义思想，明确了党对教育工作领导的体制机制和评价办法，强调各级党委要认真落实破“五唯”的领导责任。中国人民大学党委班子形成共识：破“五唯”必须紧紧依靠党对教育工作的全面领导，必须强化党委的责任意识。学校党委将积极贯彻落实中央破“五唯”的政策要求，将其上升到事关大学办学方向的政治高度，切实加强统筹协调，通过宣传引导、督促落实，进一步营造教育评价改革良好氛围。

第二，回归教育的本质功能，在教育评价上突出立德树人。从字面上理解，“立德树人”即为立德和树人。“立德”是指树立德业，“树人”是指培养人才。从基本内涵看，“立德树人”至少包括4个方面的内容，即有德行、有才学、有根基、有格局。建设中国特色、世界一流大学必须从整体上把握“德才兼备”这一人才质量与规格。2018年5月2日，习近平总书记在北京大学师生座谈会上明确提出，要把立德树人的成效作为检验学校一切工作的根本标准，真正做到以文化人、以德育人，不断提高学生思想水平、政治觉悟、道德品质、文化素养。在这一精神的指引下，《总体方案》提出了“改进结果评价，强化过程评价，探索增值评价，健全综合评价”原则。“四个评价”是科学评价观的具体体现，关注学生学习过程，突出评价的激励与促进作用。在“双一流”建设中，中国人民大学坚持把立德树人融入思想道德教育、文化知识教育、社会实践教育各个环节，重视学生的独立思考、个性发展与综合发展，构建多元评价体系，致力培养有才华、有品格、有担当的国民表率、社会栋梁。

第三，适应教育的发展阶段，在教育评价上助力高等教育内涵式发展。实现内涵式发展已成为我国高等教育自身发展的内在需要，也是建设中国特色世界一流大学的必然要求。高等教育内涵式发展强调从“量的扩张”到“质的提高”的转变，是一种使外延的扩展与内涵的丰富相互促进的发展。世界一流大学有两个比较公认的指标，一是拥有世界一流的教学，培养出世界一流的人才；二是拥有世界一流的科学研究，产出世界一流的研究成果。而一流的教学和一流的科研取决于一支世界一流的教师队伍。教师是立教之本、兴教之源。健全高校教师评价体系，引导高校教师回归教书育人的初心，是克服“唯文凭论”“唯论文论”“唯帽子论”的出发点，也是推动高等教育内涵式发展的应有之义。在“双一流”建设中，中国人民大学积极落实中央破“五唯”的要求，扭转不科学的教育评价导向，清理人事

管理制度中“唯”的事项，努力完善教师职称评审工作机制，并逐步建立起一套综合多元的教师评价体系。

第四，强化教育的服务导向，在教育评价上引导大学服务国家战略。发展教育事业必须面向和解决中国问题。习近平总书记指出：“我们中国共产党人干革命、搞建设、抓改革，从来都是为了解决中国的现实问题。”建设中国特色世界一流大学必须扎根中国大地，遵循社会需求逻辑，树立强烈的问题意识。“双一流”建设必须面向世界科技前沿、面向经济主战场、面向国家重大需求、面向人民生命健康。作为在我国人文社会科学领域独树一帜的高校，中国人民大学一直注重发挥自身优势，积极为党和政府决策、为地方经济社会发展服务，成为党和政府重要的“思想库”“智囊团”。2018年起，学校在科研为主型教师岗位序列下，试点设置智库科研岗，并印发《中国人民大学智库科研岗岗位设置和聘用管理实施细则》，将智库成果作为教师评价的重要方面，允许教师在职称评审和岗位聘用工作中将智库成果提交学术委员会审议，提高了全校教师服务国家和社会的积极性。

【靳诺，中国人民大学党委书记。】

——来源：《中国教育报》，2020年10月30日

教育评价改革，关键在落实

顾明远

中共中央、国务院发布了《深化新时代教育评价改革总体方案》(以下简称《总体方案》),这是贯彻落实习近平总书记关于教育的重要论述和全国教育大会精神、深化教育综合改革的重大举措,对新时代中国特色社会主义教育的发展、实现中国教育现代化具有十分重要的指导作用。

有什么样的评价指挥棒,就有什么样的办学导向。习近平总书记指出,要深化教育体制改革,健全立德树人落实机制,扭转不科学的教育评价导向。改革教育评价制度,事关培养什么人、怎样培养人、为谁培养人的根本问题,事关年轻一代的成长,事关国家之大计、党之大计。为此,《总体方案》明确提出,以习近平新时代中国特色社会主义思想为指导,全面贯彻党的教育方针,坚持社会主义办学方向,落实立德树人根本任务,遵循教育规律,系统推进教育评价改革,发展素质教育,引导全党全社会树立科学的教育发展观、人才成长观、选人用人观,破除“五唯”顽瘴痼疾,培养德智体美劳全面发展的社会主义建设者和接班人。

《总体方案》对教育评价制度和机制的改革做了顶层设计。最重要一点是,突出加强党的领导。习近平总书记讲,党政军民学,东西南北中,党是领导一切的。《总体方案》将“改革党委和政府教育工作评价,推进科学履行职责”作为重点任务,提出完善党对教育工作全面领导的体制机制,强调各级党委要牢固树立科学的教育发展理念,坚决克服短视行为、功利化倾向。在“组织实施”部分再次强调各级党委和政府要加强对教育评价改革的组织领导。这有利于保证改革的正确方向,保证社会主义的办学方向和落实立德树人的根本任务。

《总体方案》具有鲜明的方向性、导向性、针对性、系统性、可操作性。文件强调立德树人,强调树立科学成才观念。坚持以德为先、能力为重、全面发展,坚持面向人人、因材施教、知行合一。着眼引导学生立志听党话、跟党走,立志扎根人民、奉献国家。文件针对不同对象、各个学段教育特点,改进结果评价,强化过程评价,探索增值评价,健全综合评价,着力破除唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾,建立科学的、符合时代要求的教育评价制度和机制。

《总体方案》对各级各类学校、教育教学工作,对学生、教师的评价都提出了不同要求。坚持把立德树人成效作为根本标准,不得下达升学指标,坚决改变简单以考分排名评价老师、以考试成绩评价学生、以升学率评价学校的导向和做法,促进学生身心健康、全面发展。突出教师开展教育教学的实绩,引导教师上好每一节课,关爱每一个学生,减轻学生的课业负担。稳步推进中高考改革,构建引导学生德智体美劳全面发展的考试内容体系,增强试题开放性。推进高等学校分类评价,引导不同类型高校科学定位,办出特色和水平,培养一流人才、产出一流成果、主动服务国家需求。

《总体方案》强调,全社会都要参与教育评价制度的改革。习近平总书记在全国教育大会上指出,办好教育事业,家庭、学校、政府、社会都有责任。因此,教育评价制度的改革不仅学校要改革,家庭教育也要改革,要树立正确的人才观、教学观,重视人格的培养。同时政府、学校、家庭、社会多元参与评价体系,营造良好的教育氛围。《总体方案》一个很值得称道的亮点是,强调社会用人评价制度的改革,通过扭转“唯名校”“唯学历”的用人导向,促进平等就业、树立教育自信、服务全民终身学习,保障每个学生都有人生出彩的机会。

《总体方案》对教育评价制度改革的要求和举措都非常具体，关键在于落实。我们要深刻领会《总体方案》的精神和要求，全社会都要树立正确的教育观、人才观、教学观、用人观，全面正确地认识教育评价的性质和功能，不断探索，认真落实，使教育评价促进教育的发展，早日实现教育现代化，办好人民满意的教育。

【顾明远，国家教育咨询委员会委员、中国教育学会名誉会长。】

——来源：《人民日报》，2020年10月21日

深化改革，立好教育评价指挥棒

教育评价事关教育发展方向，有什么样的评价指挥棒，就有什么样的办学导向。不久前，中共中央、国务院印发《深化新时代教育评价改革总体方案》。作为新中国第一个关于教育评价系统改革的文件，总体方案的出台实施，对全面贯彻党的教育方针，完善立德树人体制机制，引导全党全社会树立科学的教育发展观、人才成长观、选人用人观，具有重大而深远的意义。

当前，我国教育改革正在围绕教育发展的重大问题和群众关心的热点问题向纵深推进，教育评价改革正是其中重要一环。教育评价“评什么”，在一定程度上决定了老师“教什么”、学生“学什么”、社会“用什么”。教育评价改革总体方案针对学校评价，提出坚持把立德树人成效作为根本标准；针对教师评价，提出把认真履行教育教学职责作为评价教师的基本要求；针对学生评价，提出促进德智体美劳全面发展；同时还明确了用人评价，建立以品德和能力为导向、以岗位需求为目标的人才使用机制，共同营造教育发展良好环境。建立完善多维度的综合评价体系，努力实现教育评价的科学化、制度化、规范化，才能让教育评价的指挥棒树得稳、用得好，保证教育的正确发展方向。

教育评价的指导性和可操作性，在于教育评价标准对教育规律和时代要求的高度契合。为克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾，提高教育治理能力和水平，教育评价改革总体方案坚持破立结合、多措并举。比如，学校评价方面，破除重分数轻素质等片面办学行为，确立立德树人落实机制；教师评价方面，破除重科研轻教学、重教书轻育人等行为，确立潜心教学、全心育人的制度要求；学生评价方面，破除以分数给学生贴标签的不科学做法，确立德智体美劳全面发展的育人要求。教育评价机制改革之所以吸引众多关注目光，原因就在于针对不同主体和不同学段、不同类型的教育特点和教育规律，改进结果评价，强化过程评价，探索增值评价，健全综合评价，努力办好让人民满意的教育。

解决好教育评价指挥棒问题，是全面充分激发教育事业生机活力的关键一步。近年来，各级各地在教育理念更新、教育评价变革上不断探索。有地方试行中小学生学习质量“绿色指标”，结合本地实际，建立起一套替代单一分数排名的评价体系；有地方注重用多把“尺子”丈量学生才华，鼓励多元发展；有地方使用“教育评价应用云平台”，对课堂教学进行智能分析，使其成为改进日常教育教学的新抓手……这些改革尝试，对促使教育教学回归到促进学生身心健康、全面发展上来有很大帮助。下一步，我们必须拿出攻坚克难的勇气、久久为功的韧劲，确保教育评价改革任务落实到位。

教育是国之大计、党之大计。把教育评价改革这样一项世界性、历史性、实践性难题破解好，必须加强党对教育工作的全面领导，不断提高各级党委和政府科学履行职责水平。按照改革方案，到2035年，我国基本形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系，这必将为加快推进教育现代化、建成教育强国打下坚实基础。

——来源：《人民日报》，2020年11月3日

打好教育领域综合改革“关键一役”

近日，中共中央、国务院印发《深化新时代教育评价改革总体方案》。这是新中国第一个关于教育评价系统改革的文件，也是指导深化新时代教育评价改革的纲领性文件。方案的发布对于全面贯彻党的教育方针，完善立德树人体制机制，破除“五唯”顽瘴痼疾、引导全党全社会树立科学的教育发展观、人才成长观、选人用人观具有重大意义，也标志着我国向建设富有时代特征、彰显中国特色、体现世界水平的教育评价体系，迈出了关键一步。

当今世界正经历百年未有之大变局，知识创新速度加快，科技变革升级，教育与人才竞争日趋激烈，建立一套科学的、符合时代要求的教育评价制度和机制，对于培养担当民族复兴大任的时代新人，对于进一步释放人才创新活力、提升原始创新能力，对于适应当前国内外环境发生的深刻复杂变化、适应我国“十四五”时期以及更长时期的发展对教育发展、科技创新提出的更为迫切的要求，都至关重要。因此，作为全面深化教育领域综合改革的重要环节，深化新时代教育评价改革，既是国家所需，也是时代所需。

《总体方案》坚持问题导向，从党中央关心、群众关切、社会关注的问题入手，注重改进结果评价、强化过程评价、探索增值评价、健全综合评价，破立并举，求真务实，力求推进教育评价关键领域改革取得实质性突破。其中，针对党委和政府教育工作评价，“破”的是短视行为和功利化倾向，“立”的是科学履行职责的体制机制；针对学校评价，“破”的是重智育轻德育、重分数轻素质等片面办学行为，“立”的是立德树人落实机制；针对教师评价，“破”的是重科研轻教学、重教书轻育人等现象，“立”的是潜心教学、全心育人的制度要求；针对学生评价，“破”的是用分数给学生贴标签的不科学做法，“立”的是德智体美劳全面发展的育人要求；针对用人评价，“破”的是文凭学历至上等不合理用人观，“立”的是以品德和能力为导向的人才使用机制。

特别是针对长期以来困扰教育发展的重点、难点问题，《总体方案》鲜明亮剑，明确提出：不得下达升学指标或以中高考升学率考核下一级党委和政府、教育部门、学校和教师，不得将升学率与学校工程项目、经费分配、评优评先等挂钩，不得通过任何形式以中高考成绩为标准奖励教师和学生；不得将论文数、项目数、课题经费等科研量化指标与绩效工资分配、奖励挂钩；不得把人才称号作为承担科研项目、职称评聘、评优评奖、学位点申报的限制性条件；不得通过设置奖金等方式违规争抢生源；不得将毕业院校、国（境）外学习经历、学习方式作为限制性条件……聚焦热点焦点问题、剑指沉疴旧疾，彰显了抽薪止沸、治疾于本的智慧与魄力。

当然，教育评价改革是一项世界性、历史性、实践性难题，涉及历史文化传统、经济社会发展水平、思想观念等多重因素，涉及不同主体，牵一发而动全身，被喻为教育综合改革“关键一役”和“最硬一仗”，须以攻坚克难的勇气、久久为功的韧劲，系统设计、辨证施治、重点突破，更须因地制宜、协同推进、坚持不懈。

一方面，要落实改革责任，各级改革主体要从理念上加强理解，深刻认识到教育评价改革的重要性和紧迫性，结合本地、本校实际，明确落实举措、落细落实；一方面，要加强专业化建设，既要打造专业评价体系，也要严格控制教育评价活动数量和频次、创新评价工具、完善评价结果运用，提升效率，切实减轻基层和学校负担；此外，还要尽快增进社会共识，在全社会形成科学的选人用人理念，树立正确的教育观和成才观，为改革提供良好的氛围与土壤。

志不求易，事不避难。作为指导深化新时代教育评价改革的纲领性文件，《总体方案》既大力破除不科学、不合理的教育评价做法和导向，又着力建立科学的、符合时代要求的教育评价制度和机制。在教育领域综合改革全面深化，教育的系统性、整体性和协同性亟待增强之际，《总体方案》的出台实施，必将进一步深化育人方式、办学模式、管理体制、保障机制改革，为推进教育现代化、建设教育强国、办好人民满意的教育，注入强大的基础性和牵引性力量。

——来源：《人民日报》，2020年10月14日

构建合实际、高水平的教育评价体系

教育评价事关教育发展方向，有什么样的评价指挥棒，就有什么样的办学导向。近日，中共中央、国务院印发了《深化新时代教育评价改革总体方案》（以下简称《总体方案》）。作为新中国第一个关于教育评价系统改革的文件，《总体方案》明确了新时代我国教育评价改革的任务书和路线图，对于全面贯彻党的教育方针，完善立德树人体制机制，破除“五唯”顽瘴痼疾，培养担当民族复兴大任的时代新人，培养德智体美劳全面发展的社会主义建设者和接班人具有重大意义。

新时代新形势，改革开放和社会主义现代化建设、促进人的全面发展和社会全面进步对教育和学习提出了新的更高的要求。加快推进教育现代化、建设教育强国、办好人民满意的教育，科学的教育评价体系至关重要。党的十八大以来，习近平总书记就深化教育评价改革作出一系列重要指示批示。在全国教育大会上，习近平总书记指出，要深化教育体制改革，健全立德树人落实机制，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，从根本上解决教育评价指挥棒问题。在教育文化卫生体育领域专家代表座谈会上，习近平总书记强调，要抓好深化新时代教育评价改革总体方案出台和落实落地，构建符合中国实际、具有世界水平的评价体系。习近平总书记的重要指示批示为深化新时代教育评价改革指明了前进方向、提供了根本遵循。

值得注意的是，教育评价改革是一项世界性难题，涉及多重因素，影响广泛，牵一发而动全身。近年来，中高考改革逐步推进、职称评价标准进一步完善，涉及教育评价方面的改革举措不断。但要真正扭转评价“指挥棒”导向，更需要从上到下的系统设计、辨证施治、重点突破。

“育才造士，为国之本”，教育评价改革要坚持以立德树人为主线。长久以来，我国各级各类教育中还或多或少地存在有重智育轻德育、重分数轻素质等片面办学行为，在某些地方，升学率成为考核指标，甚至与经费分配、评优评先等挂钩，严重影响了学生的身心健康、全面发展。此次《总体方案》着眼于全面贯彻党的教育方针，明确把落实立德树人根本任务、培养德智体美劳全面发展的社会主义建设者和接班人作为主线贯穿于各项改革任务始终，强调要坚决克服教育中的短视行为、功利化倾向。但在这个过程中，还要充分考虑到各级各类教育的不同特点，制定符合学段教育特征、学生成长特点的评价标准和内容，真正将立德树人落在实处。

哪里是痛点、难点，哪里就是教育改革的重点，教育评价改革要坚持以破“五唯”为导向。针对不同主体和不同学段、不同类型教育特点，改进结果评价，强化过程评价，探索增值评价，健全综合评价，着力破除唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾，建立科学的、符合时代要求的教育评价制度和机制。《总体方案》强调破立结合，既要大力破除不科学的评价，又要着力建立科学的、符合时代要求的教育评价制度和机制，提出的改进结果评价、强化过程评价、探索增值评价、健全综合评价等举措为改革提供了指导方针。但在具体实践中，各级教育主管部门、学校、教师还要结合工作实践提出具体办法，形成完善的规章制度与可操作的实施细则。

非知之艰，行之惟艰。《总体方案》提出，到2035年，基本形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系。要实现这样的目标愿景，既需要教育系统的守正创新，敢于“啃硬骨头”，又需要社会与文化的系统性变革。只有这样，才能让教育评价回归教育的本质、规律和初心，为新时代中国特色社会主义事业提供源源不断的人才支持。

——来源：《光明日报》，2020年10月14日

指引新时代教育评价改革的纲领性文件

近日，中共中央、国务院印发《深化新时代教育评价改革总体方案》。《总体方案》就完善立德树人体制机制，扭转不科学的教育评价导向，克服“五唯”顽瘴痼疾，提出了5个方面22项改革任务，目标是到2035年，基本形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系。《总体方案》坚持以立德树人为主线，以破“五唯”为导向，以党委和政府、学校、教师、学生、社会五类主体为抓手，系统、整体、协同谋划，为教育评价改革明确了方向，是当下和未来教育评价改革的路标和航向。

《总体方案》的出台具有重要的时代意义，它催生于时代，服务于时代，是新中国第一个关于教育评价的系统性改革文件，也是指导深化新时代教育评价改革的纲领性文件。教育评价改革是一项世界性、历史性、实践性的难题，是教育改革发展过程中的“龙头之战”，是最难啃的“硬骨头”。《总体方案》的出台是落实习近平总书记关于教育的重要论述和全国教育大会精神要求的生动实践，对引导全社会树立科学的教育发展观、人才成长观、选人用人观，具有重大的现实意义和重要的指导意义，必将有利于推动解决当前教育评价领域存在的一系列问题，培养担当民族复兴大任的时代新人，培养德智体美劳全面发展的社会主义建设者和接班人。

贯彻落实《总体方案》是当前教育工作的重要聚焦点，要把破“五唯”作为当前教育评价改革的重心。《总体方案》坚持问题导向、“破”“立”结合，既为扭转不科学的教育评价导向擘画了蓝图，也为破解教育评价中的现实难题提出了具体要求，极具宏观性和针对性。深化教育评价改革，要全面贯彻党的教育方针，立足全局，统筹兼顾，针对不同主体、不同教育领域和不同学段的特点，分层分类施策。要通过评价改革，推动党委和政府科学履行职责，推动学校扎实落实立德树人根本任务，推进教师践行立德树人使命，促进学生德智体美劳全面发展，营造良好的教育发展环境。

贯彻落实《总体方案》，要为问题的解决寻找到符合时代要求的路径。《总体方案》既是一个目标性愿景，又是施工图和路线图。《总体方案》分两个阶段提出了深化新时代教育评价改革的目标，充分考虑到教育评价改革的艰巨性和长期性，将引领教育评价在改革中完善和发展，并最终形成有中国特色、世界水平的教育评价体系。在教育评价改革的过程中，要以现实问题为导向，在改革的“破”“立”之中，在螺旋上升之中不断修正目标，调整方法和策略，建立科学的、符合时代要求的教育评价制度和机制。

贯彻落实《总体方案》，需要全社会合力推进，做好保障工作。深化教育评价改革，是一项复杂的系统工程，实现改革目标、落实改革任务，是包括党政机关、各级各类学校、用人单位、新闻媒体、教师、家长等在内的各个主体共同的责任，只有各方通力配合、协同推进才能确保改革取得实效。要夯实各级党委和政府落实改革措施的责任，树立正确的教育政绩观、科学履职。要凝聚起全社会的共识、夯实群众基础，为深化教育评价改革营造良好的社会氛围。

我们要扎实学习《总体方案》的精神，理解贯彻落实《总体方案》的紧迫性和严肃性。拿出攻坚克难的勇气，以久久为功的韧劲，对照任务清单，积极稳妥推进，以科学理性的精神和扎实的行动推动改革任务在教育改革实践中圆满完成。让我们以习近平新时代中国特色社会主义思想为指导，坚持立德树人，牢记为党育人、为国育才使命，充分发挥教育评价的“指挥棒”作用，全面深化教育领域综合改革，加快推进教育现代化、建设教育强国、办好人民满意的教育，为实现中华民族伟大复兴的中国梦而努力奋斗。

——来源：《中国教育报》，2020年10月14日

分类评价：指标设计、操作程式和结果应用 ——以上海高校分类评价为例

长期以来，尤其是高校扩招以来，高校发展同质化的问题一直困扰着我国高等教育的改革和发展。

为了扭转这种不良倾向，自1993年起，《中国教育改革和发展纲要》《国家中长期教育改革和发展规划纲要（2010-2020年）》《国家教育事业“十三五”规划》《关于深化教育体制机制改革的意见》等国家重要文件都提出了建立高校分类体系、实施高校分类管理、促进高校特色发展的目标和要求。但从高等教育改革和发展的实践来看，由于评价标准单一，高校发展同质化倾向一直没有得到有效遏制。要打破这种局面，必须按照习近平总书记在全国教育大会讲话所强调的，扭转不科学的教育评价导向，从根本上解决教育评价指挥棒问题。

一、分类评价：源起与发展

高校分类理念和政策的提出，最早见于1993年中共中央国务院印发的《中国教育改革和发展纲要》，之后随着“211工程”“985工程”和“双一流建设工程”的实施而不断推进和完善。

（一）高校分类评价改革之肇始

1985年中共中央《关于教育体制改革的决定》首次明确提出，“教育行政部门要组织教育界、知识界和用人单位定期对高等学校的办学水平进行评估。”1990年，原国家教委颁布了《普通高等学校教育评估暂行规定》，之后，陆续开展了一系列高等教育评估工作。由于当时我国高等教育整体发展水平普遍不高，学校办学经费困难，办学条件较差，因此，很长一段时期内各类高等教育评估定位的重点是“以评促建”，即通过统一的标准来推动各级各类高校建设，包括基础设施建设、教师队伍建设和学科专业和课程教材建设、教学科研管理制度建设等。这种由政府主导、统一的评价制度，在推动政府加大高等教育投入、改善学校办学条件、推动学校规范教育管理、提高教育质量等方面都起到了重要作用，极大地推动了我国高等教育的发展。但由于评价标准过于单一，对不同起点、不同类型的同层次高校都采取同一个评价标准，忽视高校之间的差异；加之由政府主导的评价往往与政策支持、资源配置挂钩，学校出于趋利避害的现实考虑，为了在各类评价中获得好成绩，纷纷对标这种统一的标准，甚至唯这种统一的标准是从，结果造成了严重的同质化现象，主要表现为：在目标定位上，追求综合化、上层次、上规模；在人才培养模式方面，培养目标、课程体系、培养方式方法上趋同。

高校发展的同质化既不利于高校形成自身的特色和个性，也无法满足社会和个体的多元化需求。鉴于此，1993年中共中央、国务院颁布的《中国教育改革和发展纲要》明确提出，“制订高等学校分类标准和相应的政策措施，使各种类型的学校合理分工，在各自的层次上办出特色。”这是国家层面最早提出高校分类思想的权威性文件。为了落实《纲要》精神，教育部启动了“211工程”建设，之后又启动了“985工程”和“双一流”建设工程。其目的是通过重点建设，使一批高校和重点学科在教育质量、科学研究、管理水平和办学效益等方面有较大提高，解决高层次人才培养、经济建设和社会发展重大关键技术依赖外国的问题，同时发挥这些高校的示范作用，带动、引领其他高校的发展。通过20多年的持续投入和建设，这一举措取得了明显的成效，一批高校进入世界一流大学行列，形成了一批具有世界学术影响力的高水平学科，培养了一大批高质量、高层次的优秀人才，建设了一支具有国际视野和水

准的教师队伍，产出了一批事关经济社会发展关键技术的重大科研成果，中国高校的国际竞争力和影响力显著增强。但在肯定这种重点建设策略取得巨大成效的同时，我们也不得不承认，进入重点建设行列的“211 高校”“985 高校”和“双一流高校”总体上数量不多，总共只有 100 多所，只占 2600 多所高校的左右，这些学校的跨越式发展并不等于中国高等教育整体水平的普遍提升，甚至由于“马太效应”的存在，加剧了高校的两极分化，阻碍了其他高校的发展。

（二）中长期规划纲要的推动

1999 年全国教育工作会议做出高等教育大扩招的重大决策后，我国高等教育规模迅速发展，普通高等教育本专科、在学研究生分别由 1999 年的 413.42 万人、23.35 万人迅速增长到 2009 年 2144.66 的万人、140.49 万人，高等教育毛入学率由迅速提升到 24.2%。在高校扩招如火如荼地进行的同时，对高校的评价仍然是“用一把尺子量所有高校”。如 2002 年起试行的第一轮本科教学水平评估就是以统一的标准来衡量所有本科院校，其单一的指标体系和评价方式受到社会各界的诟病，甚至被认为是导致本科教学水平评估诸多乱象的根源。“这样同一把尺子的评价标准必然抹杀学科之间的差异，这样的评价结果的公平性存疑。可见，这种由政府主导、统一的评价制度，这种资源倾斜式的重点建设策略，不仅没有办法解决高校同质化发展的问题，甚至在高校大扩招后加剧了这种倾向。为了扭转这种不良倾向，《国家中长期教育改革和发展规划纲要（2010-2020 年）》明确提出，“建立高校分类体系，实行分类管理。发挥政策指导和资源配置的作用，引导高校合理定位，克服同质化倾向，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，争创一流。”为推动中长期规划纲要的实施，国务院成立了国家教育体制改革领导小组，各省市也纷纷成立省级教育体制改革领导小组，高校分类管理、分类建设、分类评价提上了高等教育改革和发展的重要议事日程。但总体而言，多数省市仍然沿袭了重点建设的思路，将高校区分为重点建设和非重点建设两类，并推出了对重点建设高校的各类评价，对高校的分类管理并没有做到真正意义上的全覆盖，高校分类管理、分类评价工作推进情况总体上并不尽如人意。

党的十九大报告提出要“加快一流大学和一流学科建设，实现高等教育内涵式发展。”实现高等教育内涵式发展的着力点之一是解决高等教育同质化问题，而解决同质化问题的关键是高校分类管理，高校分类管理的前提又是如何科学地对高校进行分类。作为全国教育改革先行试验区，上海根据人才培养和科学研究的主体功能，将高校分为学术研究型、应用研究型、应用技能型和应用技术型四种类型。根据主干学科专业的数量，将高校分为综合性、多科性和特色性三种类型，由此形成了上海特色的“二维”“十二宫格”高校分类图景，并以《上海高等教育布局结构与发展规划（2015-2030 年）》《上海市高等教育促进条例》等重要政策法规文件加以固化，这为上海推进高校分类评价提供了法理依据。

二、指标设计：多方共同参与博弈

如前所述，上海以教育综合改革为契机，以规划和立法的形式解决了高校科学分类的标准问题。但分类标准的确定并不意味着高校会自然而然地各就各位。如果仍然沿用整齐划一的评价方式来衡量所有的高校，高校内涵发展、特色发展也就无从谈起。要让高校各安其位、特色发展，必须实施分类评价，而开展分类评价的前提是设计针对不同类型高校的评价指标。

（一）确立指标设计的方法和原则

一般而言，评价指标的设计有三种方法可以选择：一是从内涵分析入手，抓住事物的本质属性，然后把这一属性的现象性外观表现确定为指标；二是从分析事物间的相互联系开始，

抓住事物变化后产生的效应，把事物变化所产生的效应确定为指标；三是抓住事物的全部属性或相关属性，把因素群作为相关指标。相比而言，第一种方法最简捷、最有效，第三种方法最不经济。上海在设计高校分类评价指标时主要采用了第一种方法，即先把高等教育的本质属性具体化为高等教育五大职能，先后召开市教卫党委、市教委处室负责人、不同类型高校专家代表相关会议各二十余次，请他们列出与高等教育职能相关、体现不同类型高校特征的指标，并经过多轮次的意见征集、汇总和分析，把一些非本质、非关键、不具操作性的指标剔除后，形成分类评价指标征求意见稿。

在分类评价指标征求意见稿的研制过程中，强调了三个基本原则：首先是突出中国特色，正如习近平总书记在2016年全国高校思想政治工作会议特别强调的，我国独特的历史、文化和国情决定了我国必须走中国特色的高等教育发展道路。什么是中国特色社会主义高等教育道路？习近平总书记在年全国教育大会提出的“九个坚持”是中国特色的最好诠释，其中居于核心和统领地位的是“坚持党对教育事业的全面领导，坚持把立德树人作为根本任务。”因此，我们将党对高校全面领导的落实情况、立德树人根本任务的落实情况等指标列为一票否决的关键核心指标。其次是聚焦高校职能，即中国特色高等教育五大职能——人才培养、科学研究、社会服务、文化传承创新、国际交流合作在指标体系中都要有所体现，不可或缺，而且在权重分配上要合理配置。如果片面地强调体现某一个或某些职能的指标，往往会产生错误的导向，不利于高校职能的全面发挥。再次是体现学科专业差异，通过采取单设“办学特色”指标并赋予其10分的最高权重、设置个性化评价标准、设置权重可选指标等多种策略，引导学校根据自身学科专业基础办出特色、争创一流。比如，根据教育部《普通高等学校基本办学条件指标（试行）》及评估实践确定的优秀标准，将综合类、理工类、医学类、体育类、艺术类等不同类型高校的生师比的满分标准分别设定为16、16、14、11、11，这样既给高校未来发展预留空间，又不至于浪费宝贵的人力资源。再比如，在应用技术型和应用技能型两种类型的评价指标中，“科研项目”这个二级指标的权重总分都是“4”分，但在“师均科研经费”、“师均企事业单位委托项目数”两个三级指标中，分别设置了“2分/2分”、“1分/3分”、“3分/1分”三组权重，供具有不同学科专业优势的高校选择。

（二）广泛征求利益攸关方的意见

分类评价指标征求意见稿形成后，开始广泛征求各方意见，对指标内涵、权重、评价标准进行修改完善。首先是充分听取高校的意见。分别于2017年6月和12月，两轮次在全市所有高校进行了全覆盖式的意见征集。各高校对分类评价指标的研制高度重视，组织专家对学校所在类型的指标体系进行了专门的研究，并经学校党委会、校长办公会审议后反馈给教育主管部门。意见征集后，主管部门根据学术研究、应用研究、应用技术和应用技能四种类型，分门别类地进行了汇总，对高校反馈的意见逐条进行了分析和讨论。之后，又分七批次与二十多个对分类评价指标高度重视、意见比较集中的高校进行了面对面的交流和沟通。其次是听取相关委办局和科研院所的意见和建议，包括市发展改革委员会、市财政局、市人力资源与社会保障局、市编办、市教育卫生工作党委和市教委的意见。其中，市教育卫生工作党委和市教委征求意见十余轮次。其他委办局征求意见两轮次。在将意见汇总分析后，本着“能采尽采”的原则，采纳了其中80%以上的意见和建议，对无法采纳的意见和建议都通过多种形式向各方说明了无法采纳的理由和依据。

经过多部门多轮次广泛征求意见，学术研究、应用研究、应用技术、应用技能四种类型高校的分类评价指标终于出炉。据不完全统计，从指标研制到意见征集再到修改完善和最终

定稿，保守估计有超过人次的各类专家参与其中。如此广覆盖、大批量的专家参与，保证了指标内涵、权重分配、评价标准的信度和效度，为上海高校分类评价奠定了坚实的基础。可以这么说，指标研制的过程既是广集众智、反映各方利益诉求的过程，也是多方博弈、逐渐达成共识的过程。指标设计时的多方共同博弈，特别是高校的多轮参与，为评价结果的广泛认可奠定了扎实的基础。

三、操作程式：自我评价和专家评价相结合

自我评价的过程是被评价对象搜集、整理评价信息、进行自我诊断的过程，有助于评价对象的自我改进，但也存在“报喜不报忧”、夸大成绩、掩盖不足的弊端。专家评价是对自我评价的再检查，有助于提高评价的客观性、科学性和权威性，但也可能存在由于专家掌握信息不够系统、全面、准确等因素，导致评价结果有失偏颇的现象。教育是一个非常复杂的社会现象，教育评价是一项专业性非常强的工作，因此，现代教育评价一般都由自我评价和专家评价两个阶段组成，上海高校分类评价同样如此。

（一）明确指标内涵和具体要求

不论是自我评价还是专家评价，评价者和被评者对指标内涵和相关要素的精准把握和客观公正的态度是确保评价结果信度和效度的前提和关键。为此，我们首先对高校相关人员和评价专家进行了专门的培训和辅导，确保各方对指标内涵、统计要素、时间节点等评价标准的认识达成一致。其次要求被评者和评价者都签订承诺书，学校必须承诺所提供的数据和材料是客观、真实的；评价专家必须承诺依据自身的专业素养做出客观评价，同时实行回避原则，确保评价公平公正。

自我评价是学校为专家评价提供充分信息的过程。在自我评价阶段，学校需要根据评价指标体系和具体工作要求提供自评报告等相关材料。一方面学校要为量化指标提供准确数据，如科研项目数、学生重大赛事获奖数等；另一方面学校要以案例的形式充分展现自身在人才培养、科学研究、社会服务等方面的办学特色。在这个过程中，要求学校保持客观的态度，展示成绩、提炼经验固然重要，但绝对不能粉饰太平、回避短板，甚至弄虚作假。

（二）综合应用多方专家评价的结果

与学科评估、专业认证等专项评价不同的是，上海高校分类评价本质上属于综合评价，涉及高校职能的方方面面，是一项非常复杂的系统工程，其指标内涵丰富，涉及的学科专业领域和具体管理环节众多，仅由一次专家评价是很难完成的。因此，上海高校分类评价中的专家评价实际上是一个由多方专家构成、多轮次的评价过程。一是教育行政部门各职能处室根据职能分工，结合日常管理工作的实际，组织专家对业务性强、职能指向明确的指标进行评价。如高等教育处在日常管理的过程中，经常会以各种形式组织专家对高校教学质量保障体系建设、教学质量监控实施、教学质量反馈与改进、教学质量信息及利用等工作进行评价，对高校教学管理水平和教学质量高低有全面的了解。因此，“教学质量保障水平”这一指标的评价结果就由高等教育处结合历次专家评价给出。二是由教育督导部门组织专家对职能指向不是很明确的综合性指标进行评价，如规划管理水平等。“建立以规划为引领的市级统筹机制”是上海教育改革的成功经验，得到教育部领导的高度肯定。在市级规划的统筹指导下，各高校都制定了学校的发展规划，但在规划管理的过程中往往存在“重编制、轻执行、无评价”的现象，导致规划的执行情况大打折扣。为了扭转这种不良倾向，上海高校分类评价时引入了“规划管理水平”这一指标，对规划的编制质量和执行情况进行评价。由于发展规划涉及到学校教育教学工作的方方面面，必须由对学校情况有全面了解的专家队伍才能做出准

确评价。为此，我们组织了由市内外担任过省级教育行政部门或高校主要领导或分管领导组成的专家组，对学校发展规划的执行情况及成效进行集中评价，并给出评价结果。三是在部分纯客观的量化指标上，直接采信政府行政部门和有较强公信力、影响力的社会第三方评价结果。如“国家级教学成果奖数”指标采信了教育部公示后无异议的结果；“国家科学技术奖数”指标采信了科技部公示后无异议的结果。四是开展第三方调查，如对“学生满意度”这个指标，利用上海市教科院的专家力量，从“学生对学校的满意度”“用人单位对毕业生”的满意度两个维度进行较大样本的调查分析。

对自我评价的明确要求和多轮次应用第三方评价结果使得分类评价的结果更能反映不同高校的发展现状和水平，提高了各界对分类结果的接受度和认可度。从2018年和2019年两次高校分类评价的实施来看，不管是哪一种类型的学校，只要在人才培养、科学研究、社会服务、文化传承与创新、国际合作与交流五大职能方面表现突出，都可以在评价结果中得到反映。例如，2018年，侧重文理学科和侧重工科的两所综合性学术研究型高校双峰并峙；某纯文科的特色性应用研究型高校在同类型高校中拔得头筹；某以艺术类专业为主的特色性高校在应用技能型高校中名列前茅。

四、结果应用：发挥评价多重功能

随着教育评价理论和实践的发展，评价的功能从最初的鉴定、选拔、判断逐步发展到诊断、问责、改进。从我国教育评价的发展来看，“基于研究视角的学校评价指标、评价标准、结果分析等方面的探讨比较深入，但对教育管理者和实践工作者而言，如何应用学校评价结果更为重要。”实际上，不管是什么评价，都有结果应用的问题，只不过应用的主体、应用的目的、应用的力度、应用的方式不同而已。一个明显的例子是，尽管各界特别是学术界对各大排行榜诟病很多，各类主体却都在使用其评价结果。比如，针对目前影响比较大的高校四大排行榜（Us News、Qs、THE、ARWU）等，有学者认为它们“不严肃，不科学，不可信。”但事实上，教育部门往往用它来说明教育改革和发展成就，学校往往用它作为自身水平和地位提高的标志，学生和家长往往根据它来选择学校和专业。从国际经验来看，评价结果的应用也相当普遍。在英国，学科评估结果会影响政府对大学学科建设和科学研究的拨款份额。在法国，高等教育评估委员会组织实施的高等教育质量评估结果，会决定和影响政府与大学之间所签订的教育与研究合同。在美国，只有顺利通过六大区域认证机构或相应专业认证机构认证的高校才能优先得到联邦或州政府的经费资助。鉴于此，上海在实施高校分类评价之初，就明确提出要充分应用分类评价的结果。目前，评价结果已应用到高校内涵建设经费分配、高水平地方大学建设遴选、教师绩效工资、高校党政班子考核等诸多方面，发挥了评价的多重功能。

（一）发挥评价的诊断、改进功能

每次分类评价后，都会以集中反馈和个别交流的方式及时将评价结果反馈给学校。具体方式是以图表的形式向学校反馈其在同类型高校的地位，同时反馈同类型高校每个指标的平均水平和学校自身的状况，让学校知道自己的长处和短板，找到差距，明确努力方向，推动学校教育教学改革，提高人才培养质量和科学研究水平，增强社会服务能力，进而促进学校发展，体现了分类评价的诊断、改进功能。

（二）发挥评价的选拔、激励功能

2015年国务院启动“统筹推进世界一流大学和一流学科建设工程”以来，为了进一步提升高等教育整体水平，上海根据《上海高等教育布局结构与发展规划（2015-2030年）》，在

学术研究型、应用研究型、应用技术型、应用技能型四种类型高校中陆续启动了地方高水平大学建设工程，并给予重点支持。如何在众多高校中遴选出高水平大学建设高校？在分类评价中名列前茅是入选地方高水平大学建设行列的前提和先决条件（之后还要组织专家对学校的建设方案进行专门的评价），这充分体现了分类评价的选拔功能。与此同时，在安排每年约 15 亿元内涵建设时，打破了以往学术型和应用型、本科层次和专科层次不同标准的惯例，在统一不同类型高校校均内涵建设经费标准的基础上，主要根据分类评价结果在各自类型进行分配，这样就抑制了高校盲目上层次、换类型、不安自身定位的冲动，起到了引导不同类型高校在各自类型中各安其位、努力办出水平和特色的作用，体现了分类评价的激励功能。

（三）发挥评价的问责功能

20 世纪 80 年代以来，西方发达国家率先启动新公共管理改革，绩效优先成为席卷全球的教育治理的主导理念和行动策略。“换言之，政府做成了什么比做了什么更重要。”在我国，随着“管办评”相分离的教育体制机制改革的不断深入，政府在赋予高校更多办学自主权的同时也要求高校不断提高绩效。2015 年 10 月，国务院印发的《统筹推进世界一流大学和一流学科建设总体方案》中提出了四个基本原则，其中之一就是“以绩效为杠杆”。而要通过绩效来激发高校改革和发展的内生动力，强化对领导班子的问责制是必然的选择。正如美国著名高等教育专家克拉克所言，大学的兴旺与否取决于其内部由谁控制。上海在应用分类评价结果时，对在分类评价中表现优异的高校党政班子给予奖励，对表现不佳的高校党政班子进行问责，包括将评价结果作为遴选、调整领导班子的重要依据，作为提高或降低班子成员考核等第和个人绩效工资系数的重要依据等，体现了分类评价的问责功能。

教育评价是政府治理高等教育的重要抓手和有效途径，也是高校教育教学工作的“指挥棒”。科学、合理的评价有利于提升政府教育治理水平，有利于推动高校特色发展。在评价过程中，如何兼顾各种多元价值、如何平衡不同主体的利益诉求、如何实现评价的多元目的，是一个考验管理者智慧和能力的问题，也是衡量一个地区现代教育治理体系是否完善、教育治理能力高低强弱的重要标志。上海紧紧抓住教育评价这个“牛鼻子”，打破大一统的评价模式，针对不同类型高校的定位和目标，研制了分类评价指标体系，实施了分类评价，并通过分类评价结果的应用，基本形成分类规划、分类投入、分类评价的上海高等教育治理新格局。上海高校分类评价在指标设计、操作程式、结果应用的整个评价活动中，所有评价利益相关方都积极参与到评价过程中来，充分表达了自己的观点，形成了有利于改进主管部门教育管理和学校教育教学的“共同心理建构”，具有第四代评价理论的典型特征，也可以说是第四代评价理论的成功实践。

——来源：《教育发展研究》，2020 年 10 月

遵循与超越：基于《柏林原则》的高校分类评价透视

——以上海为例

一、解读《柏林原则》：对大学排名的元评估

（一）高等教育自主持续健康发展受到挑战

在过去的十多年中，随着大学国际排名激增，高等教育受外部排名的影响越来越大，越来越多的排行榜很难也不可能全面反映大学的发展情况。因此，人们开始关注如何客观认识这些排名，对这些排名的可信度和可靠性开始反思，特别是对判断世界大学的表现和其相对地位的衡量标准进行质疑。作为对这种关注的一种回应，联合国教科文组织下面的欧洲高等教育中心（UNESCO-CEPES）和美国高等教育政策研究所（IHEP）于2004年成立了国际排名专家组（IREG）。该排名专家组中的等级评定者和其中所涉及的其他人联手制定了一套指导方针，以制定和评估不同等级的质量及其测度，即《柏林高等教育机构排名原则》（The Berlin Principles on Ranking Higher Education Institutions, 简称BP）。

（二）《柏林原则》的实践

这些原则包括有关数据收集、排名方法、权重分配及其结果发布等在内的16条通用指南，被作为大学排名审核实践的基础。2008年4月18日，成立了学术排名与卓越协会（Observatory），由其对提出申请的大学排名机构进行审核，并为通过评估的排名提供“IREG认可”的质量印章，以此作为排名认可的证明。2010年秋季，在柏林会议和网站上征求与会者的意见后，2011年5月，IREG学术排名与卓越协会执行委员会通过了最终排名审计规则和程序。2013年5月，有两个排名第一次被授予“认可”的称号：波兰Perspektywy大学排名和国际世界大学排名。

（三）大学排名的16点原则性共识

《柏林原则》的组成经过国际排名专家们的讨论，达成了16条原则性共识，公开发布了今后国际排名的《柏林原则》。该原则由四个部分组成，分别为“排名的目的与目标”、“排名的指标设计与权重”、“排名的数据收集与处理”和“排名的结果公布”，16条原则分布在四个部分中（见表1）。

表1 《柏林原则》的组成

评价要素	原则内容
目的与目标	1. 仅是评估高等教育的投入、过程、产出之方法之一。
	2. 必须明确排名的目的以及排名的目标群体。
	3. 认识到高等教育机构的多样性并考虑到他们不同的使命和目标。
	4. 提供清晰的排名信息的来源范围和指标内涵。
	5. 针对被排名的高校，要考虑不同语言、文化、经济和历史背景。
指标设计与权重	6. 排名方法保持清晰、透明。
	7. 指标选择要考虑相关性和有效性。
	8. 优先考虑产出而不是投入。
指标收集与处理	9. 确保权重的稳定性与尽可能稳定性。
	10. 注意伦理标准并吸收好的做法。
	11. 尽可能使用经过审核的、可验证的数据。
	12. 数据收集经过适当的程序。
	13. 对排名过程采取质量保证措施。
	14. 运用组织保障排名的可靠性。

结果公布	15. 让使用者清晰理解用于排名的所有信息，并提供排名呈现方式的选择。
	16. 编制方式要能够消除或减少原始数据中的误差，组织和公布的方式要能够使出错被校正。

注：此表是在原文的基础上，加上作者的翻译形成。其中原文来自美国高等教育政策研究院（IHEP）的官网：<http://www.ihep.org/sites/default/files/uploads/docs/pubs/berlinprinciplesranking.pdf>

2. 《柏林原则》的特点

该原则设计的初衷是对今后所有的大学排名和排行榜即高等教育排名机构提出通用性的原则，具有普适性作用的同时还具有约束性。但在实际审核过程中则遵循自愿性原则（事实上也不可能强制），同时明确指出，审核结果仅公开正面的审核决定，并且任何时候都不会发布单个审核分数。

二、上海高校分类评价概述

（一）上海高校分类评价

上海高校分类管理首先需要解决的是高校分类的问题，该问题的破解涉及到以下三个子问题：

1. 厘清分类的目的是什么

多年来，上海高校与全国其他高校一样，面临着发展同质化、片面追求大而全、学科建设缺乏顶层设计与统筹协调等问题，为了让学校专注于内涵建设，安于自身定位，潜心围绕自身的发展定位进行顶层设计，科学合理布局，上海市先后出台《上海高等教育布局结构与发展规划（2015-2030年）》和《上海现代职业教育体系建设规划（2015-2030年）》，为保障这两个规划的落实，上海市政府建立了三年一调整的动态机制，配套出台了《上海高等学校学科发展与优化布局规划（2014-2020年）》和《上海职业教育高质量发展行动计划（2019-2022年）》。

2. 如何解决分类的问题

在《上海高等教育布局结构与发展规划（2015-2030年）》中构建了高校分类的标准，即实行二维分类法：第一个维度是从人才培养主体功能和承担科学研究类型等差异性，将高校划分为“学术研究、应用研究、应用技术和应用技能”四种类型；第二个维度是按照主干学科门类（本科与研究生）或主干专业大类（专科）建设情况，将高校划分为“综合性、多科性、特色性”三个类别，由此形成“四横三纵”的“十二宫格”的分类格局，使全市高校由“一列纵队”变为“多列纵队”。

3. 如何让高校安于定位、长远谋划

上海市政府通过出台《上海市高等教育促进条例》和坚持开展高校分类管理的举措，从立法和行政层面确保规划的落实，构建相应配套资源分配和评价机制，引导高校明确在二维分类体系的定位，在各自的类型中追求卓越。

（二）上海高校分类评价指标体系

1. 形成过程

高校分类问题解决之后，接下来要解决的是如何对不同类型的高校进行评价。这涉及到评价标准的问题。上海高校分类评价的核心问题是设计符合上海高校分类特点的评价指标体系。经过长达数年的研究，上海高校分类评价指标体系于2017年基本定型，2018年首次开始试评估，2019年正式推行，年进行个别指标修正形成2.0版本。

2. 指标体系

该指标体系由四套指标体系组合而成，分别是学术研究型、应用研究型、应用技术型和应用技能型高校指标体系。每套指标体系的一级和二级指标甚至部分三级指标都是相同的，这表明上海所有大学都姓“高”，都必须体现出高等教育的特征。比如，党对高校的全面落实情况、立德树人根本任务的落实情况、高层次人才数等。同时，每一类型高校的区

别性特征也必须完整体现出来，为此，在三级指标的设计时加以重点研究并体现，在目前的指标体系中，每类型高校的独特性指标占到近 30%。这些独特性指标可以是完全与其他类型不一致的指标，也可以是在指标名称上与其他类型相同，但在具体的指标内涵界定中有着明显的区别。比如，高层次人才数。这一指标在四种类型的高校中都有，但对于研究型高校该指标选取的是将代表研究水平最高的“两院院士”“杰青才人”等作为重点纳入；而对应用型高校而言，则将“技术技能大师”“大国工匠”“教育部教学指导委员会主任和委员”作为高层次人才加以吸纳（详见表 2）。

表 2 上海高校分类评价指标体系(2020年修订版)

一级指标	二级指标	三级指标				
		学术研究型	应用研究型	应用技术型	应用技能型	
A. 办学方向与管理水平	A1. 办学方向	1. 党对高校全面领导的落实情况				
		2. 立德树人根本任务的落实情况				
		3. 发展规划水平（内涵不同）				
	A2. 管理水平	4. 教学质量保障水平				
		5. 财务管理水平				
B. 办学条件与资源	B1. 师资队伍	——	——	生师比	同前指标	
		6. 高层次人才数		同名称不同内涵		
		7. 高层次教学、研究团队数		同名称不同内涵		
		——	——	双师型专任教师占比		
		——	——	企业兼职教师占比		
		——	——	专业教师企业实践情况		
	B2. 教学资源	8. 具有国（境）外学习工作经历的专任教师占比		——		
		9. 本科生均课程数	同前指标	生均课程数	实践课时占比	
		10. 优质课程数	同名称不同内涵			
	B3. 支撑平台	11. 规划教材、精品教材数		同名称不同内涵		
		12. 教学、科研、实验平台数		同名称不同内涵		
	B4. 科研项目	13. 创新创业示范基地（项目）数				
		——	——	产学研合作情况	校企合作情况	
		14. 师均科研经费		同名称不同内涵		
	C. 办学质量与水平	C1. 学科专业	15. 师均科研项目数		同名称不同内涵	师均企事业单位委托项目数
16. 学科专业布局优化情况			同前指标	同前指标	专业布局优化情况	
17. 高水平学科数			同前指标	优势学科数	产教融合型专业数占比	
18. 高水平专业数占比		同前指标	同前指标	入选重点专业（一流专业）数		
C2. 创新成果		19. 学生奖项数		同名称不同内涵		
		20. 重要教学获奖数		同名称不同内涵		
		21. 师均高水平论文数		同前指标	——	——
		22. 高质量专利数		同前指标	——	——
		23. 科研奖项数		同名称不同内涵		
C3. 国际交流	24. 在校学历教育国际学生占比		同前指标	在校国际学生占比	与前同名称	
	25. 学生国（境）外学习、实习占比		同前指标	——	——	
	26. 国（境）外办学机构（项目）数		与国（境）外高校学分互认高校数			
	——	——	——	国（境）外培训项目数		
C4. 社	27. 技术转移情况		同前指标	转化科技成果情	同前指标	

注：该指标体系根据 2020 年上海高校分类评价指标体系整理而成，其中标为“——”的地方为无对应指标。

三、对标《柏林原则》：上海高校分类评价的五维度透视

（一）《柏林原则》的五个维度

《柏林原则》的 16 条通用原则实际上是对大学排名的五个维度进行细化和明确，即对评价的清晰度、透明度、完整度、相关性和合法性的约束和确定。具体到与 16 条通用性原则的

对应关系可用表 3 来表示。通过表 3 可知, 16 条通用原则均可以从以上五个维度进行归纳和概括, 但个别通用原则可能与以上五个维度的两个甚至三个维度有关联。比如, 就与“清晰度”“完整度”“相关性”均有关联, 与“清晰度”“合法性”有关联, BP-15 与“清晰度”“完整度”相关联。

表 3 《柏林原则》五维度与原则对应关系

维度	对应原则
<i>Clarity</i> 清晰度	BP-2. 必须明确排名的目的以及排名的针对群体。
	BP-4. 提供清晰的排名信息的来源范围与指标内涵。
	BP-9. 确保权重的显著性与尽可能的稳定性。
	BP-15. 让使用者清晰理解用于排名的所有信息, 并提供排名呈现方式的选择。
<i>Transparency</i> 透明度	BP-3. 认识到高等教育机构的多样性并考虑到他们不同的使命和目标。
	BP-5. 针对被排名的高校, 要考虑其不同语言、文化、经济和历史背景。
	BP-6. 排名方法保持清晰、透明。
	BP-11. 尽可能使用经过审核的、可验证的数据。
<i>Completeness</i> 完整度	BP-4. 提供清晰的排名信息的来源范围与指标内涵。
	BP-15. 让使用者清晰理解用于排名的所有信息, 并提供排名呈现方式的选择。
<i>Relevance</i> 相关性	BP-4. 提供清晰的排名信息的来源范围与指标内涵。
	BP-7. 指标选择要考虑相关性和有效性。
	BP-8. 优先考虑产出而不是投入。
<i>Legitimacy</i> 合法性	BP-1. 仅是评估高等教育的投入、过程、产出之方法之一。
	BP-9. 确保权重的显著性与尽可能的稳定性。
	BP-10. 注意伦理标准并吸收好的做法。
	BP-12. 数据收集经过适当的程序。
	BP-13. 对排名过程采取质量保证措施。
	BP-14. 运用组织保障排名的可靠性。
BP-16. 编制方式要能够消除或减少原始数据中的误差, 组织和公布的方式要能够使出错被校正。	

(二) 基于五维度分析上海高校分类评价

1. 评价的清晰度

(1) 目的与目标的清晰度。

BP-2 指出, 大学排名必须明确排名的目的以及排名的针对群体。对照该原则, 上海高校分类评价的首要的目的就是解决上海四种类型高校的分类评价标准问题, 直接服务于上海高校分类管理的目地, 实现政府对高校分类指导和高校分类发展。从这个角度看, 上海高校分类评价的目的是极为明确的。同时, 该评价在这一目的的指引下兼具其他目标, 比如, 通过评价结果与内涵建设经费、民办教育专项资金、高级职称职数设置和绩效工资分配动态调整机制、高水平建设高校遴选、高校党政负责干部选拔调整和绩效考核“挂钩”。

信息来源与指标内涵的清晰度。BP-4 强调, 评价应提供清晰的排名信息的来源范围与指标内涵。对照这一条, 上海高校分类评价指标体系中, 指标的数据来源坚持“尽可能使用公开发布的权威数据”这一原则下对不同数据来源做了十分精细的规定。目前上海高校的数据和信息来源主要有四个渠道: 一是来自第三方高等教育评价机构的调查与评价结果。采信已公开发布、可公开获取、社会认可度较高的国内外第三方评价结果, 如“优势学科数”采用

了教育部学位与研究生教育发展中心全国一级学科评估结果、美国科技信息所（ISI）基本科学指标数据库（Essential Science Indicators, 简称 ESI）的学科排名以及 QS 公司的全球大学排名；“毕业生满意度”则采用了第三方开展的毕业生满意度专项调查结果。二是来自政府相关教育行政部门的督察与评价结果，如“党对高校全面领导的落实情况”“立德树人根本任务的落实情况”“财务管理水平”等指标来自政府相关教育行政部门的检查评价结果。三是来自专家组的评价结果，如“发展规划水平”“教学质量保障水平”“办学特色”等指标均由专家根据高校提交的相关材料进行评价打分。四是来自高校办学的各项基本数据，如“科研经费”“高层次人才数”等。

权重显著性与指标稳定性。BP-9 认为，大学排名还应确保权重的显著性与尽可能的稳定性。上海高校分类评价在适应原则方面做到了极致。该评价在找到高等教育通用性指标的同时，还大力研究四种类型高校的区别性指标，并对区别性指标赋予显著性的权重。此外，由于国家教育政策导向的调整，上海高校分类评价与时俱进，在保持评价指标体系相对稳定性的前提下，经过两轮评价之后对部分指标做了微调，但总的原则是保持体例和权重的相对稳定性。从以上两个角度来看，上海高校分类评价吻合了 BP-9。

提供清晰全面信息。BP-15 认为，应让被排名者清晰理解用于排名的所有信息，并提供排名呈现方式的选择。这一原则指出了大学排名过程中要求对信息做全面性的解释和说明，让被排名者理解指标的内涵，还要让被排名者知道指标选取的原因，更要让被排名者明白今后的努力方向。与此同时，在排名的呈现方式上要让被排名者有一定的选择权。上海高校分类评价根据队内高校的学科差异性，在个别二级指标设计中对三级指标设置了可选项，目的在于尽可能弥补学校因学科差异导致的不公平。比如，在“社会服务”二级指标下设置了“技术转移情况”和“决策咨询报告采纳数”两个三级指标，并允许学校选极值权重。这样就兼顾了人文社会科学学科为主体的高校与理工科院校的公平性问题。

（2）评价的透明度

大学排名的生命力在于其排名目标明确性以及为达到此目标所采用的数据的可信度和排名方法的科学性。在《柏林原则》中对评价的透明度要求很高，至少应在以下四个方面得到体现：

首先，要透明地认识到被排名对象的多样性。指出，要认识到高等教育机构的多样性并考虑到他们不同的使命和目标。上海高校分类评价的直接目的就是为上海高校分类管理、分类指导和分类支持提供科学依据。因此，针对上海四类高校设置了四套分类评价指标，充分考虑到大学的多样性问题以及形成的不同高校所承担不同的使命和目标。这点在指标设计时就对不同类型的高校设置了个性化的指标，比如，要求研究型高校更多地向世界一流大学对标，要求应用型高校更多地对接地方经济社会发展的需求。

其次，要透明地分析排名的背景。BP-5 还提到针对被排名的高校要考虑其不同语言、文化、经济和历史背景。这一原则设计的初衷是在国际化背景下如何解决不同国家的大学排名的可比性问题。具体到上海高校分类评价而言，由于该评价覆盖的高校是同一个地域不同高校之间的比较，基本不存在语言、文化、经济等大方面的巨大差异。但是，与该原则精神实质相呼应的是，该评价更多考虑的是对上海区域内高等教育机构的数量、各大学自身的体量、各大学的学科、各大学自身的发展特色以及自身的文化历史传承等背景因素的全面考虑。因此，该评价在四套评价指标中留出了不同高校对自身特色表达的空间，比如设置“办学特色”指标，对高校在上一年度的人才培养、科学研究、社会服务和文化传承等方面所做出的重要

贡献，选取其中最特色的部分形成典型案例予以呈现，并以行内专家评价评判的方式进行打分，形成定量结果予以呈现。

再次，要透明地公布排名的方法。BP-6 提出大学排名要对排名方法保持清晰、透明。在这点上，上海高校分类评价简直做到了极致。该评价不仅对综合结果公布了排名的方法，而且对每个三级指标的评价方法做了详细的交代。总的思路是，对所有的三级指标进行队内相对比较，也就是说，每一个三级指标都有一个标杆院校（一般来说，原始数据最优的高校，也有个别指标是绝对比较，如生师比指标），该院校该指标就可以得到权重中设定的满分，其他学校根据所在学校的原始数据与标杆院校进行相对比较之后进行比较得分。同时，考虑到个别指标具有强烈的导向和激励作用，设置了最低分下限标准，如党对高校全面领导的落实情况。

最后，要透明地使用各类数据。BP-11 是这样表述的：“尽可能使用经过审核的、可验证的数据。”在这一点上，上海高校分类评价一开始就坚持“尽可能使用权威的、第三方数据”，对于经过审核的、可验证的数据是“能用尽用”“应用尽用”。之前本文已对数据四种来源进行过介绍，对此需要进一步说明的是，上海高校分类评价的所有数据都是经过审核和验证的，即使是高校自己提供的数据，也至少有三个手段予以保障其真实性和可靠性：一是高校承诺制，上海高校分类评价是政府行为，因此作为教育行政部门有权利和义务对高校提出相关要求，其中包括信息和数据真实性的要求。二是数据核查机制。该评价所有的数据和信息都是通过专业的第三方数据平台进行采集，目前系统平台有强大的技术支撑力量，可以做到对学校所有关联数据进行相互验证和对异常数据实施预警；三是实施队内高校间数据公示机制。数据和信息采集结束后，允许高校在一定时间内对队内其他高校的数据进行比对和核验，对于数据造假行为进行严格核查，一经查实严肃处理。

此外，后期教育主管部门还将对部分高校的数据进行抽查，实地走访，加强督导。总之，经过之前的自我诚信保证机制、系统平台的核查机制，加之高校间的公示机制以及政府的主动督导机制，上海高校分类评价的数据是经得起检验的，是对《柏林原则》的数据透明度要求的超越。

（3）评价的完整度任何大学排名要想得到世人的高度认可，其评价的完整性十分重要，对这种完整性的要求，在《柏林原则》中主要由以下两条体现。这两条原则在之前的清晰度和透明度中也有所涉及，但在此处完整度可以有另外一番解读：

首先，排名指标要完整。在 BP-15 是这样表述指标的完整度的——“让使用者清晰理解用于排名的所有信息”，这意味着任何一个大学排名需要有自己的排名逻辑。众所周知，世界上没有一个排名能完美地评价大学的综合表现，任何排名都是对某一类型、或某一学科、或从某一角度对大学的表现进行排名。但无论是哪一个大学排名都应该有自己完整的排名逻辑和体系。就拿泰晤士高等教育（THE）大学排行榜来说，THE 自称，“相信我们已建立了大学表现国际比较的金本位”，“我们坚信，在我们已创造的排名代表了全球高等教育的最精确图景”。的确，THE 大学排行榜较合理地解决了研究型大学综合表现的横向比较，可以从他们最新的排名指标体系（见表 4）看出，该指标体系由 5 个一级指标、12 个二级指标组成。

表4 THE大学排名指标体系(2019版)

一级指标	二级指标	权重
教学（学习环境）30%	教学声誉调查	15%
	员工与学生比例	4.5%
	博士学位与学士学位授予比	2.25%
	博士学位授予学术人员比例	6%
研究（数量，收入和声誉）30%	师均收入	2.25%
	声望调查	18%
	研究收入	6%
引文（研究影响力）30%	研究生产力	6%
	研究影响力	30%
国际展望（工作人员、学生和研究生）7.5%	国际学生比例	2.5%
	国际工作人员比例	2.5%
	国际合作	2.5%
产业收入（知识转移）2.5%	机构收入	2.5%

注：该指标体系为THE官网发布的2019年大学排名指标体系。

通过表4可知，作为对研究型高校的综合表现进行评价的排名榜，该指标体系是相对完整的，选取的指标分别是大学的教学、科研、国际化和社会服务四个方面的指标，其中对大学的研究功能特别重视，其指标的权重具有显著性，占60%以上。上海高校分类评价则更是在此基础上通过4个一级指标、个二级指标和30多个三级指标完整呈现各类型高校的综合表现。四个一级指标主要从“办学方向与管理水平”、“办学条件与资源”、“办学质量与水平”、“办学声誉与特色”维度综合反映高校的办学表现。

其次，排名信息要完整。在BP-4中对信息除了清晰度的要求之外，还有对信息完整度的要求。此处的信息可以理解为所有指标下位的相关信息要完整，主要包括指标的内涵界定、数据来源、信息审核、数据分类与处理、评价方法等各要素齐全。上海高校分类评价在实施过程中对每个指标进行深入研究，专门有相应的指导手册进行说明，从指标定义、数据来源、评价方式等各个角度予以详细说明，并对数据的真实性和可靠性采取一系列手段进行保障（前文已述），对数据的分类和处理列表单呈现，对定性材料的撰写提出明确要求，对各指标的权重和权重处理方式都给予以一定程度的说明，如此等等。

（4）评价的相关性

相关性是指两个变量的关联程度。一般地，从散点图上可以观察到两个变量有以下三种关系之一：两变量正相关、负相关、不相关。大学排名的相关性原则在《柏林原则》中被明确提到，主要体现在以下原则中：

指标要有相关性和适用性。BP-7是一条真正的指导性原则，大学排名“指标选择要考虑相关性和有效性”。这一原则是指大学排名的所有指标之间要有相互关联作用，但是彼此之间又要有一定的有效性，因为大学本身就是一个复杂多面的有机体。正如美国著名大学校长克拉克·克尔在其专著《大学之用》中所述，现代大学已经演变成“多元化巨型大学”，简单地通过一个排行榜反映大学的全部越来越不真实和不完整。因此在构建和运用大学排行榜时，对指标的相关性和适用性需要重点关注。

指标要尽可能考虑产出。在《柏林原则》中还提到了“指标要优先选择考虑产出而不是投入”的原则（BP-8），这一原则实际上代表了当前大学发展和评价的一个重要导向，即对大学绩效（accountability）的关注。上海高校分类评价指标体系的原型就来自于上海地方高校绩效评价指标体系，至今该指标体系仍可以看出对绩效重点关注的痕迹，在100分总分中，四类高校指标体系中“办学条件与资源”明显是属于投入性指标，该部分的权重占到17-20%左右，而“办学质量与水平”这类产出性指标则占到左右。总体而言，明显是导向高校的绩效，即重点看学校的产出。

（5）评价的合法性

合法性，顾名思义就是指合乎“法律”“法规”“法则”“制度”“程序”等方面的规

定。在西方学者规范中，“伦理”也是“法理”的内涵之一。因此，在《柏林原则》中，有以下角度的原则对其进行框定和阐述：

首先，对大学排名的局限性予以明确。BP-1 即排名“仅是评估高等教育的投入、过程、产出之方法之一”。这一法则明确说明了任何排名只是对大学复杂有机体的认识方法之一，对大学的评价不应该只有排名这一方法。作为高等教育机构，每一个个体都是合理的存在，都有其存在的价值。鲜活丰富的教育活动不可能仅仅用冰冷的数据和排序就可以反映出来。这一判断极具警示作用。上海高校分类评价也充分认识到指标和排名的局限，因此，在选用部分定量数据的同时，逐步增加能立体化反映高校丰富办学活动的定性指标，现行的 2.0 版与 1.0 版相比，最大的区别之一也许就是定性指标数量的增加。

其次，对大学排名的稳定性予以明确。大学排名一旦定位准确，目标明确，其指标体系也应尽可能保持相对稳定，不可随意对指标设定的逻辑进行变动。这样一方面可以让高校逐步适应评价，另一方面可以保持排名的公信力。这一点在 BP-9 得到了呈现。上海高校分类评价作为政府推行的一项评价活动，更应在指标稳定性上做出典范，确保政府评价的公信力和权威性。事实上，在使用第二版指标体系时，为了保持指标的延续性，为了使高校更好地开展内部质量监控，该评价要求高校在系统平台填报信息时，仍将第一版的删除指标作为一项研究性指标予以保留并要求填写，但不计分。

再次，对大学排名伦理予以明确规范。在和 BP-12 条中，对大学排名的伦理要求是这样说明的：排名应“注意伦理标准并吸收好的做法。”此处的伦理主要对应的是程序的规范，包括对数据收集的程序（对应 BP-12）、对信息加工和处理的程序、对指标的权重和计算方法的程序、对最终排名结果的公布等。只有这样才能经得起被排名者和社会的质疑，赢得公信力。上海高校分类评价的操作程序更为系统化和机制化：首先指标的设计是广泛听取包括被评者为主体的各利益相关方的意见，并形成机制；二是对指标的权重采用被评者相互博弈之后的结果，即所有的三级指标都是经由大学自己设定权重，然后在队内高校间取均值。三是对指标的建议动态收集，目前在信息填报过程中，对每一个指标设置了建议窗口，所有被评者都可以就这个指标的方方面面提出建议。四是结果的公布程序，目前仍只在一定范围内公布，没有向社会和全教育系统公布。

最后，对大学排名纠错的机制予以强调。对此在《柏林原则》中专门有三条原则加以说明，分别是、BP-14 和 BP-16。这三条原则分别从质量保障措施、保障组织和纠错程序三个角度予以阐述。上海高校分类评价在纠错机制方面的做法如下：围绕分类评价的全过程，针对每个环节设计相应的保障机制，无论是数据采集环节还是数据核查环节以及后期的结果公布反馈环节都有一系列的保障措施并形成机制。就数据采集环节来说，该评价在实施这一环节时采取了填报前有培训、填报中有指导、填报后有公示等措施加以保障数据填报的质量。总之，上海高校分类评价对自身的纠错机制特别重视，并以固化程序的方式稳定下来，形成持续改进机制。

四、结束语

大学自身是一个复杂而丰富的有机体，要对其进行真实全面评价愈加困难。《柏林原则》的出现为我们重新认识大学排名提供了一个客观的视角，排名可以补充政府、认证机构和独立检查机构的工作。特别是它能加深我们对高等教育的了解，认识到排名的有限性和可能性。上海分类评价是对《柏林原则》的一种遵循与超越，是对自身追求科学合理排名的一种探索与尝试，是在构建中国特色世界一流大学的背景下一个难能可贵的中国大学排名方案，因为它既有理论研究又有行动实践，但它不是完成时，而是进行时，永远在路上。

——来源：《教育发展研究》，2020 年第 19 期

以首善标准建构首都高质量教育体系

在“两个一百年”历史交汇点上，党中央召开十九届五中全会，具有全局性、历史性意义。全会站在党和国家事业发展全局高度，将“十四五”规划与2035年远景目标统筹考虑，描绘了我国进入新发展阶段的发展蓝图，为全党全国各族人民夺取全面建设社会主义现代化国家新胜利指明了前进方向、提供了根本遵循。

特别是，全会做出了建设教育强国、文化强国、人才强国等一系列重大战略决策部署。如何以首善标准构建高质量教育体系，是首都教育必须回答好的时代课题。

第一，牢牢把握“定海神针”，就是坚持党对教育工作的全面领导。全会明确了“十四五”时期经济社会发展指导思想，提出了五个必须遵循的重要原则。坚持党的全面领导，摆在“五个坚持”首位。贯彻落实全会精神，首要就是加强党对教育工作的全面领导，以习近平总书记关于教育的重要论述为根本遵循，从“国之大计、党之大计”的战略定位来考量教育，坚守“立德树人”初心，牢记“为党育人、为国育才”使命，努力培养德智体美劳全面发展的社会主义建设者和接班人。

首都教育系统将深入学习宣传贯彻好总书记给首都教育的系列回信精神，以钉钉子精神打好高校党的政治建设攻坚战，积极推进思政课守正创新，开好“习近平新时代中国特色社会主义思想概论”课，扎实开展“使命在肩、奋斗有我”主题教育活动。将深入推进党建工作创新，引导广大师生相信党、热爱党、跟党走。同时，坚定不移推进全面从严治党向纵深发展，大力营造风清气正的政治生态和山清水秀的育人环境，更好地保证首都教育健康地向前发展。

第二，牢牢把握“根本标尺”，就是办好人民满意的首都教育。全会强调要“坚持以人民为中心”，这也是社会主义教育的本质要求，是首都教育要始终坚持的原则。人民对美好生活的向往，就是我们的奋斗目标。贯彻落实全会精神，首都教育将持续深入落实立德树人根本任务，培育和践行社会主义核心价值观，坚持“五育并举”，加强和改进学校体育美育，广泛开展劳动教育，发展素质教育，促进学生全面发展，培养学生爱国情怀、社会责任感、创新精神、实践能力。

持续落实“七有”“五性”要求，深刻把握人民群众对优质教育的热切呼唤，深入化解不平衡不充分的结构性矛盾，努力让每个孩子享有公平而有质量的教育，有人生出彩的机会。将进一步加强师德师风建设，全力打造让党放心、让人民满意的教师队伍，不断促进首都教育发展成果更多更公平惠及人民群众，不断增强人民群众的获得感和幸福感。

第三，牢牢把握“发展主题”，就是构建高质量教育体系。全会强调，“新时代新阶段的发展必须贯彻新发展理念，必须是高质量发展”。深入贯彻落实全会精神，就要深化教育改革开放，构建高质量教育体系。首都教育将聚焦国家重大需求和首都“四个中心”建设，持续推进“双一流”建设，进一步推动市属高校分类发展，不断提升首都高校学科建设水平和人才培养质量。

首都教育将进一步深化关键领域改革，认真贯彻落实《深化新时代教育评价改革总体方案》，切实破除“五唯”顽瘴痼疾。以首善标准推进招生制度改革，确保首都高考改革平稳有序、稳中求进。首都教育将不断优化教育空间布局，全力做好副中心教育供给；用好沙河、良乡两个高教园区，带动功能疏解和人口调控，高质量服务京津冀协同发展。

第四，牢牢把握“两件大事”，就是统筹抓好教育发展和安全。全会提出要“统筹发展

和安全，建设更高水平的平安中国”。在关于“十四五”规划建议的说明中，总书记再次强调了统筹发展和安全。保证教育系统安全，特别是政治安全，是压倒一切的工作。

贯彻落实全会精神，首都教育系统将巩固好安全稳定局面，坚持底线思维，强化风险意识，用大概率思维应对小概率事件，确保“精益求精，万无一失”。将紧盯敏感节点，旗帜鲜明抓好意识形态管理，防范各类“黑天鹅”“灰犀牛”。

将进一步教育引导全市教育系统广大党员干部，不断提高统筹教育发展与安全能力、教育规律探究与趋势把握能力、风险识别管控与化解能力、教育改革政策创制能力、新时期思政和群众工作能力、资源重组和机制重构与力量协同能力、政治生态营造与党建创新能力、校园文化培育传承与繁荣发展能力。通过这八种能力，全面提升教育系统治理能力现代化水平。

学习好、宣传好、落实好全会精神是当前和今后一个时期首都教育系统的首要政治任务。我们将扎实开展系统学习、加强宣传引导、深入研究阐释，认真编制好首都教育“十四五”发展规划，推动五中全会精神落地生根、形成生动实践。

——来源：《中国教育报》，2020年11月24日

江苏：积极参与学科评估，努力推动内涵发展

近日，中共中央、国务院印发了《深化新时代教育评价改革总体方案》，为教育评价改革明确了方向，是当下和未来教育评价改革的路标和航向。学科评估是教育评价的重要方面，是把握学科前进路线的“方向盘”，是研判学科状态的“听诊器”，是反馈学科质量的“风向标”，是改进学科工作的“加速器”，对指导和推动学科建设具有重要而独特作用。

我国实施的学科评估制度，是对具有博士、硕士学位授予权的一级学科开展的整体水平评估。自2002年启动实施以来，以其调查范围的广泛性、组织机构的权威性和评估立场的客观性，受到社会各界的高度认同。近期，为深入贯彻习近平总书记关于教育的重要论述和全国教育大会精神，认真落实《深化新时代教育评价改革总体方案》，《第五轮学科评估工作方案》（以下简称《工作方案》）正式出台，努力推动教育评价改革任务在学科建设实践中得到落实。总体看，《工作方案》做到了“三坚持”和“三统一”。

一是坚持问题导向，体现“破”与“立”的统一。

哪里是痛点和难点，哪里就是教育改革的重点和关键。“五唯”是当前教育评价指挥棒方面存在的根本问题，只有把“五唯”问题解决了，才能从根本上扭转功利化倾向，从根本上祛除浮躁之弊，还教育清静、清朗、清新之风。从《工作方案》看，第五轮学科评估工作展现了破除“五唯”顽疾的决心和行动，坚持先行先试、重点突破、系统治理、整体提升，实施了许多大胆的改革和探索，对扭转不科学的评价导向具有重要意义。比如，对师资队伍的评价，不唯“学历”和“职称”，没有设置人才“帽子”指标，避免以学术头衔评价学术水平的片面做法，而是采用“队伍总体结构与代表性教师相结合”的方法评价教师队伍质量，引导高校优化师资队伍结构，强化师资队伍内涵建设，这对推动高校解决“挖人才”“抢帽子”不良竞争具有重要意义。同时，评价注重把师德师风作为教师评价的第一标准，明确将教书育人作为教师评价的首要因素，这对推动高校打造有理想信念、有道德情操、有扎实学识、有仁爱之心的高素质教师队伍具有积极的引导作用。可以说，这是在推动破“五唯”方面的诸多有益探索，效果值得期待。

二是坚持立德树人，体现“育人”和“育才”的统一。

人才培养、科学研究、社会服务、文化传承创新、国际交流合作等是高等教育的基本功能，但人才培养是最基础、最核心、最根本的任务。习近平总书记多次强调，高校立身之本在于立德树人，要把立德树人的成效作为检验学校一切工作的根本标准，关键是要形成更高水平的人才培养体系。《工作方案》将“聚焦立德树人”作为评估工作的第一基本原则，将“进一步落实立德树人根本任务”作为第一改革举措，坚持以树人为核心，以立德为根本，把人才培养质量作为首要评价标准，构建了“思政教育”“培养过程”“在校生”“毕业生”四维度的评价方法。此外，此次评估将通过大规模的“学生调查”和“用人单位调查”，将教学质量和毕业生质量的评价话语权拓展到“需求侧”，引导学科既注重对培养过程的投入，也注重对学生成长的关心，还注重对职业发展的关注，具有重要意义。《工作方案》适应新时代新要求，首次增加了“思政教育”二级指标，把思想政治教育放在人才培养首位，把德智体美劳全面发展贯穿于评价在校生成和毕业生始终，这对有效推动学科建设更加聚焦“四为”方针具有重要的导向作用，体现了学科建设的鲜亮底色。

三是坚持分类评价，体现“共性”和“特色”的统一。

《深化新时代教育评价改革总体方案》强调，要推进高校分类评价，引导不同类型高校

科学定位，办出特色和水平。经过多年的发展，我国高校办学定位与学科建设呈现出多元化、多层次、多类型格局。此次学科评估制订了一个动态、开放、多元、包容的评估指标体系，在分类评价方面较之前四轮学科评估有了更加显著的进步。在宏观层面，针对自然科学、哲学社会科学不同学科门类特点，进一步完善了分类评价指标体系，强化“代表作”和“典型案例”评价，设置开放性“留白”，这对推动高校更加关注学科特色和内涵建设、遵循教育规律、突出质量导向具有重要意义。在微观层面，评价指标体系既考虑共性标准，也兼顾不同类型、不同层次的特色发展，构建了统一的指标体系框架，具体包括4个一级指标、10项左右二级指标和20项左右三级指标，并按照一级学科分别设置99套指标体系，充分体现了分类评估、分类引导、特色发展、差异发展的思想。

学科评估工作历经近20年的发展，具有深厚的工作基础和丰富的实践经验。第五轮学科评估着眼于全面贯彻党的教育方针，以“立德树人成效”为根本标准，遵循“质量、成效、特色、贡献”的价值导向，将对构建中国评价体系、树立中国评价标准、形成中国评价范式具有重要意义。

江苏是教育大省，也是研究生教育大省。江苏一直以来积极支持高校参加学科评估工作。在第四轮学科评估中，江苏有466个学科进入全国前70%，总数位居全国第二；80个学科进入A档，总数位列全国第三；A、B、C三类学科数均居全国总数的10%左右，学科建设呈现合理的分布格局，体现了百舸争流的良好生态。多年来，江苏充分利用学位中心提供的全方位、全覆盖、客观公正的学科评价服务，持续优化江苏高校优势学科建设体系，最大限度实现了以评促建、以评促改的功能。

在即将开展的第五轮学科评估中，江苏将以更大力度、更实举措支持高校凝聚共识、积极参评，以此来检验江苏学科建设水平，诊断学科发展问题，衡量人才培养质量，推动学科稳健发展。同时，江苏将积极引导高校正确理性看待学科评估结果，引导高校把参与评估的过程，变成梳理学科家底、规范学科管理、谋划学科发展的过程，积极营造风清气正的评估氛围，努力促进高等教育实现内涵式发展，推动江苏高校学科建设为实现民族振兴、社会进步做出更大贡献。

——来源：教育部网站，2020年11月3日

浙江：评价改革从区域层面破题

《深化新时代教育评价改革总体方案》旗帜鲜明地提出“坚决克服短视行为、功利化倾向”的要求，明确“三不得一严禁”，并将对教育生态问题突出、造成严重社会影响的地区，依法依规问责追责。这些关于评价问题负面清单的表述，反映了党中央改革区域教育发展评价、革除地方教育管理弊病、恢复良好教育生态的决心。

新时代推进教育评价改革，必须聚焦区域教育质量管理难题，以生态文明思想指导地方政府端正教育质量观，构建以教育生态为核心的区域教育发展评价体系，按照“健全综合评价，改进结果评价，强化过程评价，探索增值评价”的思路推动区域教育发展评价。

用绿色指标描述区域教育发展状态

2013年，教育部启动中小学教育质量综合评价改革，通过品德发展水平、学业发展水平、身心发展水平、兴趣特长养成、学业负担状况等5个方面20个关键性指标，探索建立了一套全新的“绿色评价”体系，以切实扭转评价学生只看分数、评价学校只看升学率的倾向。

这一改革是新时代教育评价改革的前奏。变单一评价为综合评价，从唯分数走向全面发展，是基础教育新的“指挥棒”。这一指标体系看似是学生综合素质发展的评价，其实是区域教育发展“健康体检”的框架。它引导全国中小学教育进入“绿色GDP”时代。

浙江省在这一框架的指引下，构建了“学生发展+成长环境”的中小学教育质量综合评价指标体系，尤其重视做好“四个关注”。

一是关注学生的全面发展，从“五育并举”的高度，重在记录学生的品德行为、身心健康、兴趣特长和劳动实践等方面的进步；二是关注学生的学习品质，不仅关注学生学业达标情况，还包括对高阶学习、学习品质、实践创新等方面的考查；三是关注学生的成长过程，重视师生关系、同伴关系、亲子关系，重视学生体验与负担感受的分析，研究学生学业负担状况与心理环境的安全性；四是关注影响学生成长的环境因素，从教师、学校、家庭以及区域等方面，综合评估成长环境对学生发展的影响。

教育是一个复杂系统。综合性是教育评价的重要特点。浙江的探索发现，区域教育发展评价既要在评价目标、评价内容、评价方法、评价结果运用等方面体现综合性，还要将学生发展评价与成长环境评价相结合，形成改进教育教学的综合对策。

用“结构型质量”替代结果式的评价

“维护教育生态，走向优质均衡”是区域教育发展的目标。但长期以来，一些地方仍将高考一本率、录取清华与北大的人数（以下简称“清北数”）、中考高分人数视作区域教育质量的标志，热衷于显性指标的横向比较。但这只能代表部分尖子学生的表现。只聚焦这些“结果型质量”常常会掩盖区域教育发展中的机制性问题，没有均衡为基础的“优质”难以持续。

浙江省以反映区域教育优质均衡发展状况的“结构型质量”替代纯粹结果式的评价，促进教育系统内部均衡协调发展。以城区公办学校与城区民办学校在学生发展、学校教学管理和教师教学方式等方面的差异度来评估公民办教育的协调发展；以城区公办学校与农村公办学校的差异度来评估城乡教育的均衡发展；以贫困家庭子女与富裕家庭子女在学业发展上的差异度来评估学生学业成就与家庭经济地位的相关性，反映区域教育的公平度。将这些差异状况综合起来的“结构型质量”，可以反映某一区域教育均衡发展的水平。

为了对高水平均衡与低水平均衡形成差别导向，还可以将各地学业水平中等率或后进生率

转化为调校系数，结合到“结构型质量”的计算中，以综合反映某一区域教育的优质均衡水平。

当“结构型质量”成为区域教育质量评价的新模式，就突破了以往教育评价仅仅聚焦“结果型质量”的局限，有利于引导区域教育政策与实践更好地面向全体学生、促进全面发展、优化教育生态，制止当前部分地区功利驱动的“逆向决策”。

评价须从“重结果”转向“重过程”

区域教育管理中存在三种“质量”。一是“结果型质量”，反映学生个体的学习进步情况；二是“过程型质量”，反映学校的课程实施与教学管理的水平；三是“结构型质量”，反映区域内教育资源配置的科学与合理性，以及各类学校发展的均衡态势。

综合体现区域教育质量的“结构型质量”并非孤立评判生成。浙江采用由不同群体的“结果型质量”与“过程型质量”的差异度来量化评价。科学合理的“结构型质量”与良好的“过程型质量”是体现全面发展的“结果型质量”的基础。强化过程评价，须建立关于学校层面的“过程型质量”的观察、分析与反馈体系。

评价的目的不是为了证明，而是为了改进。新时代的教育评价须从“重结果”转向“重过程”。区域层面须通过对学校教育教学的观察与评价以及对区域管理过程的评估与反思，诊断学校与区域在教育及管理过程中存在的问题，发现实践经验，形成指导区域完善学校管理与评价的科学决策。

教育质量评价在本质上不应是区分性的评价，而应是指导性的评价。旨在诊断与改进的评价，须弱化以往政绩考核常用的总分累计和横向评价的做法，根据评价内容的不同性质，采用分项等级的形式描述评价区域与学校的质量状态，以便于被评价者积极地理解与运用评价结果。

评价要向被评价者传递积极的力量

无论是发现经验，还是问题诊断，都应让被评价者感受到被帮助和被激励。但任何区域与学校都有不同的背景，其质量自有长期发展形成的过程。片面地采用统一标准的横向比较，对多数地区与学校的影响是消极的。探索增值评价的意义在于以发展性评价的思路，激励每一个积极进步的个体。

新时代教育评价改革须变横向比较为纵向比较，看趋势比增量。在崇尚排名施压的管理评价氛围下，往往只有三分之一的被评价者能感受到积极力量。但当评价关注转移到被评价者与其历史的比较时，被评价者会感受到无可推卸的责任，并打开努力作为的空间。

如，在某些区域，公办教育与民办教育的差异很大，当然这也是历史形成的，评价的重点可以关注公民办教育差异是否在缩小，同时底部是否有抬高。

新时代教育评价改革须变统一评价为分类评价，尊重区域与学校的实践背景，以差异化的标准形成对不同评价对象的不同激励。对于城市化程度不同的县（市、区），区域教育发展评价的制度设计应有所不同、区别考虑。特别对于城乡学校，不少地方简单化地采用区域性统考和排名，挫伤了大部分农村学校的积极性，要分类制定努力可及的评价标准，引导城乡学校尽己所能地进取。

——来源：《中国教育报》，2020年11月10日

教育评价：应然性与实然性的博弈及超越

刘志军 徐彬

【摘要】教育评价的开展是其应然性与实然性之间相互博弈的过程，其博弈的过程又不断促进教育评价走向革新与发展。当下的教育评价在此博弈过程中，其弱势的应然性与强势的实然性导致了教育评价的目的性追求被固化、教育的目的性价值被异化以及教育改革的理念与实践被分化的困境。合理有效的教育评价应超越其当下应然性与实然性博弈过度偏向其实然的状态而趋近于其应然的状态，即厘清教育评价与教育目的之间的关系，合理定位教育评价；积极开展反思性评价，促使个性化评价的发展；提升教育评价研究和实践的专业性，推动教育评价走向专业化。

【关键词】教育评价；专业性评价；反思性评价；个性化评价

在教育改革和发展的新时代，教育评价作为对教育活动或现象进行价值判断的过程备受关注。教育评价的开展是其应然性与实然性之间相互博弈的过程，其博弈的过程又不断促进教育评价走向革新与发展。探讨和解决当前教育评价中存在的各种问题，要深化教育体制改革，健全立德树人落实机制，扭转不科学的教育评价导向。在当前教育处在知识更迭速度呈倍数增长、互联网加快知识传播速度和人工智能促进深度学习变革的时代，教育怎样改革才能契合社会快速发展的需要，教育教学怎样进行才能更好培养适应社会快速发展的人，教育活动如何评价以符合教育适应社会快速发展的节奏等问题，亟待教育评价的深度介入，充分发挥其应然的导向、激励、调节、育人等功能。

近年来，学界对教育评价的理论问题进行了积极探讨和研究。如有研究从哲学视角系统审视了教育评价中的哲学问题，并倡导一种弱功利的教育评价哲学理念。也有研究从历史视角深入反思了教育评价与教育质量之间的关系，还有研究从宏观视角全面探讨了教育评价发展和改革中的理论问题及改进方向。这些研究无疑促进了我国教育评价理论的发展与革新，但鲜有研究对教育评价内部的应然与实然之间的理论关系，及其对教育发展的影响作系统的澄清。本文从分析教育评价的应然性与实然性的博弈关系及状态着手，探讨博弈失衡导致的教育发展诸多问题及超越路径，以期对推动以教育评价促进教育改革和发展，发挥教育评价应有的功能和价值有所裨益。

一、教育评价的应然性与实然性博弈状态分析

应然性与实然性源于理性主义和经验主义两种不同的哲学传统关于应然与实然或价值与事实的认识论问题。应然性体现的是在自身性质、范畴和规律的基础上事物所应该具有或达到的状态，它与应然或价值相对应，更加重视对事物进行基于理性的价值审视和理想构建，具有一定的价值性和理想性。实然性体现的是事物在真实的世界之中所存在的真实状态和实际样态，它对应的是实然或经验，更加注重对事物进行基于经验的感性把握和真切观照，具有一定的经验性和实用性。应然性与实然性两者之间的关系是相互对立和相互统一的。对立体现在任何事物的实然性一定不会完全达到该事物的应然性，只能是趋向于或接近于应然性，两者永远处在相互背离和相互排斥的状态。统一则是体现在两者共存于任何事物之中，如果没有实然性，只有应然性，那么事物就会处于一种至善至美的状态而失去发展的必要和意义；相反，没有应然性，只有实然性，事物就会因为缺乏价值的引领和发展目标的引导而失去发展的方向和尺度。

教育评价在改革发展过程中常常具有应然性和实然性两种表现。教育评价的应然性是指教育评价应该是什么样的存在状态，即作为通过对教育现象、教育事实及教育问题的客观记录、描述，并作出相应的价值分析和价值判断的教育评价，应该为教育目的的实现而最大化发挥其应有的价值和功能。它强调评价的主观性标准和客观性标准的统一，是人们对教育评价应该是什么样的理想追求，也是教育评价实然性所应改进和发展的方向。教育评价的实然性是教育评价实际是什么样的存在样态，它因受现实中政治经济、社会文化、科学技术、人的心理等因素的影响，单纯强调教育评价的主观性标准或客观性标准，而偏离或远离应然的教育评价。理论上讲，教育评价的实然状态因偏于主观性标准或客观性标准而存在两种状态。偏于主观性标准的教育评价不强调评价内容的客观性，而重视评价主体的经验和权威，这易使教育评价走向随意化和经验化。相反，偏于客观性标准的教育评价则强调评价内容的客观性，不重视评价主体的专业性，这易使教育评价走向刻板化和统一化，这种评价样态多出现在以纸笔测试为主的考试之中。应然性与实然性对立统一的关系决定了教育评价的应然性与实然性是对立统一的，而对立统一意味着两者存在相互排斥、冲突和竞争，对立、排斥、冲突和竞争又蕴含着博弈。

一般而言，教育评价的应然性与实然性的博弈具有两种状态。其一，教育评价以其应然性为主，这种教育评价则是一种趋近于其应然的评价样态。这种博弈状态合理地规约和调节着教育评价的实然性过于经验化或统一化的倾向，有效兼顾了教育评价中价值与事实、理想与现实、主观和客观的统一，从而是一种良性的博弈状态。其二，教育评价以其实然性为主，这种教育评价则是一种远离或偏离于其应然的评价样态。这种博弈状态则使教育评价或倾向经验化，或偏于统一化，或注重主观性标准，或强调客观性标准，从而是一种不良的博弈状态。

就目前而言，教育评价的应然性和实然性之间的博弈处于第二种状态，而且是以过于强调其客观性标准的实然性为主的博弈状态，主要表现为教育评价强势的实然性与弱势的应然性两者博弈力量的失衡。教育评价的实然性或现实状态是人们将教育评价的部分结果作为教育活动的目的性追求，亦即分数、成绩和排名成为人们普遍在教育活动中的目的性追求，偏离了教育服务人的和谐发展的本真追求，这是一种过分强调客观标准的教育评价样态。其强势性主要体现为目前我国教育评价的主流地位是以各类考试为基础的评价，产生的结果必然是分数成为衡量教育的一切，分数成为教育的一切追求。在这种情形之下，对教育评价来说，其本身是否具有价值就是当下所需全力关注的分数和成绩问题，至于评价可能引发的诸多后果，则无须考虑。其行动为追求功利的目的所驱使，势必会漠视人的情感、精神价值的实质，把功利目标视为唯一目标，容易导致行为方式的“常规化”，使社会生活丧失多元价值的创造性；使分数被人们视为在教育活动中的唯一目标，导致考试成为教育评价的常规方式，教育生活和学生发展丧失了多元价值的创造性。而教育评价的应然性或理想样态应该为培育全面而有个性的人的教育目的服务，这是一种强调主客观标准相统一的教育评价样态。其弱势性主要表现为在分数成为人们普遍在教育活动中的目的性追求的压力下，教育评价的目的和相关政策因其导向乏力而与实践严重脱节，教育评价的应然价值和发展意义被不断消解，人的发展也被束之高阁。

二、教育评价的应然性与实然性博弈失衡引发的教育困境

教育评价是教育改革和发展的重要引导力量，犹如一根指挥棒引领着整个教育活动的各个环节。教育评价强势的实然性与弱势的应然性之间博弈的不对等和不均衡必然会弱化其对

教育活动的引导力，而这也为教育的发展和改革带来一定的困境。

（一）教育评价的目的性追求被固化，应然的教育评价功能发挥受限

教育评价是教育过程的一个重要环节，其根本旨趣是服务于教育目的的实现和促进教育的发展，它在整个教育过程中具有强烈的工具性。关于教育评价的含义，有学者梳理归纳为三种主要观点，即教育评价是对教育活动进行价值判断的过程，是提供评价信息的过程，是一种共同建构的过程。无论是强调价值判断，或着重评价信息，或注重共同建构，三者共同隐含着教育评价目的的导向性和工具性，其本质都是为实现教育目的服务。就教育评价的功能而言，它通过对教育活动的价值进行判断、发现和创新，不断地实现提升教育价值或教育价值增值的目的，其工具性价值更加明显。从以上不难发现，教育评价的涵义特性和应有功能从根本上规定着其具有服务于教育目的的工具性。教育目的主要表现为社会对人发展的基本需要以及人自身发展的需要，它是教育评价活动的出发点，指引着教育评价活动的目标和方向；它也是教育评价活动的归宿，促进着教育评价活动的变革和发展。无论是个人本位论下的教育目的，还是社会本位论下的教育目的，抑或文化本位论下的教育目的和生活本位论下的教育目的，可以归为其主要是通过一定的教育活动，促使个体的社会化和社会的个性化，从根本上来说体现的是人的社会性和个性化的和谐发展。因而，教育评价之于教育目的具有强烈的工具性和服务性，教育目的之于教育评价则具有一定的规约性和指导性，不同的教育目的指导下的教育评价会有不同，教育目的的微调甚至会促使整个教育评价系统发生结构性的转变。

以成绩为基础、以分数为主参照的考试是现今教育评价实然性的主要构成，通过评价而得的成绩分数自然成为教育各方利益相关者竞相追逐的目标，而人的发展需要和社会发展的需要被淹没在成绩分数之中。以分数为主的教育评价普遍被人们作为教育活动的目的性追求，造成了促进人全面而有个性发展的教育目的对教育评价的导向力不断式微，教育评价应然性功能发挥不足，尤其是其发展性功能发挥更为受限。分数至上或教育评价成为教育活动的目的性追求的原因不是教育发展的自然结果或必然结果，而是教育多方利益相关者的共同合谋造成了“剧场效应”绑架教育的结果。“剧场效应”是社会学 and 经济学上的概念，两种学科视角下的解释稍有出入，但个人身份与社会环境的失衡所引起整体的无序是其共同点。反映在教育领域，“剧场效应”的产生有其一定的发生逻辑，在教育目的制定与推行后，最初整个教育都会围绕教育目的有序进行，但后来某位学生发现通过不断的题海战术训练、做题技巧可以提高分数，某位教师发现通过不断灌输强化、增加上课时间和重复的作业训练可以大幅度地提高学生成绩，某位家长发现通过金钱消费补课可以提高孩子的学习成绩，某个学校发现学生的高成绩可以为学校发展带来诸多的名誉和财富，某个教育行政部门发现学生的高成绩可以为自己带来政绩，其他学生、教师、家长、学校、教育行政部门便会对其进行效仿，从而使整个教育形成了追求分数最大化的局面。当每个人都在追求分数利益最大化时，最终会使每个人的利益得到一定的损伤，整个教育活动也会陷入一种无序的状态，形成“剧场效应”。教育评价就是在这种多方合谋的“剧场效应”过程中逐步固化为人们在教育活动中普遍性的目的性追求。

（二）教育的目的性价值被异化，教育促进人全面发展的应然价值彰显不足

教育的目的性价值在于使人成人，所成之人指的是追求个人社会化和个性化和谐共进的完整人，是不断追求知识寻真、道德向善和生活审美的完满人，是不断超越自我和提升生命质量的生命人，是生活更加快乐、更加美满、更加文明的幸福人。教育评价则是实现教育的

目的性价值的必要衡量尺度和重要实践工具。但在现实中，“剧场效应”绑架下的教育造成以成绩分数为导向的教育评价被人们普遍作为对教育活动的目的性追求，而促进人全面发展的应然教育价值的彰显尤为不足，这从根本上导致了教育评价生态的失衡和整个教育活动的无序，从而教育的目的性价值被异化为使人成为追求成绩分数的工具，人的全面而有个性的发展被淹没在成绩分数的洪流之中，或被作为神祇之物而束之高阁。总言之，追求分数的最大化成为教育活动的目的性价值所在。教育的目的性价值被异化主要表现为三个方面。其一，加剧教育中的“马太效应”，进而导致整个教育生态的失衡和无序。“马太效应”是社会学和经济学中常用的概念，其典型特征是强者愈强，弱者愈弱。在教育中，分数导向的教育评价促使教师给予少数拔尖学生过多的关注，忽视大多数学生，造成教育不断走向精英化。同时，对于分数的追求还刺激着教育资源的过度竞争和消费，拥有更多财富的人去购置和占有丰富的教育资源，他们通过代际的传播和积淀，逐步掌握了拥有教育资源的主动权和便利权，进而从根本上实现了对弱势群体教育权的“掠夺”。这些都加剧了教育中的“马太效应”，进而破坏了整个教育生态，使教育在背离公平的道路上愈行愈远。其二，消磨学生的创新和批判精神，使学生沦为单向度的人。教育乃是个人自我实现的过程，个人就是自己生活的创造者，自身的创造者，因此，用一个统一的价值标准去限制个人的发展，去评价教育的结果，从根本上说是不可接受的。但分数导向下的教育评价具有强烈的普遍统一性评价标准，以此长久作为评价教育的结果，则使学生逐步沦为考试的机器和分数的奴隶，学生在共同评价标准之下追求高分，在应试的评价过程中只是一味求同，其创新和批判精神则不断地被磨灭和压制，学生同质化和平庸化的现象也愈发严重，学生逐渐沦为马尔库塞（Marcuse, H.）所言的单向度的人。其三，钳制学生的个性，导致学生成为一个不完整的人。理想的教育在于让个体在不同发展阶段都能依照自身的内在基础和发展特点充分地显现自身，使得个体在不同发展阶段都能达到自身功能的最大化，个体生命在此过程中逐步趋于完整而健全。但实然的教育评价迫使学生成为竞相追逐分数的工具，在此过程中学生的个性不断被磨损而走向同质化，学生的智力潜能得到较好的挖掘，智育得到充分的发展，而学生的审美能力、道德能力等综合素质则不断地被搁浅，最终，偏智轻质的扭曲现实使学生成为一个不完整的人。

（三）教育改革的理念与实践被分化，教育理论与实践的应然联系不够

教育改革是教育发展和进步的重要动力，是保持和增强教育活力的必要手段。综观国内外教育发展史，教育改革在教育的发展中无疑扮演着具有关键意义的角色，每次教育改革都或多或少为后人留下或反思或借鉴的教育思想遗产，而足够成功的教育改革大都能在一定程度上更新人们的思想观念。因而，教育理念的转变和调整才是教育改革最为根本的问题所在，也是教育改革最为艰难的工作，教育理念的转变与否关系着教育改革践行程度的深浅以及教育改革的生死存亡。教育改革的理念受制于教育目的，并最终服务于教育目的的实现，这和教育评价的应然性是相符合的。现实中，教育评价强势的实然性倒逼了其应然性的虚无，教育的目的及其目的性价值被教育评价的实然性所遮迷，教育改革的实践受其导向影响逐步与教育改革的理念不断分化，从而加剧了教育理论与实践的应然联系不够的紧迫感。

教育改革是一个综合性和丰富性的概念系统和实践领域，包括课程改革、教学改革、考试改革、师资改革等方面。为了更清楚地说明教育改革的理念与践行之间的分化，以课程改革为例进行阐释。课程改革是人们对课程进行的根本性改造，与一般的、具体的课程改进相比，综合性更强、范围更广、程度更深、持续时间更久，它往往涉及有关课程的价值观念、理论基础、指导思想、制度政策等方面的重大变化或方向性调整。课程改革的理念与践行的

分化主要有以下五方面。第一，在课程目标层次上，新课程改革强调的是三维目标的达成及核心素养的培育，但在课程目标践行时，知识与成绩仍然是实然的课程所追求的目的，分数仍旧成为课程目标达成的重要关注点，三维目标久久难以落地生根，核心素养也难以与实践发生深度联系。第二，在课程结构方面，新课程改革强调设置符合社会发展和人的发展的合理均衡的课程结构，重视学生的综合化发展，以及提升学生对课程的选择权，但在践行时，由于成绩分数的引领，综合化的课程设置流于形式，学生的课程选择转变为被选择，合理均衡的课程结构经过实践后因有名无实和枯燥无味变得失序错乱。第三，在课程内容层面，新课程改革重视学生的经验化知识和综合实践活动知识，但在践行中，知识的选择和学习都是以追求成绩分数为导向，学生通过具身生活和亲身实践而得的知识与经验被抛掷一旁。第四，在课程实施层面，课程实施主要通过课堂教学而进行，新课程改革下的课堂教学主要强调的是“以需定学”、“以学定教”，主张探究教学、合作教学、个性化教学，但在课堂教学之中，考什么、教什么、学什么，以及怎样考、怎样教、怎样学的现象导致了课堂教学囿于“考一教一学”的循环怪圈而难以突围。第五，在课程评价层面，新课程改革强调的是教学的改进和学生的发展，但在践行时，依然过分关注评价的甄别和选拔功能，忽视合格性考试评价的价值，更遑论评价的改进教学和促进发展功能价值的彰显。综上所述，由于受强势的实然性教育评价的影响，课程改革无论是在课程目标、结构、内容、实施和评价哪个层面上的理念都与实然的践行不协调，课程理论与实践之间的应然联系逐渐脱节，加剧了教育理论与实践不断分化的进程。

三、教育评价的应然性与实然性博弈失衡的超越路向

教育评价的应然性和实然性之间的博弈是教育评价中一个重要的现象，其实然性胜于应然性的博弈状态使教育评价和教育发展面临诸多亟待解决的问题。问题的解决需从源头着手，调整教育评价应然性和实然性间的力量对比，构建合理有效的教育评价，使教育评价的实然性趋近于其应然性。

（一）厘清教育评价与教育目的之间的关系，合理定位教育评价

教育评价弱勢的应然性和强势的实然性之间的博弈导致了实然性的教育评价被作为教育活动的目的性追求，这种普遍的追求现象强化了教育评价与教育目的之间关系的混乱程度。厘清教育评价与教育目的之间的关系有利于调适教育评价应然性和实然性之间博弈力量的对比。从评价与教育的关系来看，教育评价活动只是实现教育目的的手段，它永远不可能上升为教育目的本身，更不能超越教育目的。教育目的只是教育评价设计与实施的根本指导和发展方向，它永远不可能等同于教育评价，更不能被教育评价所压制。概言之，教育评价与教育目的是手段与目的的关系。教育评价是实现教育目的的手段，主要表现为直接手段和中介手段两种方式，即教育评价自身可以直接推动教育目的实现，其他主体如教师、学校、课程、教学等也可以通过教育评价实现教育目的；教育目的是设计、开展和反思教育评价的依据，主要表现为在设计教育评价方案时需要教育目的作为根本指导，开展教育评价方案的过程中需要教育目的进行追踪和调适，反思教育评价方案时需要教育目的作为根本参照镜像。

教育评价的定位从逻辑上来讲主要有三种定位。一是底线评价，以强调公平性和基础性的合格性评价。这种评价倡导学生基于基础知识学习和基本素养养成的全面发展，注重学生教育起点的公平性和普及性，并提倡教育资源的共享共用、教育的精准扶贫和城乡教育一体化、区域教育均衡化的发展。二是中线评价，以着重竞争性和选拔性的等级性评价。这种评价虽有一定的合理性，但它是关注部分学生的、不全面的、考试竞争成绩的评价，是集合优

质教育资源、扶持优势利益群体、限制教育均衡发展、鼓励精英教育的评价。三是顶线评价，以注重个性化和人性化的发展性评价。这是一种比较理想化的教育评价样态，它通过发挥评价的引导、育人和发展功能，发现每个学生的个性特长和兴趣爱好，并将学生合理地进行分流，使每个学生各行所需、各得其所、各安其身并实现个性化发展。当前，教育评价的实然性定位主要是游离于底线评价和顶线评价之间的中线评价，这种评价定位忽视了学生全面发展，压制了学生个性化发展，加重了教育的不公平，破坏了教育生态的平衡。教育评价的合理定位应以底线评价为基础，合理调整中线评价，突出发展顶线评价，如此，这种定位既体现了学生全面发展和个性发展的统一，也调和了合格性评价和等级性评价的矛盾而走向发展性评价。

（二）积极开展反思性评价，促进个性化评价发展

超越教育评价应然性和实然性的博弈，寻求其实然性趋近于应然性的发展路径，彰显和弘扬教育使人成人的目的性价值，还需在教育中积极开展反思性评价，促使个性化评价的实现和发展。反思性评价是建立在评价者自我反思和自我评价的基础之上的具有发展意义的一种评价，它所蕴含的价值在于追求人的主体意识的自觉和实现自我完善、自我发展的意向性。我国优秀的传统文化中已有反思性评价的描述，如“吾日三省吾身，为人谋而不忠乎，与朋友交而不信乎，传不习乎”，“学而不思则罔，思而不学则殆”，“学，然后知不足；教，然后知困”等，这些都属于反思性评价。但由于“高分才是王道”等一些思想观念固化为大多数人的文化基因和社会心理，外在标准成为衡量一个人的重要尺度，甚至是唯一度量，人的价值被外在标准所奴役，自我反思和自我评价的重要性与必要性不断被以分数为主的外在标准消解，反思性评价未能达成共识。

个性化评价是建立在评价者的自我反思和自我评价基础之上的评价，是开展反思性评价的必然走向，也是反思性评价的本质体现和价值追求，旨在尽可能发现评价者的多维可能性和发掘评价者的综合发展潜力，促使评价者在得到一般发展的基础之上成长为一个具有个性而独特的生命体。个性化评价的构建和开展需要从标准、评价者、制度和文化的四个方面着手。

其一，构建个性化评价的关键在于个性化评价标准的厘定，在构建个性化的评价标准时需要注意的是所选取的标准不应过度追求精确和精细，而在于标准之于评价的意义性和有效度；不苛求标准的全面和详细，而在于标准的针对性强，且注重评价者反思意识的提升和反思习惯的养成；不盲于评价标准的单一和同质，而在于标准的丰富性和可选择性，尽量使每个评价者都能得到适合自己的评价。其二，个性化评价的实施主要是评价者依据个性化的评价标准进行自我反思和自我评价，需要评价者自身确立以发展为主的评价观念。同时，个性化的评价是以过程性评价为主的评价，注重的是评价者生活和学习过程中的经验反思、问题总结和思想改进，这需要评价者将个性化评价的践行纳入自身的日常工作和学习生活之中，以此实现个性化评价的实施常态化。个性化评价并不排斥其他评价主体的参与，相反，个性化评价的有效实施正是以自我评价为主、他人评价为辅的一种评价体系。其三，个性化评价的实施需要建立相应的激励制度、监督制度、审核制度、申诉制度等保障，通过制度的激励和保障促使个性化评价走向持续发展和保持长效性。其四，个性化评价的持续开展和长效使用促使个性化评价文化的培育成为必要，个性化评价文化的养成需要将其与教学文化相结合，以评促教、以评促学、教学评合一，推动教学文化与评价文化的深度融合，增强教师和学生在学习中的心智投入度、批判创新力和文化反思力，促使个性化评价成为一种深入人心的教

育评价文化，从而推动个性化评价的长远发展。

（三）提升教育评价研究和实践的专业性，推动教育评价走向专业化

教育评价在整个教育活动中发挥着指挥棒和风向标的作用，其专业性程度的高低关系着教育评价的科学与否和信效度高低，影响着教育评价的应然性与实然性之间的力量对比，进而影响着教育理论与实践联系的密切度。因此，弥合教育改革理论与实践之间的裂痕，调适教育评价应然性和实然性之间的博弈力量，提升教育评价研究和实践的专业性，推动教育评价走向专业化是必然选择。教育评价的专业性价值蕴含在教育评价研究和实践两个领域之中。在专业的研究领域，教育评价有利于教育真问题的发现和研究，丰富教育研究的问题和方法域，促使教育研究范式的更新和转换，不断推动教育研究走向科学化，从而促使整个教育事业的进步和发展。作为一种专业实践，教育评价专业化有助于使教育摆脱自上而下的行政化倾向，让教育建基于自下而上的专业判断之上，让教师和学生成为教育活动的合法评价者之一，实现教师和学生的个性自由发展。

促进教育评价专业化建设可从以下三个方面着手。第一，增强评价主体的专业性。这需要推动专业的教育评价人员加强评价实践经验的提升和评价理论的反思；通过简便易行的评价知识推广和评价实践操作演示给予教师、学生等评价参与人员专业的评价培训；不断促进第三方教育评价机构去行政化和激活纯民间性质的第三方评价机构的活力。第二，加强评价方案设计和实践的专业化。评价方案的厘定从目标的筛选、人员的选定、准则的制定、方法的选择、操作程序的安排到方案的元评价都需要专业人士从专业的角度秉持专业负责的科学精神进行反复商讨论证，并且方案的实践还需专业的评价人员遵循科学合理的程序进行，才能保证评价的信度和效度。第三，加大教育评价研究的专业化。专业化的研究是教育评价专业化的重要支撑和保障，需要高校加强教育评价学的学科建设，使其成为教育学科中独立的研究领域，并为教育评价研究培养和储备一定高质量的人才；而教育评价领域内的专家学者则应加强对教育评价认识和实践的自我反思以及强化对教育评价元研究的意识和能力。同时，国内教育评价的研究还应致力于构建中国特色的教育评价理论体系和话语体系，增强我国教育评价研究的能力与特色，以此推动教育评价研究的专业性建设。

【刘志军，河南大学副校长、教授；徐彬，河南大学教育科学学院博士生。】

——来源：《教育研究》，2019年第5期

教育科研评价如何走出困局

孟照海 刘贵华

【摘要】构建中国特色教育科学三大体系的关键在于评价体系改革。在强调技术理性和绩效评价的社会背景下，教育科研评价目的逐渐异化为对新颖、绩效和功用的追求；学术共同体因行政力量的强大和第三方评价的量化推动而逐渐被弱化；评价标准简单划一，忽视不同知识类型、研究类别和知识情境中的科研差异；评价方法片面追求量化指标的简单加总，造成“劣币驱逐良币”；评价结果易受非学术因素影响，致使公信力不高。究其原因，作为评价对象的教育科研通常具有知识生成的复杂性、质量显现的滞后性和成果影响的间接性等特点。评价标准随社会政治和知识范式的变化而变化，并强调个体化和特殊性知识，同时整合了多维度的知识和研究方法。在评价制度上，存在推动创新与排斥创新、内在评价与外在评价以及技术理性与总体判断的矛盾。改进教育科研评价体系，需要实施分类评价、完善同行评价、强化综合评价、推动多元评价、健全公开评价，使评价真正发挥鉴别、诊断、规范、引导和激励作用。

【关键词】科研评价；教育科学；质量标准

构建中国特色的教育科学学科体系、学术体系和话语体系，关键在于科研评价体系改革。破除“五唯”的总体部署以及《深化新时代教育评价改革总体方案》都对评价改革提出了重要性和紧迫性的要求，“改进结果评价，强化过程评价，探索增值评价，健全综合评价”，更是将评价改革摆在了更加突出的战略位置上。科研评价是对学术品质和学术贡献优异程度的鉴别与认可，它既是学术知识生产的内在动力，也是学术资源分配的主要机制。

科研评价在项目评审、人才评价、机构评估中发挥重要作用。从世界范围来看，科研评价在很大程度上支配着科研机构及专业人员的行为；把控着学术论文和专著的发表以及学术会议的主题发言；主导着学术人员的聘用和晋升；决定着科研资助机构的基金项目分配以及国家的基本科研经费分配；对高等学校的研究生招生同样也发挥着重要影响。推进教育科研高质量发展，必须紧紧扭住评价改革这个“牛鼻子”，克服科研评价中过度数量化、简单化和形式化的倾向，使教育科研评价真正发挥鉴别、诊断、规范、引导和激励的作用。

一、教育科研评价如何误入歧途

学术评价归根到底是一个“分配正义”的问题，是通过对每个人学术劳动的准确判断，给予他们与其贡献相当的回报和相应的社会责任，其最终目的是选拔优秀人才，激励先进成果，推动学术进步。科研评价是形式评价、内容评价和效用评价的“三位一体”，包括评价主体、客体、目的、标准、方法与制度六个要素。教育科学的繁荣发展有赖于合理的评价体系和奖励系统，正如培根（Bacon, F.）所言，“得不到奖励的事业就不会繁荣”。教育科研具有特殊性和复杂性，在强调技术理性和绩效评价的社会背景下，教育科研评价在目的、主体、标准、方法和结果等方面容易发生异化或错位，致使科研评价未能有效、合理地发挥规范和引导作用，激励和奖励体系的识别偏差也带来了教育科研的虚假繁荣和学术泡沫。

（一）评价目的异化，片面追求新颖、绩效和功用

在教育科研领域，评价目的异化主要表现为三种评价倾向。一是为创新而评价。从学术共同体内部而言，教育科研的实质性评价是对学术原创性的鉴别，归根结底要实现知识的增长。但在科研评价中，真创新与假创新通常难以区分。美国社会学家默顿（Merton, R.）就

指出，原创性评价的最大难题就是，把值得承认的真正的创新与不值得承认的仅仅是新颖之物区别开来。在教育科研领域，由于教育世界的人文创造性，研究者常常把自己的情绪、态度、偏好等主观因素掺进研究过程，把自己所意愿或期盼的研究结论或主张混同于研究证据所支持的客观结论，研究的主观性、私人性和相对性也被幻化成研究的创新性。在追求标新立异的科研评价导向下，教育研究常被各种意见和臆断所包围，研究者和实践者要“穿越教育概念的丛林”，同时要面对教育研究的标签化趋同和低水平重复研究。二是为绩效而评价。从外部的资源配置而言，教育科研评价是为了资源使用效率和效益的最大化，因而要以物化的研究成果来证明“物有所值”或“物超所值”。在实际评价中，由于形式评价和效用评价便于测量和计算，能够更有效地对科研活动进行鉴别、区分和排名，因而更受管理机构和科研资助机构的青睐。此外，改革开放以来的教育资源配置十分倚重各种工程和项目，而其筛选有赖于容易操作的简单指标，从而驱使高等学校和科研机构片面追求科研的绩效和功用，不遗余力地制造甚至包装各种评价指标，以争取资源、获得头衔或更好的排名。因陷入这种竞赛机制中，许多人都在努力地做着他们明知道不对的事情。为绩效而评价的倾向带来了有数量缺质量、有专家缺大师的状况。三是为功用而评价。教育研究要能够有效解决实际的教育问题，科研评价通常比较重视研究成果是否“及时、有用和管用”。决策者和实践者对科研质量的评价也是强调即时有用性，特别是具体的、可操作性的建议，从而逐渐使教育科研成为一种纯粹的技术性活动。正如美国学者拉伯雷（Labaree, D.）所说，当教育研究者一心寻求科研的有用性时，他们有时就会削足适履，让自己的研究满足此时此地的教育实践需求，致使教育科研成为“短视”行为。而一些抽象的理论构建尽管在研究时看不到用途，却在将来大有用途。我们必须为“天马行空”的基础研究保留空间，以思考不可思议之事。我们需要研究者挑战基本假设和正统观念，从而使政策议题发生根本转向，如果研究者不去为之，就更不可能指望其他任何人。

（二）评价主体失范，学术共同体独立性和自主性力量不足

科研评价通常包括学术性评价、行政性评价与社会经济效益性评价。作为一种价值判断活动，它面临学术标准与非学术标准及不同评价主体、评价程序之间的多重矛盾。学术力量、行政力量和社会力量在科研评价中的力量并不平衡，在权力关系的支配下，评价主体的失范是造成科研评价乱象的重要原因。一是行政力量过强。国家是科研资源的主要供给者，从总体上调控着科研的方向和知识生产。教育研究具有鲜明的意识形态属性，政治体制和政治制度对教育科研评价发挥关键作用。面对同样的事实，在科学问题上，决定对与错的是其真假；而在价值问题上，决定对与错的却是评判者的立场。同样的事实可以为不同的甚至截然相反的价值问题提供支持，关键就在于证明者站在什么样的价值立场上解释这些事实。以学术评奖为例，学界和民间组织的权威性评奖很少，且让人议论纷纷，而主管部门组织的评奖则由于主管官员主导的色彩甚浓，官员或准官员所占的比例过高，使评奖的学术水准很难让人信服。换言之，行政力量既被排斥，但又被需要。二是学术力量被扭曲。在学术评价中，学术共同体的独立性和自主性因内外因素的影响而难以真正实现。在项目评审中对政策热点和政策意图的过于迎合，在成果评价中对领导批示的过于推崇，都将政治逻辑置于学术逻辑之上，最终损害学术评价的公信力和权威性。同时，由于学派和利益纷争，评价者发生角色错位，成为某部门或某个人的代言人，使正常秩序下的公平竞争蜕化为无序竞争。在不同的研究范式之下，对相同的教育事实也可能有不同的看法。在学术评价中掌握话语权的评价者就可能将自己的偏好强加到评价标准上，从而限制了学术创新和教育研究的多样化。三是第三方评

价重形式轻内容。由于同行评议的公信力不高，为显示客观和公允，以文献计量为主的评价机构日益膨胀，由此带来量化评价的泛滥和学术评价的异化。这些评价机构主要包括各种核心期刊与来源期刊、各种文摘类期刊以及各类排名机构，它们是量化评价和“出身”评价的主要推手，“以刊评文”和计件制是其评价的主要方式。这些评价机构的兴盛，正是由于同行评价的作用发挥不充分。

（三）评价标准单一，忽视教育科研的类型差异

评价标准是科研评价的关键，标准不当将会破坏整个科研生态。教育学科具有人文学科、社会学科和自然学科的多重属性，其知识是一种复合型知识。从教育科学的发展来看，尽管规范性的人文社会科学研究范式仍然占据主流，但进入21世纪以来，学习科学、脑科学等自然科学知识的引入使教育科研日益重视随机实验和准实验的实证研究方法，讲求客观、可量化、可验证、可重复等相关特征。在教育科学追求科学化的道路上，评价标准日趋单一。一是简单参照自然科学的评价模式，忽视教育科研评价的独特标准。在自然科学和社会科学模式的影响下，教育科研的评价标准比较强调严谨性、原创性和影响力，特别是定量和定性的研究方法及其科研的社会影响力。然而，教育研究不像自然科学那样具有极强的和直观的累积性，文献之间的相互引用有时显得可有可无。单纯以学术成果发表数量和引用率等作为质量评价标准，不一定适合教育科研的独特性。以文学和艺术史等人文科学的研究为例，研究者就比较重视促进文化记忆以及反思批判性等学科独有的标准。同时，即便是相同的指标，不同学科的侧重点也不一样。就原创性而言，人文与历史学研究者强调的是研究视角以及史料的原发性，社会科学研究者强调的是研究方法的原创性。因此，在教育科研评价中，必须凸显教育研究的独特性。二是按照一种知识逻辑评价不同类型的研究。基础研究、政策研究和实践研究遵循不同的知识生产模式和知识逻辑。基础研究侧重于对根本问题的原创性研究，遵循纯粹知识生产的模式，追求的是不以应用为目的的“无用”知识。政策研究侧重于政策的可行性和社会合法性研究，要兼顾知识逻辑和政策逻辑，追求的是学术创新与特定政治制度和政策环境的适切性。实践研究侧重于理论知识的实践应用，遵循知识逻辑和实践逻辑，追求的是实践者的认知和行为改变。而以同一种知识逻辑评价不同类型的研究，则会加剧教育研究的同质化。三是在不同情境中采用一种评价标准。好的教育研究不是一个抽象的概念，而是体现在不同的具体评价情境中。在教育科研领域，作为知识创新的科研成果不是一个客观的结构实体，而是主体与环境互动的产物，科研创新在不同的情境中被重新界定，产生不同的意义。通常，学位论文的评价侧重于学术规范的掌握和知识的原创性；学术期刊论文的评价侧重于填补知识空白和研究方法的合理性，以及与刊物风格的适切性；竞争性科研项目的评价，侧重于研究的学术价值和社会政治价值，以及研究的前期基础和研究设计；委托研究项目的评价，侧重于研究者的学术声誉及完成研究任务的时效性和经济性。对于不同评价情境中的教育科研，应该在坚持一般科研评价标准的基础上有所侧重，根据评价的对象构建个性化和针对性的评价指标体系，而不能采用一刀切的方式进行评价。

（四）评价方式简单，以数量指标取代质量判断

《第四次全国科技工作者状况调查报告》显示，论文仍是科技评价的最主要手段，62.1%的大学教师、44.9%的科学研究人员在各种形式的科研成果中最看重论文。学者可能在某个时期会出一些好的成果，但是不可能一辈子都出好成果，这是符合人的基本成长规律的。学术评价必须遵循知识创新规律和人才成长规律。首先，教育科研评价方式的简单化表现为数字崇拜和指标崇拜。这种评价方式以客观化的指标和数字抽离了不同个体学术生产的情境性、

复杂性和差异性，将学术生产作为简单的、程式化的流水线生产过程，把学术发表的期刊或出版社级别、成果的字数、科研项目的级别及经费数量作为评判学术水平高低的依据。进入 21 世纪以来，尽管社会科学引文索引（SSCI）上发表的中国教育研究论文数量猛增，但大陆学者在国际合作网络中仍处于“学术加工”的地位，因而要提升中国教育研究的国际影响力，需要纠正当前评价指标体系过度强调国际发表数量的导向，建立理性的评价体制，引导学术研究转向精耕细作的发展方式。其次，评价方式的简单化表现在不同成果之间的可换算性。在一些评价中，数量指标可以替代质量指标，可以用数量更多的较低质量的成果经过一定的换算，代替高质量的成果，这就使得“十年磨一剑”不如“著作等身”，造成“劣币驱逐良币”现象。以阿伦特（Arendt, H.）提出的“平庸之恶”来看，当前在人文社会科学领域，平庸的“伪学术”论文占用了大量的学术资源，研究缺乏“真问题”，不讲逻辑，“伪综述”等都是其重要的表现形式。“伪学术”无关社会生活，无关学术责任，比以剽窃与不当引用为主要特征的学术不端更为普遍、更为严重，因而也就更难以揭露、更难以治理。教育科研的评价方式要适应创新性学术的生成规律。创新性学术的生成需要天赋、勤奋和灵感的三位一体，优质学术研究者的稀缺性要求学术评价机制能够真正识别出学术天才、常才和庸才；学术成果形成的长期性要求不能急功近利年年考评，要使优秀的学术作品能够有一个如琢如磨的生长过程；学术评价的公正性要求期刊的运行要规范，避免平庸之作进入权威期刊而将学术庸才装扮为学术天才。再次，科研评价的简单化还表现在过于强调结果导向，忽视不同学科、研究者和研究条件的差异，完全“以成败论英雄”，造成研究者避重就轻、舍精求多，热衷于“短平快”的研究，甚至为了多出成果而篡改数据，产生学术腐败。罗尔斯（Rawls, J.）的经典著作《正义论（Theory of Justice）》用了 10 年以上的时间写作，而在这段时期中，他没有发表过一篇文章，但在这“没有成果”的 10 年中，他的待遇完全没有受到影响，这样的事情在中国几乎是不可想象的。

（五）评价结果的公信力不高，易受非学术因素影响

教育科研评价要面对技术的复杂性、社会的复杂性和人的复杂性，而标准缺失、权力支配和人情主导是造成评价结果公信力不高的重要因素。科研评价体系的问题表现在三个方面。一是评价标准的主观性和模糊性使评价结果难以获得普遍认可。当前，管理层面和学术共同体，对教育科研评价的根本原则还缺乏共识，甚至出现相互背离的现象。以科研影响力评价为例，英国学者霍金森（Hodkinson, P.）指出，科研发现或政策建议只有符合决策者的期望或预期时，才最有可能对政策产生影响，而这实际上与科研质量几乎没有任何关系。挪威科研评价委员会在教育科研评价中就出现了科研有用性和质量的对立，研究者认为追求有用性实际上损害了科研的质量，这使科研成为一种工具和附庸，从而限制了批判性和原创性研究的空间。可见，决策者和研究秉持不同的科研质量标准。正如美国学者韦斯（Weiss, C.）所言，“评估者认为相对无效的项目却得以延续甚至拓展，而评估者认为有效的项目却被削减”。对于此种研究与决策之间的鸿沟，韦斯将其归结为两种文化的差异，并提出了“两个共同体”理论。二是科研评价容易受到行政力量和人情关系的影响，在一定程度上学术依附权力，权力绑架学术，科研评议容易被人为操纵而丧失公信力。我国教育学科的同行评议制度不完善，突出表现在地方庇护主义和政府介入两个方面。从地方庇护来看，评审专家对自己所在学术机构、朋友和其他关系亲近者的过度保护，常以牺牲质量标准和其他学术机构、研究者的利益为代价。此外，由于教育研究具有较强的实践性，同时教育在政府部门中又有相应的、明确的管理机关，与其他人文社会科学相比，行政力量对教育学术界的同行评审的介入程度更

深一些。三是评价过程不透明和缺乏监督。部分科研工作者利用评价机制的不完善谋取不当利益。同时, 现有制度对评价者的评价行为监督相对有限, 发生道德风险的概率较大。许多由政府机构控制的评审项目, 无论立项还是奖励, 举凡专家的遴选, 评议对象的选定, 评议程序的安排, 评议结果的确定、公布和利用, 尽管越来越强调同行专家的参与, 但仍主要由管理部门操控。在中国这样的“人情社会”中, 学术问题时常被“人情化”, 使学术评价成为人情交易或交换的手段。

二、教育科研为何难以评价学术

评价就像量子测量, 遵循测不准原理, 一评就变味, 所以永远评不准。科研评价的制度化的从一开始就面临着多重矛盾, 知识与非知识的界限、“以学术为业”与竞争性表现的冲突、鉴别内在品质与授予社会认可的差异及不同社会文化和政治制度对评价的影响, 都使科研评价成为一项复杂而艰巨的任务。对于科研质量标准几乎没有广泛的共识, 而对于何为好的研究也几乎没有准确的界定。教育科研评价除了要面临一般科研评价的难题之外, 还要适应教育科研评价对象的特殊性、评价标准的模糊性及评价体制的矛盾性等难题, 这使得教育科研评价异常困难。

(一) 评价对象的特殊性

一是教育科研知识的复杂性。教育学研究中长期存在实证主义和现象学范式的争论, 两者在历史上有过三次大的交锋, 但现在已逐渐退潮, 有趋于相互吸收融合之势。格林奈尔(Grinnell, F.)指出, 日常的科学研究既不遵循实在论范式, 也不遵循社会建构论范式, 而是依据不同的情境在二者之间实现一定平衡, 研究发现和可信性是科研的两个主要特点。从教育学知识生产的历史来看, 教育研究知识在很大程度上是非累积性的, 它们并不是建立在之前研究的基础上, 没有通过证实或证伪、拓展或完善, 或者以更好的证据或理论取而代之, 而是在老问题上不断开辟新路, 形成平行发展的知识结构。按照英国学者伯恩斯坦(Bernstein, B.)的观点, 平行知识的发展“并不是透过整合而达成, 而是透过新语言(新概念体系)的出现, 从而构建出新视域、新问题、新联系、新质疑以及最重要的是一群新的发言人”。由于平行知识的属性, 在教育科研领域中, 无的放矢、没有问题交集的集体独白现象仍然具有一定的普遍性。此外, 由于情境和文化差异, 教育研究结论几乎不可能重复验证。美国学者麦克尔(Makel, M. C.)和普拉克(Plucker, J. A.)按照5年间的影响因子对排名前100位的教育类学术期刊的研究发现, 仅有0.13%的教育类论文是可重复性研究, 并且48.2%的重复性研究是发表最初研究成果的同一个研究团队进行的。在社会学思想和方法论的影响下, 教育研究者倾向于根据研究热点变化而不是问题解决与否, 经常变换研究主线, 从而难以像自然科学那样产出以事实为依据的累积性科学知识。近些年, 为促进教育学知识的增长, 构建中国教育学科的知识体系, 崇尚实证研究的风气逐渐形成, 教育研究者特别重视确凿知识的积累。教育科研知识的复杂性, 给教育科研评价带来了挑战, 使得科研评价要综合考虑教育科研知识的多重性和多维性。对于教育科研而言, 在后现代主义范式下, 要从扩大知识生产主体、承认知识观多元化、开放知识生产组织形式三个维度对教育学知识生产方式进行重构。

二是质量显现的滞后性。与自然科学相比, 人文社会科学的质量通常需要更长的时间才能达成共识。人文社会科学研究对象的差异性和复杂性更大, 在科学研究从特殊到一般的抽象过程中, 对学术卓越的判定更要整合时间维度和社会—空间维度上更大的差异性, 这也使得教育科研质量的显现通常需要更长的周期。在历史上, 具有真实独创性价值的东西常

常因为其非正统性而被忽视，被视为异端学说，直到另一个时代才被接受。在现有评价体系之下，突破性成果不符合评价标准之事时有发生。学术成果的价值必须经过一定的时间才能看清楚，无论是评价成果、评价学者还是评价机构，都应当留出一定的滞后时间，不能患有“近视症”。衡量学术成果的价值，最终标准只能是历史和实践。因此，教育科研质量判定的一个重要难题就是质量显现的滞后性与学术承认的即时性之间的矛盾。按照默顿的说法，既要避免因荣誉来得“太迟”而变得“太轻”，也要避免因荣誉来得“太早”、“太多”而导致研究者自满地停滞不前。

三是科研影响的间接性。教育科研在实践和政策中的应用不是工程模式，而是更多地发挥启发作用。由于教育实践的极端复杂性，教师在课堂教学中主要依靠自己的常识、经验、判断和价值观念等缄默知识，而不是基于教育科研。对于决策者而言，教育科研知识首先作用于他们的“工作知识”（working knowledge），而后通过改变此种知识间接对决策者行为发生影响。教育科研的政策影响力主要表现为概念性应用，科研成果常常以“滴灌”的渗透方式或“知识潜入”（knowledge creep）方式影响决策和实践。以素质教育为例，20年来尽管竞争性表现主义的社会现实为素质教育带来本体论、知识论和方法论上的问题，但它是承载着中国人对“好的教育”的愿望最简练、最适切的合法表述，并作为一种理念和合法化话语潜移默化地影响教育决策和教育实践。由于政策逻辑和学术逻辑的不同，好的教育研究不一定会转化为教育政策。教育科研的这些特征决定了教育科研评价不能也不应当简单套用自然科学的评价模式。

（二）评价标准的模糊性

一是评价标准随社会政治和知识范式的变化而变化。教育研究的对象处于多样化和变动的政治社会情境中，不同的利益相关者对教育科研知识的类型和供给方式提出了不同的要求，外部的决定和程序直接或间接地界定了科研质量标准，并影响研究者的方法选择。以英美等国的教育研究为例，20世纪60年代，对教育不平等问题的关注和社会科学的影响，好的教育研究体现在对社会弱势群体的教育社会学研究上；20世纪80年代对学校“效能”的关注使得教育经济学研究成为科研的典范；进入21世纪，美国更是强调基于证据的科学的的教育研究，并将其作为评判教育科研质量的主要标准。教育研究与教育实践一样，它的变化看起来更像艺术或时尚品位的钟摆变化，而不像科技进步那样的逐步改进。

二是对特定性和个性化知识的评价标准更难把握。教育研究是一门最难的科学研究，研究者常常需要掌握特定性或本土化的知识，而最容易的科学研究则追求更加普遍性的知识，显然，充分理解极其复杂的特殊性要比归纳不同情境中的一般性更难。好的教育研究不仅仅是提供普遍性的知识，而要对具体的教育情境进行分析，找准每种特殊情景中教育发展的敏感因子和钝感因子。教育学的研究对象是人与人之间的相互影响，其复杂程度显然处在最高一级水平。教育学所探索的不仅是必然性，而且可能更多时候是或然性。复杂系统的一个固有性质是，它在微观尺度上的非线性动力学和对于初始条件的敏感性依赖，不允许人们对于系统的终态作出预见。因此，对于教育结果的评价就不可能严格按照既定的目标进行考核，而应以一种动态的、模糊的评价方式来鉴别潜在的可能空间。

三是教育研究的复杂性和多维性增加了评价的难度。默顿对83种学术杂志的稿件拒稿率的研究发现，人文学科方面的拒稿率最高，其次是社会科学和行为科学，再次是数学和统计学，物理学、化学和生物学的稿件拒稿率最低，只有人文科学的三分之一。拒稿率高说明，编辑和评议人与那些撰稿人之间，对于什么是合乎标准的学术成果几乎是各执一端。尽管美

国等把基于证据的研究作为教育科研质量的判定标准，但不可否认其他研究方法仍然可以产出高质量的科研成果，正所谓“条条大路通罗马”。美国学者福伊尔（Feuer, M.）等人也指出，研究方法本身没有好坏或科学与否之分，对科研质量的判断主要看某种方法对于特定问题的适切性。教育科研与教育领域一样是一个复杂的生态系统，理应包含多种研究方法，这给教育科研评价增加了难度。

（三）评价制度的矛盾性

一是推动创新与排斥创新的矛盾。科研评价体系旨在给予知识创新者应得的承认和肯定，从而激发更大的创新活力。但制度化和体制化的学术评价制度是按照科学的常规范式，对作为反常行为的创新进行评价，在常规范式过于强大时，这种“反常”常常被排斥在外，使学术创新得不到应得的承认。在制度化的评价体系中，评价者扮演“守门人”的角色，要通过评价实现学术分层和学术区隔，维持不平等的学术等级秩序，真正的学术创新很难通过评价得到承认。同时，教育研究的“平行知识”特征使教育科研评价催生出诸多标新立异的伪创新。麦克尔和普拉克的研究也表明，研究教育事实比追求标新立异的创新更加重要，过分追求和倡导原创性成果，是造成今天不可重复的研究越来越多的重要原因。为追求新颖，浅尝辄止的研究和跟风式的研究越来越多，而专注某一个问题的长期追踪研究却越来越少。可以说，大家都去挖坑了，结果金子不一定越来越多，但坑一定是越来越多。学术评价是评价者与被评价者之间的对话和争鸣，缺少这种互动，科研评价也就背离了初衷。

二是内在评价与外在评价的矛盾。科研评价包括对内在品质的评价及对外在表现的评价，默顿将前者称为工具性承认，将后者称为荣誉性承认。由于内在品质难以把握，评价者倾向于把外在表现作为内在品质的代理变量，但是二者在很多情况下并不同步或等同。正如在大学评价中内在卓越与外在声誉的关系一样，后者是前者的“症候”，尽管声誉好的大学通常的确具有卓越的品质，但显然外在声誉不能确保内在卓越。教育科研质量有时会依据外在的学术声誉、学术头衔和刊物级别来判定，尽管高质量的教育科研通常具有这些“症候”，但具有这些外在特征并不能确保内在的科研品质。

三是技术理性与整体判断的矛盾。出于绩效评价、科研竞争等技术治理的需要，科研评价会追求技术的效率，强调科研的可计算性、可比较性，并把不同类型的科研进行化约，按照相同的尺子进行评判和测量。这种技术理性导向的科研评价制度会催生出光鲜亮丽的科研数据，但却丧失了科研的灵魂。教育科研质量评价更多的是一种审美而非科学活动，它更像是一种鉴赏而非测量。科研质量评价如同欣赏一个人一样，要察觉出不依赖于外在装饰的“纯粹”品质；对质量的判断也不是通过定量的测量，而是建立在对被评价者充分了解和细微观察的基础上，如同没有身高体重数据同样能判断人的身材；对质量的总体印象显然也不能通过各部分加总而得出；识别出这些质量能够给拥有者带来愉悦的反应。

三、科学合理的教育科研评价体系如何构建

人文社会科学评价体系可以从三个层次来构建，一是核心层（对机构、学者和成果的评价），二是环境层（学术理念、学科结构等），三是社会层（政治经济结构、意识形态等）。构建教育科研评价体系必须秉持整体主义的观点，正如戴维斯（Davis, A.）所说，不能孤立地看待和使用评价指标，一个评价指标的使用取决于对相关指标的理解及其匹配性的认识，反之亦然，这就是解释学的循环。科研质量通常包括四个维度，即内部的可靠性（严谨性、一致性、连贯性和透明性）、外部的有用性（原创性、相关性和可推广性）、形式的逻辑性（可接受性、可理解性和可获取性）和研究的规范性（合规定性、伦理性和可持续性）。具

体而言,教育科研质量的一般标准包括严谨性、原创性、有用性、重要性、影响力、伦理性、规范性和科学性。这些质量标准相互关联,没有原创性,严谨性就显得机械刻板;不考虑学术价值,原创性就变得微不足道;而学术影响力也要通过学术价值来评价。总体上,推进科研评价改革,需要重视五种评价。

(一) 实施分类评价,回归评价初心

科研评价的鉴别、区分和排名功能建立在对科研内容质量和创新的准确判定上,过于强调科研的形式评价和效用评价无异于缘木求鱼,本末倒置。《旧金山科研评估宣言()》建议停止使用期刊影响因子等期刊计量指标来评价单个论文或学者的贡献。《科学(Science)》杂志也认为影响因子可能妨碍创新,它使科学家热衷发表高影响因子论文,追逐科研“热点”,而不是潜心科研创新。教育科研评价要超越“工具理性”,回归以评促改、以评促建的初心与使命,突出创新在科研评价中的核心地位。一是要针对不同研究类型设定不同的评价标准。基础研究、政策研究和实践研究的知识创新分别以“真、善、美”为目标,通常有不同的内容和表现形式,评价的尺度和标准要适应各自的特点和规律,不能笼统地采用统一的数量标准或形式标准(如论文数量或学术化程度)。二是科研评价要与人才评价和机构评价有所区分。对人才和机构的评价,除了学术评价之外,还要有针对性地增加其他维度的评价,比如大学教师的教学能力和师德、高等学校学位点的教学质量等。要适当降低科研评价在人才和机构评价中的比重,要各类评价回归各自初心,防止科研评价绑架甚至主导人才和机构评价。三是剥离科研评价的资源分配和利益调节功能。科研评价之所以备受重视并逐渐异化,就是承载了过多的非学术因素。在机构和个人的科研活动越来越依赖于外部资源和环境的条件下,学术共同体内部的自发的科研评价就逐渐让位给外部的学术干预或控制,容易出现“外行评内行”或非学术目的的评价。学术评价只有回归学术初心,才能获得社会的合法性基础。

(二) 完善同行评价,发挥权威认定

文献计量等定量评价方法简单易行、便于操作,且确凿的数据也更具有说服力,而同行专家评价则相对困难和成本更高,但科研评价追求的不仅仅是效率,更重要的是准确与公平。“数字会撒谎”以及定量评价的形式主义弊端,呼唤同行专家评价的回归与作用发挥。一是将同行评价建立在定量和定性的综合评价基础之上。评价标准的主观性和模糊性,是造成同行专家评审公信力不高的的重要原因。教育科研质量评价标准要适应教育科研活动的特点和规律,同时要按照分类评价的原则,针对不同类型、群体和分支学科制定各有侧重的评价标准,从形式、内容和效用等多个维度确定教育科研质量的评价标准。科研质量评价的最优方法是建立一套适应学科特点的定量指标体系,同时对学术成果进行同行评议,在此基础上,以评审小组的形式对科研质量进行综合的专家判断。二是探索建立科研奖项提名制。2019年年底通过的《国家科学技术奖励条例(修订草案)》提出,要将过去主要由单位推荐改为专家、学者、相关部门和机构等均可提名,并强化提名责任。提名制将更多发挥学术共同体和专家作用,尽可能减少行政部门对科技奖励的影响干预,诺贝尔奖、图灵奖等国际大奖都采用提名制。包括教育科研在内的人文社会科学研究成果评奖可以探索建立提名制,更大程度上发挥学术共同体和专家的积极性、主动性,遴选出真正具有学术原创性的科研成果。三是优化同行评议的人员结构,兼顾广泛认可和权威认定。华东师范大学组织的人文社会研究“思勉原创奖”通过知名学者推荐与同行大众“海选”相结合的遴选方式,兼顾了学术共同体金字塔结构的两端,实现了“同行评议”在数量讲求上的最大化和质量讲求上的最高级,从制度层面使推荐过程和结果经受学术共同体的检验,克服学者个人视野的局限,充分发挥了学术

共同体的作用。在学术权威欠缺的情况下，教育科研评价要通过优化同行专家组成，使评价结果获得更广泛的认可。

（三）强调综合评价，减少评价偏差

一是完善代表作评价制度。针对科研评价中“只数数量和影响因子、不看论文质量”、“重国外期刊轻国内期刊”等问题，北京大学等知名高等学校在职称评审中试行学术代表作制度。但人文社会科学研究复杂性、评审专家的认知偏差、学科划分、同行选择、人情关系等因素对代表作制的评价效用产生影响。当前，完善代表作评价，需要处理好同行评议与定量评价之间的关系，构建科学的评价治理体系及优化学术环境。二是建立个性化、发展性评价机制。评价最重要的意图不是为了证明，而是为了改进。我国科研评价问题的症结在于追求各项指标上的“比较”，而不重视发展性、开创性研究的实质性进步。人文社会学科更多受到了学科立场、价值偏好的影响。对于一些“冷门绝学”或基础性、开创新的研究，现行的评价标准通常会把它们排斥在优秀成果之外。对于这类研究，就不能以“显学”的评价标准来对待，而是要从“学术”与“学科”的关系、“有用”与“无用”的相互支撑等角度，考察这类研究对于促进学科体系的系统性、完整性和保护民族文化的独立性、自主性的重要价值和意义，特别要遴选熟悉学科规律、具有战略眼光的学科专家进行评价，防止此类研究在常规的评价中被弱化甚至被淘汰。教育科研评价不能仅仅满足在评价指标上的证明和比较，更应该为原创新研究开拓学术认可的空间。三是建立全过程的科研评价体系。教育科研评价不能仅仅关注作为结果的学术成果数量和质量，而是要从条件质量、内容质量和成果质量几个方面做出综合评价。动态的科研评价可以把科研活动放在一个更广阔的评价框架中，从全要素生产的角度更加客观全面地评价科研的潜力、成绩与不足，真正实现“以评促改”的目的。

（四）推进多元评价，丰富评价方式

一是要有“多种组合”的科研评价方式。要突破一元化评价范式，探索实施定量评价与定性评价、同行评价和社会评价、过程评价和结果评价、当前评价和长远评价相协调的多元化评价范式，通过多种评价方式的组合，全面、整体地反映科研的质量。二是要用“多种尺子”评价不同的教育科研。要克服科研评价中的“一刀切”，根据不同的研究类型，科学设置精准的评价标准和评价方法，如基础研究侧重于知识创新，应用研究侧重实践改进，青年项目评审侧重于科研潜力，要根据评价目的选择最适应的“尺子”。三是要有“多种技术”支撑的科研评价方式。教育科研评价要充分利用区块链、大数据、社会网络等现代技术，创新科研评价的方式方法，从而更加全面、系统、客观地展现科研人员和科研活动的发展过程和影响因素。

（五）健全公开评价，构筑良好生态

科研评价乱象的根源在于良好的科研生态尚未完全建立起来。改进教育科研评价，需要完善科研治理体系，构筑良好的学术生态。一是要提高同行评价的科学性和公开性。要通过制度创新和技术创新，最大限度地降低学术偏好、人情关系和长官意志等因素对科研评价的干扰。不断提高同行评价透明度，完善科研评价实名制和责任制，回归学术评价的原点。二是健全科研评价的社会监督和申诉机制。不断加强制度体系建设，进一步完善评价结果公示和申诉等机制，逐步建立健全信息披露、责任追溯、风险管理等机制与制度。明确学术不端行为的判定标准和处理程序，科学划分各类机构在处理学术不端行为方面的权责范围。三是优化教育科研的学术环境。教育科研的政策环境要为原创性成果的产出创造有利条件，在资

源分配、人才评价和机构评估中重视基础性、原创性研究，要通过评价引导各类研究各安其位、各得其所，坚决破除教育评价中的“五唯”现象，使科研评价真正发挥鉴别、诊断、规范、引导和激励作用。

教育科研评价体系实际上是教育科学领域中的“运行法则”，各种可见的和潜在的教育科研活动都受到了这种“法则”或“实践逻辑”的支配。由于在教育学科本体论、认识论和方法论上的分歧，教育科研评价长期以来难以达成共识，在碎片化的知识体系中，并行着各类集体独白式的质量诉说。改进教育科研评价体系，需要在科学文化和科学信念的大前提下，进一步澄清教育科学知识的复合型、复杂性和特殊性，找准不同类型、不同情景中教育科研的敏感因子并制定个性化的分类标准，秉持“和而不同”的思维构建“一与多”相统一的科研评价体系。教育科研评价不是单纯教育现象，更是社会学现象，评价体系的构建需要正确应对各种内在的矛盾性，合理解决评价的“测不准”问题以及结构化带来的负效应。只有正确认识并妥善化解这些问题，教育科研评价才能走出目前的困局。

【孟照海，全国教育科学规划领导小组办公室副主任，中国教育科学研究院副研究员；刘贵华，中国教育科学研究院副院长、教授。】

——来源：《教育研究》，2020年第10期

高等教育评价中的价值冲突与融合

李庆丰 周作宇

【摘要】我国高等教育大众化以来形成了基于不同利益相关者价值取向的高等教育评价多元话语体系，反映了高等教育评价主体与评价标准由单一走向多元的社会事实，也反映了高等教育评价范式由认识论向价值论的转变。价值论域下高等教育评价的本质是基于价值事实的价值判断，必然产生不同评价主体之间以及评价主体与评价对象之间的价值冲突。对此，应以“育人为本”作为高等教育评价的核心价值标准及价值融合的方向。

【关键词】高等教育大众化；高等教育评价；价值冲突；价值融合

2020年6月30日召开的中央全面深化改革委员会第十四次会议通过了《深化新时代教育评价改革总体方案》，会议指出，教育评价事关教育发展方向，要全面贯彻党的教育方针，坚持社会主义办学方向，落实立德树人根本任务，遵循教育规律，针对不同主体和不同学段、不同类型教育特点，……建立科学的、符合时代要求的教育评价制度和机制。有什么样的指挥棒，就有什么样的办学导向。就高等教育学段而言，随着我国高等教育后大众化和普及化时代的到来，高等教育已居于社会的中心，同时“上好大学难”等问题日益凸显，使得高等教育质量和评价受到国家、市场和社会的广泛关注和重视，针对高等教育评价实践也出现了不同的声音和话语，“不同群体利益诉求的价值取向呈现多元化趋势，对教育问题的关注之广泛、抨击意见之尖锐、观点建议之多样、价值冲突之激烈，为其他领域所鲜见”。在多元利益取向的话语背景下，如何正确认识这些矛盾与冲突，就成为科学地化解和平衡矛盾冲突，建立科学的高等教育评价机制，推动高等教育评价实践科学发展，持续改进高等教育质量保障体系的一个重要理论问题。

一、高等教育评价的话语背景

1. 国际竞争与创新的压力

2005年7月30日，钱学森对时任国务院总理温家宝说：“现在中国没有完全发展起来，一个重要原因是没有一所大学能够按照培养科学技术发明创造人才的模式去办学，没有自己独特的创新的东西，老是‘冒’不出杰出人才。这是很大的问题。”这就是著名的“钱学森之问”：“为什么我们的学校总是培养不出杰出人才？”与此相应，早在20世纪20年代就警醒国人的“李约瑟之谜”依然清晰：“为什么近代科学和科学革命只产生在欧洲？为什么直到中世纪中国还比欧洲先进，后来却会让欧洲人着了先鞭？是什么造成了这样的转变？”

2010年我国GDP总量超过日本居世界第二时，“21世纪是中国的世纪”的话语就腾空而出。考察新时代经济全球化背景下中国作为世界工厂在经济产业链中的真正位置及经济总量构成，应该说，我们在有些领域已从跟跑追至并跑乃至领跑，但总体而言还是大而不强。在许多产业领域，如果说每一件产品都要经历从设计、生产（包括设备、技术、材料与组装）、流通最后到用户的环节，每个环节都有价值生成与分割的话，显然我们是居于价值链最底端的，利润绝大部分都流向了专利、品牌和销售环节，特别是专利与品牌。近年来持续不断的中美贸易摩擦也与此有关，其中基于核心技术的专利和品牌更是成为贸易制裁的利器。

回望人类社会发展阶段和国际竞争态势，从农业社会到18世纪后期以牛顿力学和蒸汽机广泛使用为代表的蒸汽机时代，到19世纪后期以电磁学和电力应用为代表的电气时代，到20世纪中期以相对论、量子论和半导体、计算机、互联网应用为代表的信息时代，再到当前以

大数据、云计算和智能机器人为代表的智能时代，所有这些发展阶段，无不是以科技创新为驱动的。这个过程造就了 18 世纪英国的勃兴、19 世纪德国的领先、20 世纪美国的崛起。那么，21 世纪会是谁的世纪呢？有人说，21 世纪是中国的世纪。21 世纪会是中国的世纪吗？

美国兰德公司在 20 世纪 80 年代初曾提出，一个国家没有经济独立就没有政治独立。时至今日，这一说法依然成立，但还应该加上，没有科技独立就没有经济独立，从而也就没有政治独立。2004 年美国国家竞争力委员会指出，创新是决定美国在 21 世纪成功的单一最重要的因素。国际竞争就是创新之争，此为“李约瑟之谜”和“钱学森之问”所直面的问题。“钱学森之问”涉及一所大学如何才能培养出杰出人才，从而促进科学技术的创新；“李约瑟之谜”涉及一个国家的科学技术如何才能进步，从而走在世界的前列。两问虽然都引起了广泛的探讨，也无统一的答案，但教育特别是高等教育，都被认为是其根本问题之一。两问的共同本质是：中国如何发掘与发挥创造性人才的社会功能与价值，也即如何培养创新人才。这是高等教育评价所必须面对的一个沉重的话语背景或社会场域。

2. 扩张的动力与质量挑战

回顾新中国成立以来高等教育的发展可以发现，高等教育发展规模经常发生起伏变化，尤其表现在高校数量上。20 世纪 50 年代，因“院系调整”和“教育大跃进”，高校数量急剧扩张；60 年代因“文化大革命”的破坏，高校数量急剧减少；改革开放后的头十年，高校数量大幅增加；90 年代因“高校合并”，高校数量有所萎缩；进入 21 世纪后，因为“扩招”，高校规模扩张。就近考察自 1999 年以来的新一轮高等教育规模扩张则可以发现，随着招生人数的激增，办学质量下降、办学定位趋同、特色不彰的问题突显，引发了社会的广泛抨击和忧虑。

对此，21 世纪初以来，国家相继启动了“普通高等学校本科教学工作水平评估”（2003）、“高等学校本科教学质量与教学改革工程”（2007）、“高等学校创新能力提升计划”（2011）、“双一流”建设（2015）、“一流本科专业建设‘双万计划’”（2019）等。这些国家重大工程、计划和方案，都是在国家层面对从高等教育大国向强国迈进、全面提升高等教育质量所作的积极回应和具体实施。与此同时，全社会也都满怀热情，异常关心、关注、关切高等教育。政府对高等教育的投入不断增加、认识越来越深入，大学也雄心勃勃，提出争创“世界一流”、“国际知名”等激动人心的奋斗目标和发展战略，社会对高等教育的关切不断增强，以至高等教育成为出镜率与曝光率都很高的“新常态”。

一方面，全社会对全面提升高等教育质量、推进高等教育从大向强有些“迫不及待”，因而对大学倾注了过多的热情与关切，街头巷尾、茶余饭后，话语之间多多少少总会涉及大学，并往下延至中学、小学乃至幼儿园。但过多的感情和过度的关切，在某种程度上无异于拔苗助长，常常使得大学疲于应付，而真正用于提高质量的时间、精力与经费却少了，更难以静下心来思考涉及未来发展的科学前沿和重大问题。另一方面，据教育部的统计数据，2008-2018 年间，我国各类出国留学人员年均复合增长率达到 13.9%，2018 年达到 66.21 万人，且这种现象正呈向基础教育下移的趋势。虽然 2020 年以来这种趋势因受疫情的影响而有所放缓，但这在很大程度上仍给我们敲响了警钟：如果我们不能提供优质的高等教育，那么适龄学生及其身后的家长将继续用脚投票；如果社会一直这么浮躁、急迫，过度关注甚至苛求、压迫高等教育，大学就无法沉下心来面向未来有序地做好当前的每一件教育实事，高等教育强国和创新型国家的建设无疑也会遥遥无期。

从高等教育本身发展和创新人才成长的规律来看，二者都需要有合适的土壤和环境。“十

年树木，百年树人”，早在1924年，鲁迅先生就指出，在未有天才之前，就让我们都来当培育天才的土壤。在全面提升高等教育质量和建设高等教育强国及创新型国家的过程中，我们是继续“跨越式”、“赶超式”、“跃进式”发展，还是静下心来脚踏实地、一步一个脚印向前走，能否不再出现以往70年高等教育发展历程中的大起大落现象？这是高等教育评价面对的又一个社会话语现实背景。

3. 大学的透明度与公信力

大学是社会政治、经济、文化等因素共同作用的产物，社会需要是其生存和发展的唯一理由。一方面，大学接受了政府的巨额资助，而这些资助均来自纳税人，因此，大学有责任回报社会，帮助解决社会问题；另一方面，大学往往作为一种公益属性的社会机构存在，如OECD就将大学划分为公益机构，我国则将其作为事业单位。所以，不管是人才培养、科学研究还是社会服务，都具有强烈的社会性，或者说强烈的公益性或公共性，这是大学的基本特征。这样，大学不仅作为传播知识、研究学问的相对独立的文化机构而存在，而且作为承担社会责任的公共机构和重要阵地而存在，成为社会道德的楷模、公共利益的代表、人类理想的寄托及精神与文化的家园。早在1909年，卡内基教育促进基金会首任主席、MIT原校长普里切特(H. S. Pricheter)就说过：“所有高校，无论是靠政府资助还是接受私人捐助的高校，都是公共服务机构。公众有权了解它们在管理和发展方面的事实，包括财务和教育方面的事实。”2007年世界高等教育年度报告特别突出了“大学的社会责任”理念；2009年世界高等教育大会更是将“高等教育的社会责任”作为重要的会议主题；2011年召开的大学校长全球峰会明确阐述了在未来世界经济发展中高等教育所应发挥的作用与社会责任。近年来，社会更是呼吁大学应该承担更多的责任，如将可持续性和共享纳入大学的社会责任范畴等。这就要求大学的自身运行机制及发展状况要向社会开放，要向社会表明自身承担了什么样的社会责任，所承担的社会责任的完成程序等。也就是说，大学应该是开放和透明的，需要将自己所承担的社会责任及其完成情况昭告天下。

大学也是知识、高品位文化与价值观的权威，几乎垄断了诸如授予学位、表征接受了高层次教育等方面的有价值的资源，而学位和学历不仅是许多令人渴望的职业所必不可少的，而且在大众化和普及化阶段更是与普通老百姓的个人利益紧密联系在一起，在某种程度上已经成为现代社会中个人身份的象征。甚至有学者认为，知识社会是一个文凭社会，大学文凭是通向知识社会的首要通行证。所以，大学的这种权威身份在不断影响和控制着个人乃至整个社会的律动，以知识占有的层次来不断构筑社会的层次及其合理性，通过进行高层次的教育影响和造就社会成员的精神气质，以此传承文化与精神传统，并对社会现实进行批判，从而对社会的主流文化进行影响和引导。这也就赋予了大学在现代社会中具有高公信力的人格特质。

然而，近年来大学的改革与发展常常成为全国“两会”的热议话题，也饱受社会抨击。这一方面说明大学已从社会边缘走向并稳居社会中心，全社会各层面、各类群体、家庭与个体等都非常渴望把大学办好；另一方面也说明大学自身这些年的改革与发展还不能令社会完全满意，离成为人民满意的大学还有距离。这主要表现在两个方面：

一方面，大学的透明度和开放性不够。自2010年9月1日《高等学校信息公开办法》施行以来，信息公开逐渐受到大学的重视，教育部网站公布各直属高校的年度信息公开报告，各地方高校的网站上也基本上能查到其信息公开情况。但据相关研究，目前各校有限的公开制度还很不完善，仅对学校信息进行有限公开，且不易找到；而对不予公开的内容规定得又

过于笼统，且具有很大的自由裁量权，因而难以保证社会公众大学信息获取权的实现。事实上，信息公开是大学管理部门应尽的一项义务，更是现代大学治理的一项重要内容。可能有些学校担心，如果公布所有信息，它们就无法与其他学校竞争；但反过来说，维护社会公众的知情权也很重要。如果信息不公开，不仅有违国家确立的依法治校理念，还会对大学自身的信誉度造成一定的影响。

另一方面，大学的公信力在下降。教育理念、管理制度、办学行为是大学社会公信力的基本要素，高等教育服务及其质量则是大学社会公信力的核心要素。随着高等教育后大众化的到来及普及化的推进，社会感觉高等教育的公信力在下降，其原因在于：一是本不应为大学所为之事却屡屡发生，如腐败、学术诚信等。正因为高等教育的社会公益性、公共性，决定了大学机构有些事情不能做，有些事情做了就会比其他机构要受到更多的社会责难。二是本应大学承担之社会责任，并未让社会感觉到大学已承担。随着我国高等教育发展从精英化向大众化、普及化的转变，大学的结构、规模、特色、质量与效益等问题日趋复杂，其传统价值观与责任意识发生了很大变化。今天的大学在一定程度上也具有了社会其他机构普遍存在的“重眼前轻长远”、“重物质轻精神”、“重经济轻文化”、“重科技轻人文”、“重操作轻思想”等浮躁倾向，弱化了大学对社会的建设性批判和创新性引领功能，并出现高等教育的不协调发展、社会责任难以履行等问题，其直接后果就是大学的社会公信力不断下降。

以上两方面构成了高等教育评价实践所需面对和解决的又一个话语背景。

二、高等教育评价的本质反思

高等教育评价领域为什么会以上不同的话语体系？不同话语的真正内涵是什么？这些不同话语背后的共同本质又是什么？

1. 不同话语反映了高等教育评价主体与标准由单一走向多元的社会事实

根据有关话语研究，任何一种语言除了应该遵守的语法规则之外，还包括一种文化规则，即“所有语言都不可避免地承载着社会关系，包括人们之间的政治关系、经济关系和各种现实的人际关系”，所以不同的话语就“不仅仅具有一种形式的意义和价值，它实际上还包含了非常丰富的社会政治内涵，体现了现实中复杂的社会关系，甚至是一定的社会结构”。不同的评价话语产生的根本原因就在于高等教育发展模式及其赖以发展的经济、政治环境乃至社会结构发生了很大变化，而不同的话语背后所反映的实质就是，不同的话语所代表的群体所持的高等教育质量评价标准不同，其本质是对高等教育价值的认识不同，背后潜藏的则是不同话语的群体对高等教育不同的价值或利益诉求。

在计划经济时期，政府、大学与社会一体化发展，政府、大学与用人单位之间没有明确的责权利划分：政府按计划投资办学，大学按计划培养学生，学生免费上学、毕业后政府包分配，企事业单位按计划接收政府分配来的大学毕业生。这样，政府成为大学、学生与社会的核心纽带，政府、大学、学生与社会之间利益目标趋同，利益关系简单，基本上以政府的利益目标为旨归。与此相适应而建立起来的高等教育评价主要是以政府为评价主体的一元化评价，大学、学生和社会没有自己独立的利益目标，主要是配合政府开展高等教育评价。

在由计划经济体制向市场经济体制转型的过程中，政府、大学、学生和社会的利益目标逐渐分化并呈现出明显的差异性：在“小政府、大社会”的政府职能和治理结构改革背景下，政府作为市场经济背景下的高等教育决策主体，逐渐成为一个有着自利倾向的组织，也会追求自身利益的最大化；大学成为“面向社会自主办学的法人实体”，获得了追求与发展自身独立利益的合法性，但需要面向市场自筹部分经费、自主办学，历经近30年的发展，如今的

大学已成为名副其实的有着自身利益追求与目标的行动主体；学生由原来的“统招统分”到如今的缴费上学、自主择业，由此逐渐转变为需要考虑投资回报的有着自身利益追求的行动主体；企业因面向市场自主经营、自负盈亏，由此逐渐成为需考虑引进人才/大学毕业生的质量与成本以追求企业利益最大化的独立行动主体。这样，原来统一的利益目标与追求逐渐分离、分化和多样化、复杂化，一元的高等教育价值取向和评价标准逐渐向多元化的价值取向与评价标准转变。

同时，在这一过程中，我国高等教育开始了由精英化向大众化、普及化发展的进程。随着高等教育由“旧时王谢堂前燕”逐渐“飞入寻常百姓家”，在国家发展、社会进步、个人成长、企业创新等方面日益重要起来，特别是在推动社会公平、形成社会有序流动、实现社会有序整合、保持社会发展活力等方面发挥了基础性作用，使得大众化、普及化时代的高等教育与社会各群体、组织、机构等产生了密切关联，并逐渐成为各方关注的焦点与中心。过去大学质量保障常常被认为是大学内部或自身的事情，现今却受到社会非常广泛的关注。这些不同的利益相关者不仅要大学的办学情况“说三道四”，用自己的评判标准对大学进行评价，甚至有资格对大学进行必要的干预了。换句话说，在计划经济向市场经济体制转型以及高等教育从精英阶段向大众化、普及化发展的过程中，高等教育的质量观及质量评价标准逐渐由单一向多元转变；不同的话语实质上反映了不同话语的主体对高等教育发展的期待，同时这种期待成为各话语主体评判高等教育质量的标准、价值乃至利益诉求。

2. 不同话语反映了高等教育评价范式由认识论向价值论转变

高等教育评价话语及评价主体和标准的多元化不仅与计划经济体制向市场经济体制转型、高等教育由精英化向大众化和普及化发展有关，而且与教育自身的内在发展逻辑以及哲学观由认识论向价值论转型有关。

从高等教育发展的内在逻辑来看，在高等教育精英阶段，大学不仅是研究高深学问的场所，更是卓越与质量的象征。彼时的高等教育以知识传授与发现及个体理智发展为宗旨，为知识而知识，知识本身即为目的，这是一种认识论视域下的高等教育质量观和评价观。随着时代精神的变迁与哲学范式的变化，作为理性主义产物的现代大学“依靠改变自己的形式和职能以适应当时当地的社会政治环境，同时通过保持自身的连贯性及使自己名实相符来保持自己的活力”，从而使大学职能逐渐由单一的人才培养发展为人才培养、科学研究、社会服务等多种职能并重，其质量评价标准也逐渐由中世纪大学“以知识的纯粹理性（独立于一切经验的理性）和知识的内在逻辑为其质量标准”发展到洪堡时代的“以‘科学而达至修养’为其质量标准”，再发展到“以知识的实用价值为其质量标准”。由此，高等教育质量评价标准逐渐由探究知识向关注服务转变，由一元价值向多元价值转变。在 20 世纪四五十年代，随着高等教育由精英化向大众化、普及化发展，大学为社会政治、经济、文化等服务的功能日益凸显，大学开始且不得不关注利益相关者的价值诉求。

大学职能的扩展、大学理念的变迁及高等教育质量评价标准的多元化发展，反映出高等教育评价实践由关注知识的认识论范式向关注利益相关者的价值论范式的转换。不管认识论范式下对高深学问的追求还是价值论范式下对经世致用的强调，其根本目的都在于满足高等教育主体的需求。在高等教育精英阶段，高等教育主体主要是大学内部的教师和学生，高等教育的质量诉求和评价标准主要是教师探究知识、学生掌握知识本身；而在大众化、普及化阶段，高等教育主体不仅包括大学的内部人员，还包括大学的外部利益相关者，高等教育质量开始关注利益相关者的需求，质量诉求和评价标准也呈现多样化趋势。不同利益相关者开

始更加关注高等教育对自身需要及价值诉求的满足程度，从而不断增强了各自的质量保障和质量问责意识。

从前述三类不同的话语可以看出，“国际竞争与创新的压力”话语主要反映的是国家或政府的价值诉求和评价标准，期待高等教育能够培养出更多更好的德才兼备的各级各类创新人才，能够研究出更多的可以直接转化为现实生产力的科技成果，能够更多更好地服务社会并引领社会的未来发展，不断提高国家的科技和人才的国际竞争力。“扩招的动力与质量挑战”话语更多反映的是大学和政府的价值诉求，大学期待通过规模扩张迅速扩大自身的影响和办学层次，同时吸引更多更好的生源来提高自身的办学质量，进而提升自身的社会声誉与影响力；政府期待通过高等教育扩招满足社会需求和提高整体国民素质、拉动经济和增加绿色GDP，同时期望通过不断提高质量推动我国由高等教育大国向高等教育强国迈进。“大学透明度与公信力”话语更多地反映出学生、家长、用人单位、捐赠者等不同利益相关者的价值诉求，他们期望大学在市场经济环境下能够明明白白办学，能够让大众明明白白上大学，能够保持应有的“大学精神”并维护其社会公共性；学生期望自身获得全面发展，教师期望自身事业有成，用人单位期望大学培养出所需要的创新人才，捐赠者则希望自己的捐赠达到预设的目的、收获预设的价值等。可见，传统认识论下的纯粹探究高深知识的高等教育质量观与评价标准已不能完全满足大学利益相关者的多元需求，而必须与时俱进，寻找新的逻辑生长点，即转向价值论。显然，高等教育不同利益相关者的不同价值诉求使得大学必须关注社会需求，关心大众需要，关切公共利益，协调和尽量满足多方面的需要。换句话说，我们要在价值论域下来开展高等教育评价的研究与实践。

3. 价值论范式下高等教育评价的本质：基于价值事实的价值判断

评价是关于对象价值的活动，是主体依据一定的目的，通过特定方法收集关于某对象在某价值领域的数据和信息，根据特定参照系对数据信息进行分析，以识别这一对象相对状态的过程。简言之，价值论域下的评价活动本质上是一种基于价值的认识活动。广义的价值是指“主体与客体之间在相互联系、相互适应、相互依存、相互影响的互动关系中所产生的效应，既包括客体对主体的效应，也包括主体对客体的效应”。关系属性、生成属性以及多维属性是其三大基本内涵特征，同时也构成了分析价值问题的方法论，即在分析价值活动中，必须同时弄清楚并回答“有无价值，谁的价值，什么价值，多大价值，何以有价值”五个问题。对高等教育评价的研究与实践也应遵循价值问题的分析方法，即关系的、生成的和系统的思维方法。

据此考察前文所述不同话语体系的本质，在某种程度上即是回答以上五个价值问题的一种价值认识活动。具体来说，就是作为高等教育评价的不同主体，包括国家/政府、大学、社会等，根据各自不同的价值取向，确定评价的目的，再通过各自所能采取的方法收集关于高等教育在他们各自所选择的价值领域中的数据、信息，并与他们选定的参照系（评价标准）进行比较，从而对高等教育价值的有无、大小等进行价值判断的过程。正是因为不同评价主体的价值取向不同，导致评价的目的、方法、数据和信息、参照系（或评价指标体系）都不相同，最终导致对高等教育这一评价对象的价值判断和结果不同，也就导致评价过程中不同评价主体价值取向之间的冲突。例如，国家或政府对高等教育的主要价值取向和评价目的是希望在创新人才培养和科技创新方面提高国际竞争力，维护国家和社会公共利益，其制定的指标体系、采取的评价方法、收集的评价数据等都围绕这一价值取向和目的而展开；同样，大学自身的价值取向和评价目的是追求可持续发展的能力和态势，保证学术自由；社会整体

上的价值诉求是追求高等教育的公共服务能力，即大学是否彰显社会公共性，同时作为具有市场属性的社会又要求大学优化资源配置，提高资源使用效率。

从哲学层面来看，这种主体之于客体（评价对象）的价值诉求及其结果构成了主客体之间的价值事实，因此，评价本质上就成为“一定价值关系主体对这一价值关系的现实结果或可能后果的意识”，它“是以一定的价值事实为对象的反映”。高等教育评价的本质就是基于价值事实的价值判断。换句话说，高等教育评价不是一种单纯收集自然事实（物、事、人）并对之作出评价的认识活动，而是在自然事实与价值事实基础上作出价值判断的认识活动，价值事实是高等教育评价这一认识活动中最为主要的认识对象。因此，在高等教育评价中，评价主体不仅需要收集反映高等教育这个客体的基本情况数据及高等教育对利益相关者基本需要的满足情况等方面的数据，而且更需要收集高等教育利益相关者与高等教育相互关系和相互作用的实际过程及其结果所产生的价值事实方面的数据。

三、高等教育评价中的价值冲突

在高等教育由精英化向大众化、普及化发展的过程中，高等教育的评价主体和评价标准由单一走向多元，多元化的评价主体和评价标准，指向同一的评价对象，其结果必然是不同评价主体之间及主体与对象之间的矛盾与冲突不断。简言之，高等教育评价中的价值冲突是指评价主体在对评价对象进行评价时，在主体之间、主体与对象之间产生的关于评价权力、评价目的、价值判断和评价效用的认知、决策和利益上的分歧和对立。

1. 高等教育评价主体与评价对象的构成

从理论上考察，高等教育的利益相关者均可成为高等教育评价主体，有权对高等教育开展评价，一般认为，我国高等教育的主要利益相关者从宏观上看包括国家/政府、大学和社会；从微观上看，在国家层面包括中央政府和地方政府，在大学层面有学生、教师、高级行政人员，在社会层面有企业等用人单位以及社会组织、家长等，他们对高等教育均有自己明确的价值追求或利益诉求。他们是高等教育的主要利益相关者，同时也是高等教育评价的主体。

那么，作为高等教育评价的主要认识对象——价值事实又在哪呢？高等教育中的价值事实“存在于高等教育利益相关者和高等教育相互作用的结果之中”。从宏观层面来看，高等教育的主要利益相关者是国家、大学与社会，则高等教育中的价值事实就体现为国家、大学和社会与高等教育之间相互作用的结果；从微观层面来看，高等教育的主要利益相关者包括各级政府、教师、学生、行政人员、用人单位、家长等，则高等教育中的价值事实就体现为这些不同利益相关者与高等教育之间相互作用的结果。如在本科教学评价的第一轮水平评估和第二轮审核式评估中，学生和教学之间的关系是最基本的一对主客体关系，学生是价值关系的主体，教学是其客体；学生和教学之间相互作用的结果所产生的价值事实是本科教学评价过程中所应把握的最基本的价值事实。换句话说，本科教学评价中所要认识和把握的价值事实是政府、大学（主要包括教师、学生）和社会（主要包括家长、企业等用人单位）等高等教育利益相关者对“学生和教学之间相互作用的结果”所形成的各自的认识和价值判断。

2. 高等教育评价主体之间的价值冲突

高等教育评价主体之间的价值冲突主要表现为评价主体在对评价对象进行评价的过程中，关于评价权力、评价目的、价值判断和评价效用等方面的认知、决策和利益上的分歧和对立。

首先，各评价主体对评价对象的价值判断存在冲突。也就是说，宏观上国家、社会和大学，微观上各级政府、用人单位、家长、教师、学生等作为高等教育的利益相关者，都对高等教育满足自己的功能需要或自身对高等教育的价值取向有着不同认知甚至是冲突，具体的

价值判断或取向上的差异在上文已有论述。评价主体的不同认知与多元价值需要，决定了冲突的复杂性，也正是这些不同的价值判断与取向，构成了不同的评价话语体系。简言之，在大学的教学、科研和社会服务三大基本功能中，有的主体更强调大学的教学功能对自身需求的满足，如学生、家长；有的更强调大学的科研功能，如国家或政府、大学；有的更强调大学的社会服务功能，如用人单位或社会群体、组织等。

其次，各评价主体关于评价的功用和目的存在价值冲突。正是基于各自对高等教育功能的认知和期待上的差异，导致各利益相关者对高等教育评价的功用与目的的认知存在冲突。这主要表现在两方面：一方面，从横向上看，高等教育评价本身就是一种多元价值的存在，既包括教学、科研与社会服务，也包括公平、正义、理性、法治、权利等，这些价值都拥有独特的内涵和意义，它们时而包容，时而并立，时而对立，并在不同的评价主体那里以不同的结构形式呈现出来。有的认为教学即人才培养的价值最高，有的认为公平与正义的价值最高，这就决定了高等教育评价价值取向冲突的复杂性。另一方面，从纵向上看，高等教育评价的功用大体可分为两类，一是“描述—诊断—改进”模式，强调以学生成长为优先的评价功能，二是“描述—鉴定—奖惩”模式，强调以利益为优先的评价功能。对评价功能的认知差异直接导致对评价目的的取舍不同，成长优先的“改进模式”更强调评价过程中的用户导向与参与者导向，利益优先的“奖惩模式”更强调评价过程中的目标导向与管理导向。

最后，各评价主体在评价权力的分配上存在冲突。从理论上讲，不管是宏观层面的国家、大学与社会，还是微观层面的各级政府、教师、学生、行政人员、用人单位、家长等，高等教育利益相关者都可以作为评价主体。然而受计划经济体制下以国家或政府作为单一价值取向和单一评价主体的传统惯习的影响，虽然在市场经济体制转型过程中，高等教育评价的价值取向逐渐多元化、评价标准逐渐多样化，但从总体上看，评价权力主要还是集中在国家或政府手中。宏观上国家依然拥有更多的评价权力，而大学和社会作为评价主体的地位不显：大学往往以被动接受评价的对象出现，积极性与主体性没有得到应有的发挥；社会组织开展的各类高等教育评价虽然有了很大的发展，但作为高等教育评价的重要社会主体——用人单位和家长却没有或少有参与。微观上政府和大学行政人员具有更多的评价权力，而学生、教师、用人单位、家长对高等教育评价的权力没有得到应有的保障和彰显。这些矛盾与冲突在第一轮本科教学工作水平评估中充分地呈现出来，也是其广受诟病的原因之所在。

3. 高等教育评价主体与评价对象之间的价值冲突

在评价过程中，高等教育评价主体与评价对象之间会产生关于评价权力、评价目的、价值判断和评价效用等方面的认知、决策和利益上的分歧和对立。造成这种冲突的根本原因在于高等教育评价对象的特殊性，即高等教育评价对象（高等教育作为客体与各利益相关者作为价值主体所构成的价值关系及其结果）中作为需求满足方的价值主体（如学生、教师、行政人员等），一方面是作为评价客体（如学生构成及质量、教师构成及质量、行政组织构架及质量等在高等教育评价指标占据重要位置，是评价的对象）的组成部分而存在，另一方面又可以作为评价主体（即学生、教师、行政人员作为评价对象的价值主体，自身对评价过程中的价值关系及其结果也有自身的价值判断）而存在。高等教育评价就是评价主体对由构成价值关系的价值主客体之间互动而产生的结果进行价值判断。这样，评价关系中的主体与价值关系中的主体之间就可能存在基于不同高等教育价值观的价值冲突。为此，第二轮审核式评估在技术和程序上对这种冲突进行了一定的协调。

在市场经济体制下，国家（政府）、大学（教师、学生和行政人员）和社会（用人单位、

家长)等同属于评价主体,各自均有自身独立的价值诉求和利益目标。但由于行政权威还在相当程度上存在,国家或政府往往主导着评价的走向,具有更多的评价话语权,而大学与社会的评价权力没有得到应有的彰显,从而使得评价主体之间的冲突更多地表现为主客体之间的冲突,即评价主体与构成评价对象的价值主体之间的价值冲突。这样的冲突主要体现在对大学功能及评价目的的价值判断上,对此上文已有论述。

另外,高等教育评价主体与评价对象之间的价值冲突还表现为对评价内容在价值判断上的冲突。前文已述及,评价的主要对象是主客体或物、事与人相互作用而产生的价值结果或价值事实,而不仅仅只是单纯地考量高等教育中的物、事和人。比如,在我国首轮本科教学评估实践中,评价指标非常重视学校的办学条件如教室、实验室、校园面积,生均经费、生均实验仪器设备值,生师比、教师学历结构等,通常认为,只要有好的办学资源与条件,就有高的办学质量。这就是单纯的评价物、事和人,而没有从构成价值关系的主客体相互作用的结果及事实出发,少有甚至没有考虑师生对学校资源和条件的使用及受益程度。换句话说,资源与条件好只是大学办学质量高的一个单向的甚至在某种程度上说是一厢情愿的充分条件而已,并不构成双向的充要条件。这也能比较好地解释为何现在很多大学办学资源与条件均是一流了却并没有培养出一流的学生,相反大家却经常提及和怀念抗战时期办学条件简陋到不能再简陋却培养了一大批创新人才的西南联大。这也成为以往评价中不同层次和类型的高校并不满意的原因,因为至少到目前为止,大学所具备的资源与条件不是其自身所能决定的,而更多地受制于我国高等教育投资体制及其历史积累。这是评价过程中“见物不见人”的典型现象,其实质就是轻视或无视主客体所构成的价值关系所形成的现实结果是高等教育评价的主要认识对象。与此相对应的是评价过程中“见人不见物”的现象,即只考虑评价过程中价值关系主体的需要和情感,而对主客体所构成的价值关系的互动结果视而不见。比如,大学普遍实施的学生评教片面关注学生对教师“教”的感受,而忽视“教”本身对“学”所带来的客观影响或结果,同时也忽视了“教”的效果本身还受制于“学”的程度,当然也忽视了互为主客体的教师对“教”与“学”的需要与感受,这也成为学生评教中教师被动参与或消极对待,同时学生评教广受抨击的一个重要原因。

四、高等教育评价的价值融合

根据上述高等教育评价价值冲突发生的逻辑,化解价值冲突需要三个条件:一是冲突双方的价值都得到承认,即解决在多元化价值取向、多样化质量标准背景下“评什么”的问题,或者说评价标准问题;二是构建起化解冲突的公共机构,这在很大程度上是解决在多元化、多样化背景下“谁来评”的问题;三是形成一些冲突双方都能共同遵守的解决冲突的正式规范,即有利于调节冲突的评价过程规范和技术、方法问题,以解决在多元化、多样化背景下“如何评”的问题。

1. 构建能够整合多元价值取向的统一质量标准:以“育人为本”作为高等教育评价的核心标准及价值融合的方向

价值取向多元化是当今社会无可置疑的一个特征。不同个人、不同群体、不同组织对于同一教育现实的价值取向由于主客观因素的影响而存在差异,价值冲突的本质在于价值取向的多元化现实,而对于如何解决价值冲突的问题,不同的立场可能决定着不同的方法。所以,基于多元文化主义的现实,解决冲突的合理公正的方法应基于价值融合的机制,而非价值分裂的机制。这是我们关于高等教育评价的一个基本理念。

价值融合作为一个基本理念,是高等教育评价的一个核心思想与出发点。融合的可能性

即是在主体多样化、价值多元化的背景下，要有一个共同对话的基础和尺度，或需要一个共同的价值标准。这个共同的价值标准能够统领主体多样、价值多元背景下的高等教育评价的价值标准，其本质在于使多元的价值取向在评价过程中都能够得到承认。

毋庸置疑，高等教育评价首先要体现大学的本体性功能，即体现大学的人才培养功能，体现大学的公共性。而公共性恰恰又是通过人才培养来体现和实现的，其他诸如科学研究与社会服务也主要是通过人才培养来实现和体现的。同时，高等教育利益相关者如国家、大学和社会等也均可通过人才培养功能的实现而达至自己所追求的价值目标。国际竞争和创新的压力可以通过培养高质量的各级各类具有创新意识与创新能力的人才而得到缓解；高等教育规模扩张后人才培养质量的暂时下降可以通过夯实本科教育、落实学生为本、教学为本而稳步提升；大学的透明度和公信力无非是大学的纳税人及其利益关者想知道公共财政都投到了人才培养的哪些方面，培养过程中的质量是如何掌控的，大学又如何证明其人才培养和质量满足了社会的需求，是否可以与国际同行相较一二。所以，人才培养功能的实现情况，即人才培养质量如何是高等教育评价的一个核心价值取向，或者说是多元价值标准中的元标准。这一点实际上早在 1998 年世界高等教育大会发布的《面向 21 世纪高等教育宣言：观念与行动》中得到明确：“21 世纪将是更加注重质量的世纪，由数量向质量的转移，标志着一个时代的结束和另一个时代的开始。重视质量是一个时代的命题，谁轻视质量将为此付出沉重的代价。”所以，高等教育评价中多元价值的融合要以人才培养质量为基础出发点，即“育人为本”应成为高等教育评价的核心价值标准和多元价值融合的方向。

实现价值融合的办法是价值整合。价值整合是价值融合的具体执行，是操作层面的东西，其关键是如何规避、减少评价主客体之间的价值和利益冲突。在价值整合乃至价值融合的过程中，除了高等教育利益相关者的不同价值取向在评价过程中都能够得到承认（即以“育人为本”作为评价的核心价值标准）外，还需要考虑“谁来评”和“如何评”的冲突化解策略。

2. 建立能够化解价值冲突的科学评价机制：制度环境和技术环境

从高等教育评价过程来看，它通常包括建立评价机构，明确评价问题，确定评价目的，确定评价对象、内容及方法，确定信息采集的性质、数量以及分析信息的方法，科学采集和分析信息，综合描述、判断、解释信息并得出结论、形成改进建议，用适当方式向评价委托人报告评价情况，将评价结果与建议运用于教育实践，对评价过程和结果的科学性和合理性进行再评价。从我国高等教育评价实践来看，一方面，评价机构基本上由主办和管理高等教育的政府部门设立，即管办评合一；另一方面，运用评价结果改进教育实践的还很少，对评价过程、结果及效用进行再评价就更少了。在高等教育主体多元、价值多样的大背景下，高等教育评价要做到尽量规避、减少评价主客体之间的价值和利益冲突，实现价值整合乃至价值融合的目的，就必须建立起教育评价与教育实践之间良性互动的规范与制度，建立起评价过程的再评价制度，大力营造协调和化解高等教育评价价值冲突的制度环境和技术环境。

所谓制度环境，主要是指用来建立和促进评价方法技术开发和运用的政治、社会和法律制度，包括软制度、硬制度和实施机制三个层次。其中，软制度是指人们在长期的社会生活中逐步形成的习惯习俗、伦理道德、价值观念、意识形态等对人们的行为产生非正式约束的规则，是社会认可的非正式制度；硬制度就是从宪法到成文法和不成文法，再到具体细则等国家规定的正式制度；实施机制是指制度安排的具体操作过程。营造适宜的制度环境，需要从三方面着手：在硬制度层面，《国家中长期教育改革和发展规划纲要（2010-2020）年》等政策已经明确规定了高等教育“管办评”分离的原则，在此基础上，应尽快建立健全相应的法律法规，尤其要把第三方评价纳入法制化的轨道，明确第三方评价主体的地位、性质、权

利及义务，同时完善政府绩效定期评价制度、政府信息公开制度、绩效预算制度、绩效审计制度等；在软制度层面，要加强绩效评价理念、绩效文化的推广，促进政府绩效观的转变，激发公众对政府绩效进行监督的意识，提高民间组织对第三方评价的重视程度等；在实施机制层面，要落实结果应用、信息公示、评价监管等机制，使第三方评价操作流程公开透明、可信度高。

所谓技术环境，是指一个国家和地区有关高等教育评价的方法技术水平、技术政策、技术开发能力及技术发展动向等。高等教育评价技术是在高等教育领域为解决评价问题、实现评价目的的方法或手段，与通常意义上所说的“如何评”这一问题相关的各种想法和做法，都包含在高等教育评价方法技术这一概念之中。具体来说，评价的方法技术主要涉及评价技术的功能、作用对象、工具选择、工具使用、技术使用意图、使用效用六个方面。这六个方面的问题相互联系，相互影响，最终决定了一种方法技术的效用。有效运用高等教育评价方法技术解决评价问题，必须深入理解、准确把握、合理处理上述各种问题。随着高等教育大众化和普及化的推进，高等教育评价的方法技术也得到多样化的发展与完善，总体上可分为实证主义评价方法和后实证主义评价方法两种。实证主义评价“在具体的方法与技术操作层面发展了精确的以定量为主的研究方法，试图在评价中排除人的主观因素的影响，从而保证评价的客观性”；而后实证主义评价则“在评价方法和技术操作层面主要以定性研究和质性研究探索参与评价过程的各主体的需要和利益”。多样化的评价方法技术为采用多样化的评价模式、协调多元主体的价值冲突奠定了技术基础。从我国高等教育评价的现实来看，这两类评价虽都有涉及，但总体上还处于实证主义评价阶段，如清华大学、北京师范大学、南京大学等引进的本科教育学情调查（NSSE）、大学生学习经验研究（CSEQ）以及研究型大学本科教育学情调查（SERU）等。当然，在严格遵循实证主义评价路径开发以精确的量化研究为主的本土化评价方法技术方面，我们也在不断地探索和推进，如厦门大学开发了“大学生学习情况调查”、华中科技大学开发了“中国大学教师调查”，以及自2015年开始举办全国教育实证论坛等，但严格遵循后实证主义精神开发出来的评价方法技术还很少。所以，在从高等教育大国向高等教育强国迈进的过程中，我们“既要在遵循实证主义的原则的基础上竭力丰富和完善现有的评价方法论体系，也要努力跟进后实证主义评价方法论体系的发展，进而在此基础上发展有中国特色的评价方法论体系”。

3. 形成能够化解价值冲突的正式规范：处理好三对关系及其转换

价值论域下的高等教育评价是基于价值事实的价值判断，关系属性、生成属性及多维属性既是价值问题的基本特征，也是分析价值问题的方法论。换句话说，从认识论域下的高等教育评价走向价值论域下的高等教育评价，需要转换思维，遵循关系的、生成的和系统的思维方法，重点处理好主客体之间的关系、评价的目的以及评价程序的科学性与可持续性三方面的问题。

一是客体思维与关系思维的关系。二者主要就评价主客体之间的关系而言。客体思维在评价活动中单纯以主体为尺度来度量价值，主要表现为更多甚或只考虑高等教育或其因素（如学科、专业、课程等）这一评价客体对政府/国家、社会（企事业单位）、个人等评价主体的价值需求满足程度，很少甚至不考虑评价主体对评价客体的价值需求满足程度，是一种单向的价值判断。关系属性主要表现在评价互动过程中，价值的实现也是评价主客体之间相互运动、相互作用所形成的效应，既包括客体对主体的效应，也包括主体对客体的效应。也就是说，评价过程中既要考虑评价客体——高等教育或其因素对评价主体的价值（客体思维强调的价值，以主体为尺度），也要考虑评价主体对客体的价值（以客体为尺度来度量的价值），即作为评价主体的政府、社会、个人等在评价过程中对高等教育价值需求的满足程度，

如是否保证了正常的教学秩序，帮助学校发现了问题，形成了解决问题的路径方法等。这需要评价主客体在评价过程中相互对话、沟通乃至理解，以真正达至“评价是为了改进”这一根本目的。从某种程度上讲，评价过程中单纯的客体思维是导致高等教育评价饱受社会诟病的认识论根源。

二是训诫导向与反思导向的关系。二者主要就高等教育评价的目的而言。训诫导向的评价主要以鉴定高等学校的办学条件、办学水平、教育质量合格与否以及发展程度高低和办学状态优劣的结果来实施相应的奖惩（如经费的划拨、招生人数的增减、学位专业点的增减等）为目的，强调结果与资源奖惩。从高等教育评价实践来看，无论是合格评价、水平评价还是评优评价，大多是鉴定性的，许多学校将其看作是对自身办学水平的鉴定，教育行政部门也主要利用评价结果对学校实施奖惩。反思导向的评价主要通过评价过程中主客体之间的双向互动，发现高等教育及其因素在发展过程中的进步与优势，诊断和反思高等教育或学校工作中存在的问题、困难与不足，并据此提出相应的意见和建议。高等教育质量评价的目的不是为了证明，也不是为了监督和训诫，而是为了发现问题，共同改进，从而达至增值的目的，强调过程与新价值生成。与传统训诫导向的鉴定性、终结性评价不同，反思导向的教育评价越来越强调把改进学校的教育质量作为主要目的，越来越强调评价的形成性功能，以帮助学校持续不断地改进教学专业活动的效果，不断满足高等教育利益相关者对大学人才培养的期望。应该说，新一轮的审核式评价在反思导向上迈出了第一步。

三是单环学习与双环学习的关系。二者主要就评价程序科学化及可持续发展而言。根据学习型组织理论，单环学习是指组织或个人通过具体经历的观察、回顾和思考形成抽象概念，再在新环境中检验概念，发现和纠正其中的不足与错误，以寻求行为和结果之间的匹配，以保证组织的正常运转。在此过程中，强调的是对现状的认知，并不质疑和改变组织的基本意愿、信念和价值观。组织可利用单环学习来获取经验、察觉和解决问题、判断和预测未来，但组织的政策和目标并未发生变化。单环学习培养的是组织扫描环境以及控制目标实现的能力，是一种认知性学习，遵循单一、单向的解决问题的逻辑。双环学习通过对行动背后的想法加以检视，追问组织行为的前提是否恰当，反思看问题的心智模式，谋求从行为的前提变量（即行为的前提假设）上取得根本性改善。当发现错误时，强调对造成现状的原因进行反思，其改正方法常常包括对组织目标、政策和常规程序的修改。双环学习向组织中根深蒂固的观念、规范或传统惯习提出质疑和挑战，有利于人们提出与以往不同的问题解决办法，进而采取真正有效的行动。双环学习培养的是开放、变通的学习精神和鼓励怀疑、欢迎挑战、勇于放弃的态度，是一种创新型学习，遵循不同层次、双向的问题解决逻辑。单环学习在高等教育评价过程中主要表现为评价主体按照评价标准对评价客体——高等教育或大学的现状进行“认知性”的价值判断，而很少反思现状背后的历史原因、发展环境以及本身办学目标与定位等。双环学习则强调在评价过程中不仅要参照评价标准对评价客体的现状进行“认知性”的价值判断，而且更强调对造成评价客体现状的原因如办学目标、政策以及教育模式等进行反思，同时对评价目的、评价指标体系以及评价程序、方法等进行反思。换句话说，在评价过程中，评价客体或评价对象（如大学和教职员工等）主动开展自我评价，不仅根据评价标准与要求就具体的评价对象和内容，对办学条件、运行情况和教育质量进行全面深刻的反思，总结成绩经验，查找问题不足，分析原因后果，商讨对策措施，而且还自觉地再次思考自己的理想追求、教育愿景及价值定位等。同样，对于评价主体（如政府、社会、个人等）来说，一方面，他们要能够发现并准确把握评价过程中评价客体所存在的问题，反思高等教育作为一个复杂性的存在物，其办学条件、发展状况和教育质量等方面存在的问题。虽然这与评价客体——高等教育自身有着毋庸置疑的关系，但评价主体对此也有不可推卸的责任和

义务,如作为评价主体的教育行政部门是否保障了经费投入,是否给予了大学充分的办学自主权等,作为评价主体的社会如用人单位和家庭及学生个人是否一味地强调大学必须根据市场的需求来培养人才等。因此,评价主体在评价过程中也常常需要理性反思自己的工作与高等教育发展状况之间的关系,根据高等教育评价所反映的成绩与问题,探讨高等教育改革与发展的新思路、新战略。另一方面,评价主体还需要对评价进行再评价,持续考察高等教育评价方案及程序是否科学、评价指标体系设计是否合理、评价过程是否客观、评价结果是否确切等。再评价一般有两种方式:一是总结性再评价,即在评价结束后对本次评价的方案设计、技术方法、实施程序、结论质量及其产生的作用和影响作出全面的分析、评价,为最终的决策提供咨询建议或施加影响,也为改进评价活动提出意见和建议;二是形成性评价,即在评价的不同阶段开展再评价,以便及时修正评价的方案、技术或操作中可能出现的问题。当然,评价过程中双环学习的形成有赖于科学评价制度环境与技术环境的建立,也有赖于主客体在互动过程中平等、坦诚的交流、对话以及互为价值客体的关系思维的实践。

4. 遵循能够解决价值冲突的规范生成路径:交往伦理与制度生成

高等教育评价作为一种主客体之间基于价值的认识活动,其特殊性就在于评价主客体在相当程度上均是高等教育的利益相关者,同时主客体之间在互动过程中是互为价值客体而存在的。换句话说,评价主客体均对高等教育的发展及其质量负有相应的责任与义务。这就为评价主客体之间价值冲突的协调与化解奠定了共同的责任和义务基础,也为评价制度的科学化与可持续发展提供了共同的基础。与我国传统的以血缘和情感为基础的“人伦本位”交往伦理不同,评价过程中主客体之间的交往关系是建立在现代“个体本位”交往伦理基础上的。如果说传统“人伦本位”的交往伦理在一定程度上也表现出了“交互主体性”,那么现代“个体本位”的交往伦理虽然也强调交往双方互为主体性,但却是以肯定主体存在的独立价值为前提的,主张的是“我”与“你”之间的自由平等的对话与交流;人与人依赖理性意志对遇到的各种问题进行商谈、讨论,找出解决方案,并制定人类交往的规则。国内外的评价实践表明,恰恰是在评价过程中,在充分肯定评价主体和评价对象都存在独立价值的基础上,基于互为价值客体的主客体之间的交往、沟通、对话与协商,才能不断促进高等教育评价制度乃至高等教育内外质量保障体制机制的不断生成与完善。对于评价客体——大学自身整体发展或构成要素的发展来说,作为办学主体的大学和作为教育主体的教职员工的积极参与是高等教育评价活动顺利开展和达成目的的前提。更为重要的是,他们在评价过程中所开展的自我评价以及针对评价结果对学校办学情况和教育质量进行的深刻反思,不断促进大学内部治理和质量保障体系走向科学化。对于评价主体——高等教育主管部门或评价组织机构在评价过程中或评价结束后及时开展的再评价,可以渐进弥补评价中的不足,有些带有导向的问题可以得到及时纠正,从而促使高等教育评价制度及高等教育外部质量保障体系不断走向科学化与规范化。

可见,高等教育评价制度是基于价值协商基础上的相关利益者在人才培养、科学研究与社会服务活动方面基于价值事实而开展价值判断时所应遵循的惯例、习俗、规范、信仰和仪式,是高等教育评价有序组织与运行的规则系统。无论是关系思维理念在高等教育评价过程中的形成与实践,还是反思导向和双环学习的评价规范化程序的形成与实践,都有赖于建立在现代交往伦理基础上的科学的高等教育评价制度和环境的建立。

【李庆丰,北京工业大学高等教育研究院研究员,教育学博士;周作宇,北京师范大学副校长,教育学部教授,博士生导师。】

——来源:《高等教育研究》,2020年第10期

创新评估范式 引导大学评价

王战军

大学评价是大学和学科建设的风向标、指挥棒，是检验大学和学科建设成效的重要手段。6月30日，中央全面深化改革委员会通过的《深化新时代教育评价改革总体方案》中明确指出，落实立德树人根本任务，改进结果评价，强化过程评价，探索增值评价，健全综合评价，着力破除“五唯”，建立符合时代要求的教育评价机制。在近期召开的全国研究生教育会议上，教育部陈宝生部长也指出，新时代研究生教育发展要过“评估关”，扭转不科学的教育评价导向。

创新评估范式要不忘初心、牢记使命。当前，各类大学和学科排行榜对我国高等教育的学科体系、教育教学体系形成了冲击，“五唯”的形式主义现象普遍存在，简单思维、排名思维、功利思维的顽瘴痼疾使一部分人追逐排名，忘掉了大学和学科建设的初心和使命。大学和学科建设作为一个国家的高等教育水平标志，首要任务是面向世界科技前沿、面向国民经济主战场、面向国家重大需求、面向人民生命健康，以原始创新为核心，占领科学发展制高点，解决科学技术关键领域“卡脖子”问题，产出更多的具有国际影响力的重大原创性成果，着力提高关键领域原始创新、关键技术自主创新和国际学科影响力。

创新评估范式要把握大学和学科建设评价导向。大学和学科建设评估要坚持立德树人导向。大学和学科的核心任务是培养一大批德才兼备的高层次人才，教育评价必须遵循人的教育规律、人的成长规律、人的发展规律。大学和学科建设评价要把人才培养质量放在首位，培养的人才主要为了满足国家战略、经济社会发展和各行各业发展对多层次人才的需求。

大学和学科建设评估要把人才培养质量和人才培养成效作为根本标准。

创新评估范式要迎接大学和学科建设评价的新挑战。我国大学和学科建设评价起步于上世纪八十年代，发展于新世纪，已经成为大学和学科建设发展评价的重要工具。当前，眼花缭乱大学排名、学科排名使我们的大学和学科建设迷失了方向。全面贯彻落实全国教育大会和全国研究生教育会议精神，促进大学生、研究生德智体美劳全面发展，各大学要深刻认识现代大学和学科建设的时代背景和战略意义，必须要清理“五唯”思想，充分认识评估事物的复杂性，摒弃排名思维，由简单思维转向复杂思维、由形式思维转向本真思维。深化教育评价改革，必须构建具有中国特色的、符合中国国情的教育评价体系，必须探索新时代符合中国高等教育发展规律、积极引导教育服务国家战略发展大局的教育评价改革之路。

创新评估范式要树立监测评估理念。在知识经济和信息化时代，大学和学科评价亟须创新评价理念、理论与方法。大学和学科建设是一项长期性、系统性的工程。监测评估要全过程、全方位、多层面、多维度呈现建设状态、建设成效与发展水平，为多元主体提供科学的、客观的、有效的决策服务。大学和学科建设评估要树立监测评估理念，采用监测评估方法，持续提高办学效率和教学质量。

创新评估范式要应用监测评估方法。监测评估是利用现代信息技术持续收集、深入分析有关数据，直观呈现学科建设状态，为多元主体价值判断和科学决策提供客观依据的过程。监测评估方法是基于大数据驱动的现代评估方法，主要包括动态数据获取技术、大数据挖掘与分析技术、数据呈现技术等等。监测评估能够实现对大学和学科建设发展的过程评价、增值评价、结果评价和综合评价。监测评估还可以对大学和学科建设状态进行预警、预测，对教学、科研、社会服务等各方面运行情况进行常态监测和直观呈现。

创新评价范式要构建多元多维的评估体系。大学和学科建设评价要突破简单的排名思维，坚持定性评价与定量评价相结合，针对不同类型的大学、不同类型的学科，针对同一类型的大学和学科的不同发展阶段，针对不同的建设目标和要求，系统构建多元多维的评估指标体系和评价标准，开展分层评价、分类评价、多元评价和精准评价，对其发展建设情况进行整体发展水平评估、建设成效评估。大学和学科建设评估的核心是从达成度、贡献度、支撑度、影响度、引领度等五个维度回应国家和社会发展的时代需求，构建一套“中国特色，世界一流”的大学和学科建设评估体系。

2020 年是我国“十三五”规划收官之年，也是“十四五规划”谋划和“双一流”首轮收官之年。在新时代，面对百年未有之大变局，各个大学要从国家战略、社会需求、基础条件、办学定位出发，抓住机遇谋发展，为实现两个一百年目标、建设高等教育强国做出自己的贡献。

【王战军，北京理工大学研究生教育研究中心主任，人文与社会科学学院特聘教授。】

——来源：《教育发展研究》，2020 年第 19 期

声明：

《高教视点》为江苏理工学院高等教育专题研究参考资料，不定期围绕相关专题开展探索研究，内容主要来自于国内学术期刊、政府网站、报刊等文章。由高教研究室收集整理，仅供本校高教管理工作者和高教研究者交流参考，切勿另做其他用途，特此声明！

编辑：江苏理工学院高教研究室

地址：江苏理工学院厚德楼 212 室

电话：0519-86953695

电子邮箱：jsut-gjyjs@jsut.edu.cn