

中国式学科评估：问题与出路

今年四月份，教育部学位与研究生教育发展中心（以下简称“教育部学位中心”）邀请全国学位授予单位参加全国第四轮一级学科整体水平评估。随之，各个高校展开了一场大规模、高级别的学科评估申报及材料提交总动员。第四轮学科评估自发布起也引发了学界的广泛关注和热烈讨论。高等教育是中国崛起的思想发动机，关涉民族复兴的未来，而学科评估是近年来中国高等教育学科建设成就的集中展示，其意义和影响可谓深远。为了更好地推进学科评估科学进行，特别是促进高等教育健康发展，《探索与争鸣》编辑部邀请全国知名学者围绕第四轮学科评估实施过程中的问题及出路展开对话，特别就评估的目的、主体、结果使用及如何改进等做了深入探讨。

与会专家认为，学科评估作为政府履行有关职能的手段，通过评估实现以评促改、以评促建，达到质量保障和提升的目的，本身无可厚非。但是，具体设计和操作中却出现了一些问题，概括起来，主要有：第一，评估的初衷与实际效果存在偏差。本次评估宣称淡化学科排名，着重于诊断与监测的功能，但此轮评估带有较浓的官方色彩，又处于国家“双一流”建设实施方案和配套政策尚未明朗时期，导致高校把附加的利益看得太重，因过度重视而产生“恐慌”应对现象。第二，本轮学科评估所秉持的价值立场不够明确。有些表述或过于宏大笼统，或语焉不详，这让大学无所适从，欲罢不能，理性上不愿为评估所左右，但在现实中又不得不伴着评估起舞。第三，评估实施主体的独立性和公信力有待商榷。专家们认为，第三方评估的首要价值就在于，第三方评估机构与委托方和受评方利益无涉，作为教育部下属单位的学位中心尚不具备这一独立性。第四，活动组织中行政化思维依然占主导，此轮学科评估依然沿袭“短、平、快”的政治动员方式。第五，加剧了高校内部的无序竞争和生态失衡。一些高校被学科评估指标牵着鼻子走，借着“资源整合”的名义做起了数字游戏，以“砍杀”或拆并非重点、非特色、非优势学科为代价，来确保重点学科的权威地位。第六，学科评估标准“窄化”和“固化”。所谓“窄化”，即用某一门学科或者某一类学科的评估理念和技术方法，去评估所有的学科；所谓“固化”，即拘泥于某一学科的要素与特点来衡量各类学科，缺乏灵活和变通。第七，评估结果绝对化。学科评估只能用一些显性的、可量化的指标来作为评价依据，依靠显性指标和定量结果来量化学科间的绝对差距，这是不科学和不合理的。第八，评估功效绝对化。评估行为可以实现督促监管，却不能对质量提升有太多直接的作用。评估不是制造优秀大学、优秀学科、优秀学者的万能工具。第九，现有学术刊物的认定方法及相关评估技术需要改进。

如何真正做好学科评估？专家们建议，学科评估应尊重丛林生态，培育自主生态，坚持二八分治，理性把握边界，适度合理调控。主要体现在十个方面：第一，学科建设是大学的一项日常工作，是一个漫长的不断累积的过程，切不可急功近利。学科评估应能够促进学科整体梯队建设，而不是搞掐尖投资。第二，推进管、办、评分离，多管齐下培育第三方学科评估机构，政府部门应制定政策扶植和培育第三方评估机构的建立、健全和发展，同时对其实施必要的监管，使其依法依规开展评估。第三，通过融合评估的事实判断与价值判断，建构起科学的学科评估体系，这是走出具有中国特色学科评估之路的基本路径。第四，外部力量尤其是政府力量的介入应理智地保持“一臂之距”，政府监控的重心应是对学科建设和人才培养质量的底线保障上。第五，明确学科评估的根本作用在于帮助参评学科在相互比较中知己知彼，诊断问题和寻找对策，学科评估应回归到学科建设本身，培育卓越的学科群体。第六，摒弃以论文论英雄的价值取向，衡量学科和学者水平，不在于论文、著作、课题数量，而在于对本学科的贡献，对国家和民族科技进步的贡献，这是学科建设追求的根本目标，也是学科评估的核心立场。第七，文史、理工、农医、经管等学科遵循不同的规律，学科评估也要分门别类制定不同的标准，切忌一刀切。学科评估应该营造“万类霜天竞自由”的氛围，而不能用丛林法则管理丛林。第八，充分尊重高校的办学自主权，建立多方协商对话机制。通过协商对话机制来保证评估参与各方的诉求和利益得到最大限度的表达和响应。第九，建立公共数据库，学科评估尽量不要“进校”，不应影响高校正常教学秩序。第十，谨慎使用评估结果，尽量避免评估结果与资源、利益的直接挂钩。政府如要将资源分配与学科评估结果建立一定联系时，要综合考虑教育投入的公平、公正、透明、和谐和高效等原则，不仅要奖励排在最前端的大学或学科，也要托底排在尾端的学科和激励大多数大学。

——陈学飞 叶祝弟

主持人：

陈学飞，北京大学教授、博士生导师

叶祝弟，《探索与争鸣》副主编、副编审

对话人：

王英杰，北京师范大学教授、博士生导师
 龚放，南京大学教授、博士生导师
 秦惠民，中国人民大学教授、博士生导师
 施晓光，北京大学教授、博士生导师
 董云川，云南大学教授、博士生导师
 卢晓中，华南师范大学教授、博士生导师
 樊秀娣，同济大学副研究员

熊庆年，复旦大学教授、博士生导师
 胡建华，南京师范大学教授、博士生导师
 陈洪捷，北京大学教授、博士生导师
 刘海峰，厦门大学教授、博士生导师
 马陆亭，教育部教育发展研究中心教授、博士生导师
 周川，苏州大学教授、博士生导师

叶祝弟:《探索与争鸣》杂志一直致力于为不同观点的交流与对话搭建一个公共平台。第四轮学科评估,引发了学界的广泛关注和热烈讨论,甚至是激烈的争论。对于此轮评估中出现的一些问题,我们需要做审慎的观察和思考。为了中国高等教育的长远发展大计,经陈学飞教授提议,我们邀请诸位先生进行一场“中国式学科评估:问题与出路”的对话和交流,请各位先生本着正本清源、实事求是的原则,畅所欲言,并提出真知灼见。我们将整理出诸位先生的思考和对话实录,呈现给读者。我们请陈学飞教授主持这次讨论。

陈学飞:今天的沙龙请大家围绕两个主题展开讨论。第一,教育部学位与研究生教育发展中心组织的第四轮学科评估,客观上对高校产生了哪些影响?第二,大家对学科评估有一些争议,那么什么样的评估是比较科学的,能够真正评出学科发展的真实水平,并真正能够有效促进高校的学科发展?请大家先就第一个问题发表看法。

问题一:评估初衷和实际效果存在较大偏差

熊庆年:我先来讲几句我的观感,对于第四轮学科评估,我有些话如鲠在喉,不吐不快。目前正在进行的第四轮学科评估,感觉其评估初衷和实际效果之间存在较大的偏差。现在进行的学科评估,把几乎所有的高校、所有的院系都调动起来了。虽然名义是自愿参评,实际效果是,不管你愿意还是不愿意,都得参与其中,与“自愿参评”的原则相去甚远。之所以这样,因为大家心里明白,未来政府的投入将与评估结果挂钩。为了争取排名靠前,不少学校采取了“丢车保帅”的策略,围绕某些优势学科或者方向,归并整合成果、整合材料,结果造成一些严重的问题。比如,在一个学校内部,哪个上,哪个下,哪个合并到其他学科,很大程度上是权衡和算计的结果,一时间学科内部、不同学科之间激烈竞争,导致一些学科分裂,甚至学校内部院系分裂。如果围绕一个功利的目的去进行评估,破坏学科建设的稳定性,那么这对高等教育的影响将是难以挽回的。联系到最近正在进行的一流学科建设,究竟什么是一流学科,大家现在越来越糊涂了,难道把学科“捆绑”起来就是一流学科吗?

龚放:对于熊老师的这个困惑,首先,我们要弄清楚学科是什么。学科是人们认识客观事物的一个视觉通道。我们要认识客观事物,需要借助一些观察视角,形成一些概念、范畴,这就形成了一定的学科。但是,客观事物无论是自然界,还是人类社会,不是一个单一学科所能穷尽、所能覆盖的。所以,学科的发展既是人们认知客观世界的一个进步,也是大学现代化的一个标识。从传统的

经史子集到数理化、天地生、文史哲经法,20世纪20年代大体上形成一个现代学科体系,这是一大进步,但仍然存在局限。每一个学科实际上只是我们认识外界事物的视觉通道,美国学者伯顿·克拉克将其形容为“打开一盏盏聚光灯”。打开单一的聚光灯,可以把局部照得纤毫毕现,但也可能造成失真,可能隐去背景,可能造成其他困扰。好比照相,如果多盏灯打开,才能柔和,如果只打开一盏灯,比如只有脚灯打开,就有可能显得面目可憎。所以,观察世界最好打开多盏聚光灯,借助多个学科。

中国现代大学的很多学科是上世纪20年代前后形成的,那时候,近代数学、近代物理、近代化学、近代生物学、近代地理学等先后产生,法学、社会学、哲学、经济学等也逐步取代了我国传统的经、史、子、集的知识系统,奠定了现代大学的学科专业格局。但是百年后的今天,大学的学科结构不断演化、学科前沿持续推进,发生了很大变化。更何况,现实世界是很复杂的存在,今天的人类面临的都是综合性难题,这些都不是哪一个学科就能解决的,需要的是不同学科联合攻关,携手前进。所以,有见识的学者提出打破学科壁垒,提倡跨界研究。这样的呼吁应该是符合人类社会发展的规律和要求的。现在,跨学科研究、多学科合作、跨界行动,预示着问题导向、需求牵引、应用为上的知识生产方式方兴未艾,与学科导向、学术为上的知识生产模式相辅相成又相反相成。在这样的发展趋势、时代背景下,如果我们对学科建设的理解过于狭窄,过于固化,甚至用某类学科的思维定势来看待学科建设、指导学科评价(包括所谓的“一级学科评估”,以及人为划定的A刊标准,等等),那就会误导整个高等教育发展,造成难以挽回的损失。

现在不少大学并不是围绕科学技术发展的趋势、围绕国家战略发展及社会发展的需求来凝练学科发展方向,组织多学科、跨学科的合作以加快知识创新、理论创新、技术创新和集成创新的步伐,而是被大学排行榜、学科评估指标牵着鼻子走,一味琢磨如何对上排行榜的口径,如何根据有关学科评估指标体系来整合学科,归并成果,提高自己的排名。有些学科,比如物理学,有的大学分成应用物理、理论物理两个院系,为了能使学科排名得以前移,有的大学想方设法将它们成果归并起来。还有的更是着力做文字游戏、数据搬家、变相造假……这样做,对大学自身发展、对国家发展都没有帮助,真正是“可怜无补费精神”,反污了大学清誉!

陈学飞:“可怜无补费精神”,龚老师说得好!学科评估类似于高考指挥棒,搞得高校也要按照这个指挥棒来发展自己。以北京大学教育学院为例,我们在教育经济与管理方向上投入人力物力最大,产出的成果也比较多,但它属于公共管理学科,每次评估都要把它从教育学院、教育学科剥离出去。教育学院本来规模就很小,学科评估又把它

肢解了，造成北大教育学科每次排名都比较靠后。学校提出，北大只支持在全国学科排名靠前的学科，教育学科的发展就得不到必要的支持，结果势必影响学科未来的发展。所以教育学院下定决心，今年无论如何要把教育经济与管理争回到教育学科来评估。现在搞评估，就像做游戏。本来应该是客观评价，结果却像是人人在捉迷藏、猜字谜。许多经验教训都告诫我们，导向比努力更重要。我们的学科评估如果把导向搞错了，搞偏了，后果都将是长远的。

问题二：行政力量没有保持“一臂之距”

董云川：龚老师和陈老师的发言引发我颇多感触。本人多年来一直主持或不同程度地参与各级各类教育评估活动，对于中国式评估感受深切。

在加强宏观管理、促进微观放活的总体形势之下，政府职能的转变是未来教育行政改革的大势所趋，原来长期习惯于由“上级”下达以及安排的事无巨细的教育活动，不得不逐步分解、下放，而在分权放权的同时，所谓的“分级分类”职责被提上议事日程，与此相对应，各种各样的“第三方”评估机构应运而生，粉墨登场，在质量保障的口号之下各显神通。无论招牌如何，大家实际上还是不得不打政府的招牌，只是名目不同而已。如果真正脱离了与政府职能授权的干系，第三方教育评估机构基本上寸步难行。这是当下教育评估所处的真实语境。而原来对教育事业拥有绝对指挥权的各级各类职能部门，或多或少流露出失意失控的慌张，也在通过不同的方式继续强化着自己的操控权，试图守住原来的一亩三分地。

当下，政府、社会、评估机构以及作为办学主体的高等学校，大家对于提高教育质量的善意诉求毋庸置疑，然而，各方在权力制衡游戏中的角逐与博弈，依然紧锣密鼓进行中。

与此同时，各级各类学校在外力的强大作用力影响下，为了固守并抢占所谓优势特色学科在全国范围内的排名和地位，纷纷借“资源整合”的名义做起了数字游戏，以绞杀或拆并非重点、非特色、非优势学科的代价，来确保稳居马太效应中心区的学科的权威地位。历史迟早会证明这种做法之荒谬，甚至是愚蠢的。一所大学中不同学科的生长态势以及在历史沿革中的变动、组合、嫁接、交融因素千差万别，各有其历史合理性，再加入时空因素之后，更是变幻无穷，作为文化主体的大学本身在生存与发展中原本具备强大的“自组织”基因，亦会根据社会的发展变化呈现出“自组织”调节机制。纵横交错的学科构成了庞杂的学术丛林，外部力量尤其是政府力量的介入必须理智地保持“一臂之距”，否则就会因为越界而成为大学生长发育的负面动因。一言以蔽之，热带雨林里除了参天大树，还有更多的低矮草木，如若为了确保“望天树”的

辉煌地位，而铲除其他植物的话，后果可想而知！

胡建华：这次第四轮评估背后仍然是教育行政化的思维在主导。这是关系到我国学科发展的一个根本性问题，是不是只有评估才能促进学科发展？不搞评估，学科就没有办法发展吗？本来学校发展是为了培养人才，现在把排名作为学校发展目标。我觉得这是本末倒置，是一个异化。然而，行政化不去掉，这一评估的负面影响只会越来越严重。

问题三：以短平快的行政手段推出一流的做法很难持续

王英杰：俗话说，一叶遮目，不见泰山。我们可以将学科评估放在一个更大的视野下审视。今天的我们所处的是一个评估社会，在世界范围内形成了一股评估热潮，政府关注评估，专业协会/学会在评估，媒体在评估；评估的形式也五花八门，线上评估，专家组现场评估；评估的目标也各有不同，有的是为了设计最低限度入门门槛，有的是为了服务消费者而进行排行，有的是为了促进一流大学建设而进行评估。从某种意义上说，评估是乘新制度主义之风横扫世界。

何以如此？许多国家的政府发现，直接控制的传统治理手段难以为继了，因此需要从大政府转向小政府，从直接控制转向引领导向，简单说就是要简政放权。利用第三方来评估，对于政府来说最直接的好处是，可以使政府摆脱直接干预和控制之嫌，还可以使政府减少行政开支。在这些国家有比较成熟的第三方组织，比如在美国有些地区，学校认可机构和专业学会的认可机构已经存在运转近百年了，它们有效地通过评估来保证最低限度的质量。而《美国新闻与世界报道》杂志在全美甚至世界有广泛的影响，积极开展了大学排行，为学生择校提供了必要的信息。另外新媒体和新技术使得评估更简单易行，成本更低，影响面更大，于是评估开始大行其道，对政府决策、大众参与和选择都产生重大影响。现在的问题是，在新制度主义影响之下，大学较少受到政府的直接干预了，但是却日益为评估所困。从一定意义上讲，声誉是大学之本，声誉既是大学发展的动力，又是大学的沉重包袱。因此，面对评估，大学会欲罢不能，在理性上不想为评估所左右，在现实中则又不得不伴着评估起舞。

回到刚才大家所讨论的第三方评估的资格问题，我国目前的问题是，还不存在真正的第三方评估。我国还未建立起具有公信力的第三方评估机构，就开始大力开展相关评估了，结果就是政府主导评估。上一轮政府改革的时候，只是把教育部两个评估机构的帽子摘去了，而并没有发生实质性的变化，实际上还是以政府部门的名义运作，还是习惯于发红头文件。即便这样，也并没有坚持很长时间，重新又戴上了教育部的帽子。作为教育部直属的行政性事业单位，其负责人是国家公务员，

其任务是根据教育部制定的方针、政策和评估指标体系实施评估。教育部下属的所谓事业单位往往自视为政府部门，主要负责官员来自教育部，他们的做事方式也照搬政府部门的一套做法。把那些本来可以是学者的自律组织，或者学科的专业组织，变成官本位的组织，结果使得这些组织很难成为构建学科文化，研究学科发展方向，制定学科评价标准的真正的学术组织和影响学科发展的第三方机构。

上述所讲的是学科评估的主体问题。毫无疑问，政府作为办学方，要对大学进行监管，特别是在不断增加投入之后，必须了解投入的效益，对社会负责并作出说明。但是在没有具有公信力的第三方评估机构的情况下，教育部没有制定必要的政策，构建必要的政策环境，促进第三方评估机构的产生，而是反其道而行之，自己建立起评估机构，承担了办学者和评价者的双重角色。这样就使得评估具有了强制性，就使得大学不得不围绕评估起舞。公众质疑评估的客观性，也就难免了。

第二个问题，评估实际已与资源的配置密切地关联起来。这次学科评估，能搅动全国大学置日常工作于不顾，奋力一搏，成为学校工作中的重中之重，就是因为时机上与政府的双一流建设政策的出台太“巧合”了，难免不招人遐想。各个高校，特别是原985或211院校，对各自的优势学科立下军令状，誓夺第一，因为如果得不了第一，就意味着失去了国家双一流建设大学和学科的地位，就会失去巨大的资源，不仅是经费上的资源，而且事关荣誉，事关大学今后多年的发展。把资源与评估这么直接地、线性地、密切地挂起钩来，这个评估便不可避免地造成了全国总动员的形势，造成了以评估为中心来建设世界一流大学的局面。但这种方法是非常短视的，它迎合了社会的浮躁心理，认为世界一流大学可以通过这种评估，通过把资源非常集中地投入到某些大学的某些学科上，就能够在很短的时间内实现建设一流大学或者一流学科的目标。事实上，学科建设是大学一项日常性的工作，是大学不断累积的过程，学科要通过几十年不断的建设和发展才能够稳定地居于前列。

如果学科评估的结果是，大量资源集中在极少数大学，那么这样的政策效果是很难持久的。因为少了全国的基础性队伍，学科发展就很难持续，就像中国足球一样，若总是把巨额资源放在顶端，结果还是进不了世界一流队伍，最后发现还得从娃娃抓起。就像跑马拉松一样，要有第一群体或第一方阵，比如说电视报道肯尼亚有15位运动员都在第一方阵的范围内，这就保证了一流。倘若某大学的一个学科一流，其他学科却垮下去了，这一学科又怎能持久？如果未来国内某一学科主要的学术带头人、主要的研究成果和研究方法都出自一所大学，那么这个学科还有什么希望呢？独木难支，一木不成林，

道理浅显，政府部门领导要接受。所以从根本上来讲，这样的评估是没有理解大学是怎样的机构，大学的性质是什么，大学的任务是什么，如何建设大学，如何发展大学。正是这种行政化的思维方式，以短平快的办法，用我们最熟悉的运动方式来推出、捧出个别的一流，结果可能在一定时期有一定的作用，但是不可能长久，所以这个办法是行不通的。

问题四：学科评估实施主体的独立性和公信力有待商榷

熊庆年：我接着王老师所说的第一个问题，确实如此，第四轮学科评估之所以受到广泛质疑，一个重要的因素就是，大家认为承担这次评估事务的教育部学位中心缺乏独立性，换句话说，人们并不认为它是第三方机构。大家的质疑是有一定道理的。

第一，教育部学位中心是教育部直属的事业机构，所秉承的是教育行政部门的意志，听从的是相关机构领导的指令，是事实上的利益相关者，并不具备独立性。

第二，学位中心某些做法令人不解。比如，学位中心解释，A刊的遴选是“学位中心请汤森路透、爱思维尔、CNKI、CSSCI、CSCD根据文献计量学等方法遴选产生初步名单，并请全国博士生导师进行网络投票，再由国务院学位委员会学科评议组推荐确定”的。然而，我作为博士生导师，并没有接到过参加网络投票的邀请，我也询问过我认识的一些博士生导师，也没有人说参加过投票。

第三，评估运作缺乏基本规范和专业严谨。数据填报提交的截止日期两次改变，评估活动已经启动而分学科的方案还没有出来，具体的数据标准更是不断更改，让参评单位无所适从，完全不像是一个专业评估机构应有的行为。

没有真正的第三方评估机构，就不会有科学合理的评估。第三方评估的首要价值就在于，它与委托方和受评方的利益无涉，能够保证客观、公正。去年5月，教育部出台了《关于深入推进教育管办评分离、促进政府职能转变的若干意见》，提出“在做好内部评估的同时，要主动委托第三方开展全面、深入、客观的评估”，这从政策上明确了第三方评估的地位。实践上，我们不得不遗憾地看到，真正实现管、办、评分离，实现真正的第三方评估，还有很长的路要走。可以这么说，在现阶段国内还没有真正的第三方教育评估机构。无论是教育部所属的学位与研究生教育发展中心等一批中心，还是各省、市的教育评估院（所），都是准官方的教育评估机构。它们是事业编制，有行政级别，吃的是财政饭，不可能与上级部门利益无涉，不可能持独立的立场，也就不可能做到真正意义上的客观公正。第四轮学科评估中出现的乱象也印证了这一点。

问题五：利益捆绑客观加剧高校内部无序竞争和生态失衡

陈洪捷：我们可以看到，评估的指标是由专家所制定的，并且一直在不断地修改，力争能够准确地评价学科的状况，但是现在所面临的问题，主要不在评估技术本身（当然这方面也需要不断地研究和改进），而在于评估背后的理念和逻辑出了问题。实际上，学科评估并不能真正评估出学科的质量，为了便于操作，只是用一些显性的、量化的指标来作为评价依据。其结果就是吸引大家都来“做”指标，拿出漂亮的指标来换取好的评估结果。目前大学里比较普遍的急功近利的种种做法，在很大程度上其实都是“指标”的需要，未必是出于真正学术的需要。

从大学基层来看，很多老师（或者说大多数老师）不关心学科的评估问题。那么谁最关心学科评估？当然是管理者们。教师是大学的主体，但大学的主体不关心学科评估，这说明什么？很值得探究。管理者关心评估，这说明评估本身就是他们的分内工作，或者说，评估就是一种管理工作。而我们都知，管理就是服务。但当一种服务，不被服务对象所认可，还让服务对象有排斥感，应当说，这种服务是不合格的。

事实上，管理者其实并不把自己提供的这种“服务”视为服务，而是视为一种以权力和资源为后台的“管制”，当然往往可能是出于善意，认为这种管理可以促进学术的发展，提高学科的水平。管理者重要的管理手段就是通过可衡量、可计算的指标来评估，并在此基础上分配手中的资源。试想，一旦废除种种指标，管理者如何来管理？管理者业绩如何体现？所以管理者视评估为其管理的重要手段，所谓“以评促建”，就道出了这种逻辑。

秦惠民：迄今为止的学科评估，实际上是中国式学科建设的一部分，是以行政化的方式评判、影响甚至框定学科发展并以此进行资源分配的一种重要手段。由学科评估引导的学科建设，实现了行政管理和学术管理的制度性同构。无论是从教育行政部门还是从大学来说，以学科评估为导向的学科建设，给了他们一种最简便、最直观和最有效的管理手段。高校中的评职称、定编制、圈重点以及各种政策倾斜和人财物的资源配置，都通过以学科评估为导向的学科建设名义来进行，因为这样很简单、很便捷、很有效率，但这种看上去最有效的方式，其实并不符合科学和大学的发展要求。

正因为如此，人们说学科评估附着和捆绑了太多的利益因素；也正因为如此，学科评估才受到如此强烈的关注和重视。

要说问题，首先是一级学科评估的一级学科就存在问题。例如教育学只有一个一级学科，用一级学科内二级学科设置齐全和均衡发展的尺度去衡量，非师范类高校的教

育学院就很难实现结合高校自身优势的教育学科发展创新，不利于其办出特色和个性发展。因此，这种基于一级学科目录的学科评估是一种对学科发展的扼杀机制，而不是鼓励新兴学科、边缘学科创新发展的良性机制。

其次是行政化的评估方式使高校普遍陷入一种难以名状的评估焦虑。由于对评估结果不仅影响学校声誉而且有可能与资源分配挂钩的预测和恐慌，“应付”好学科评估就成了高校中学科建设的头等大事。高校为了能评出好名次，绞尽脑汁、想尽办法，不惜对学科进行关停并转、整合取舍，甚至不择手段、削足适履，无所不用其极。这样评出来的结果，纵然有好名次甚至资源获取上的优势，但对于高校和学科的长远发展又有什么正面的积极意义呢？

再有就是评估导向的行政化思维误导评估指标和评价标准。大家对第四轮学科评估所谓A刊目录、国外发文、人才标签的质疑，反映了对以行政化思维评价学术问题的诟病。如此学科评估，对高校和学术发展的不利影响极其严重。

这些问题如不解决，学科评估不如不评。否则，对高校和学科、学术发展的负面影响多于正面价值。

陈学飞：我们现在是用最简单的办法来管理最复杂的事情。

熊庆年：一些省如法炮制，采取同样的政策，省属高校也采取同样的举措。比如最近《某省人民政府关于统筹推进一流大学和一流学科建设的意见》颁布实施，有4所高校成为省重点支持的国家一流大学建设第一层次高校，8所高校成为国家一流大学建设第二层次高校。全省高校共有17个学科获批世界一流学科建设项目，37个学科获批国家一流学科建设项目。该省并不是传统的高教强省，但是政府部门如此大张旗鼓推进一流大学和一流学科建设，不顾历史积累和客观现状，很有点“大跃进”的味道。

陈学飞：上行下效，中央怎么做，省里也怎么做。

刘海峰：刚才诸位对学科评估，包括第四轮学科评估提了不少批评性的意见，不过我觉得要客观、理性、辩证地看待学科评估。正如王英杰老师所说，当今世界已经进入一个大学排行榜的时代，我们国家也进入一个排行榜的时代。1983年美国的《新闻与世界报道》做了美国大学排行，后来慢慢出现各种各样的学科排行，包括世界大学的排行也越来越多。从学校总体的排行榜来讲，尽管批评排山倒海，但排行榜还是越来越多。不同的是，大部分国家是依赖新闻媒体包括公司等非官方来操作的。中国原来也都是非官方机构，而且教育部原先对各种综合排行榜是持反对意见的。因为综合排行榜是不同的东西弄在一起，比如农林的跟工科或者美术学院全部弄在一起排名当然是不一样的，梨子和苹果不完全是同一回事，大家觉得比较科学的还是学科排行榜。学科排行榜比综合性的排行榜，还是可比性强得多，所以学科排行榜一般来讲比综合排行榜更能得到大家的认可。综合排行榜也有它存在的道理，但

是学科排行榜更能体现公正。

相对其他的学科排行榜而言,我觉得教育部学位与研究生教育发展中心主持的评估的可信度还是高的。当然有一些问题,现在问题在哪里呢?就是越到后面大家越看重。刚开始第一轮的时候,南京师大、东北师大没参加,有的是管理者疏忽没注意,有的是不想参加。到第二轮大部分参加,西南大学没参加。到第三轮差不多所有学科比较强的学校都参加。到第四轮基本上可以说没有一所大学敢不重视,虽然也有可能个别不参加。文件上明确自愿参加,而且讲清楚这个中心不是一个行政机构,是一个事业单位,所以它没有行政手段或者行政命令叫大家必须参加。第三批结果出来后,大家觉得评估结果总体上可信度比较高,就教育学学科而言,我也看了一下,跟大家心目中的排名预期相差并不太大,比如说北师大排名遥遥领先,确实总体上实力比较强。

但是现在我觉得问题出在教育评估的结果跟各种利益绑定得太厉害了。特别到了第四轮,据说,排名前1%到3%政府给800万,排名前4%到10%之间是给500万,排名前11%到20%给300万。评估结果跟利益高度挂钩,不仅位置靠前的殚精竭虑往前挤,位置在后面的很多学校现在也要求各个学科争第一,至少看看自己到底排在什么地方,所以现在名义上是自愿,实质是很多人不敢不自愿,不得不主动。其结果便是,这个评估完全被功利化了,与利益捆绑得太紧了。我们进入一个排行榜时代,不管你愿不愿意,你都要被裹挟进去。其他的一些排行榜都不是自愿的,到时候他给你排的,你也没有什么办法,像美国这样,个别大学不送资料,考生就不知道这个学校的情况,结果大家都投降了,所有的学校都向这个排行榜投降。所以我说,这是一个排行榜的时代,无论你情愿不情愿,你都得参加。现在的问题是,怎么来改进它。

龚放: 排行榜绑架了大学。

王英杰: 关于排行榜,还是有些反对声音的,美国著名的商学院前几名都退出了,中国国内非得有这样的大学带头退出,否则的话,就必然被排行榜绑架。

刘海峰: 这件事你是没法改变的。

叶祝弟: 现在虽然评估结果还没有出来,但是其厉害之处已经显示出来了,我有一个例子可以佐证大家的担忧。我看到一则报道,最近,兰州大学校方正式发布文件,宣布撤销该校教育学院及内设机构。兰州大学之举并非个案,早先南开大学、中山大学、山东大学等综合性大学,对其教育学院、高等教育研究所等教育相关机构进行了不同程度的调整或撤裁。在“双一流”建设的背景下,综合大学教育学科的命运不济,一些大学撤销教育学院建制,改为研究所或者合并到其他学院。大家如何看待近期教育学学科在一些重点大学先后遭遇的拆并的命运?

董云川: 对于教育学学科近期在一些重点大学先后遭遇

的拆并和异动现象,本人未觉丝毫意外,而且在某种程度上说早在预料之中。总体而言,我有三点看法:

第一,我们还没有进入依照教育规律来导引教育实践的时代。包括学校的生存方式,学科的存在与否,以及大多数在院墙内辗转竞争的芸芸众生,影响大家生存与发展命运的主要动因大多存在于非教育政策及其游戏规则之中。犹如改革之前的大型国有企业一样,生也好,死也好,与经济家的建言献策并无直接相关,与社会经济发展的规律也没有必然联系。高等学校的当下的活法以及活的滋味,难道不正是这样的吗?一点都不奇怪。紧跟社会资源,紧随上级政策,即可左右逢源。至于所谓的教育规律,有也罢,无也罢,均属于下位法则,只要学生扎堆,横竖不会关门,院校生存并无退出风险,更无因教育自身衰败而遭清退的个案。

第二,在当下的评估以及质量甄别的语境中,教育学科无论怎样蹬打,即便使出浑身解数也难以在指标上为所在院校做出让领导瞩目、且有助于提升院校排名的骄人业绩。原因首先在于教育学科的人文属性,极少可能产出超越时空、创新求异、惊天动地的理论和试验模式;其次在于教育学科无论是对院校发展的贡献和促进作用,还是对学生的成长引领与关怀,以至于对不同学科教育改革行为的影响往往是间接的,绝对无法直接测量。润物无声,绿叶陪衬,奠基帮扶,也许正是教育学科的理性且科学的作为方式。换言之,教育学科的宿命不是冲锋陷阵,而是辅佐培植。要求教育学科争先创优,当然是勉为其难了。

第三,教育学科由于自身内涵方面的短板以及长期以来依附于外部政策诠释的生存路径,再加之既无周期性的教育实验,也无独到的理性方法,鲜见本质意义上的实证研究,更缺少存在价值的不可替代性。长此以往,教育学中大多数人陷入了“体制化”的陷阱,逐渐适应了新的游戏法则,少数人沉醉于热闹的现象研究,个别人自身能力渐行渐软,其实也拿不上台面,于是乎,当然“有”或“无”这个学科,均可。仔细想想,古往今来,许多哲学家本身就是教育家,许多思想家其实也是教育家,许多科学家最后还是归类于教育家,甚至有一些直接推动社会发展的政治家们也或多或少扮演着教育家的角色。奇怪的是,教育学者却极少成为教育家。这难道仅仅是偶然现象吗?

这样看来,有没有必要保留教育学科,是否需要重视教育学科,在何种意义上呼唤并建设教育学科,已经不言自明!

问题六: 学科评估标准存在窄化和固化的风险

龚放: 我回应一下海峰教授的观点。双一流建设“坚持以学科为基础”的原则,意味着无论是“世界一流大学”建设,还是“一流学科”建设,其出发点和立足点

都应当是提升学科的现代性、前沿性、竞争力与影响力。因为同一学科具有可比性，因而基于学科评价的大学评价，是相对靠谱、相对合理的。海峰教授等正是在这一前提下认可并赞成学科评估的。

但是，学科评估以及基于学科评估并由其推进和衡量的学科建设，必须努力避免两个潜在的威胁，即“窄化”和“固化”的风险。所谓“窄化”，即滥用某一门学科（或者某一类学科）的评价理念和方式，去评价所有学科的绩效与成就；所谓“固化”，即拘泥于某一学科的视野与疆界，胶柱鼓瑟，不思超越，不懂变通，因而难以适应现实发展的需要和变化，贻误学科发展乃至国家发展的大计。

学科是人类认识客观世界形成的知识体系的基本要素。知识体系有诸多类别、诸多分支，犹如“龙生九子，不成龙，各有所好”，不同学科之间是有显著差别和不同特性的。我们习惯上将学科体系区分为自然科学、工程技术学科、人文社会科学等几大类。中国还按照“学科门类”、“学科大类（一级学科）”、“专业”（二级学科）三个层次来进一步区分、设置。英国苏塞克斯大学的托尼·比彻，是长期从事学科分类及其特性研究并卓有建树的学者。他在上世纪80年代即提出了学科类型的“两分法”，即“规则性结构领域（areas of contextual imperative）”与“联合性结构领域（areas of contextual association）”。前者“具有严格定型的解释顺序，每一项新的研究成果在整个知识的描述中有它适当的位置”，它们“像一棵树一样不断生长，每一新的树枝又依次生长出新的嫩枝”；后者则“由许多观念群构成，没有明确具体的框架”，好像“河在流淌，流向无常”。后来比彻又借助科尔布等有关“硬科学—软科学”和“纯科学—应用科学”的范畴，进一步创立了“学科四分法”，即包括“纯硬科学”、“纯软科学”、“应用硬科学”和“应用软科学”等四类知识体系的学科分类框架。无论我们认可、采纳哪一种学科分类法，都必须承认，不同的学科有着不尽相同甚至截然不同的特点、特性，不仅研究对象的特点不同，而且知识发展的程序和性质、研究人员和知识的关系也有很大差异，研究成果的信度和验证标准，以及研究成果的表现形式，也各不相同，各有千秋。

例如，托尼·比彻所说的“纯硬科学”，即物理学、天文学等自然科学，其使命在于格物致知，揭示奥秘，探求新知；其知识生产特征是“累积的，原子论的（晶体状的或树形的）”；其研究“与普遍、数量、简化相联系”，具有客观性和价值中立性；“对知识的验证和知识的陈旧有明确的原则”，即可以证实，或者证伪；“研究成果为某种发现或对某种现象进行解释”。而被比彻命名为“纯软科学”的人文社会科学如历史学、人类学等，知识生产过程是“反复的、有机的（与河流相似）”，研究“注重细节、质量与复杂性”；具有主观性，“受个人价值观的影响”，“对知识的确认标准和知识陈旧的标准存在争议，

就所需要解决的重大问题缺乏一致意见”，“研究成果为对某种现象进行理解或鉴赏”。很明显，这两类学科的特性与知识生产程序及验证标准、成果表现形式等，几乎完全相左。我们怎么能够以偏概全，用其中的某一类学科评估的尺度来衡量与其不尽相同甚至截然不同的众多学科呢？又如，像机械工程学、微电子光电子技术学等属于“应用硬科学”的研究，目的是解决问题，付诸实用（实践），其研究成果为产品的设计与开发，是技术的集成或更新换代，是新兴产业的发展以及市场的开拓……我们怎么能够用高影响因子学术期刊论文的数量、标准化论文引用影响指数（NCI）、被引用的频次等等来评价、衡量这些学科发展水平与研究实绩呢？再如，我们用标准化论文引用影响指数（NCI）来评价那些以“去标准化”和“受个人价值观影响”为特点的学科（如文学、艺术学等），岂不是闹了缘木求鱼同样的笑话？

问题七：本末倒置：作为预警手段的评估异化为目的

施晓光：评估很重要，也是必须的，但是正如王英杰老师所说，关键要明确为什么要评估？如何进行评估？首先，应该知道，评估只是一种手段，而非目的。从国际高等教育评估的趋势看，很多国家在开展评估的过程中，将更多的精力放在质量保障（外部保障和内部保障）上去。评估的目的是为了保障质量，评估工作的出发点和归宿都需要落实到质量保障上去。然而，现在的评估太过于看中评估结果，因为这个结果直接和利益、资源的分配挂在一起，这样的评估取向显然存在问题。

其次，我们需要对评估正本清源。评估实际上是构建一套预警机制，它绝不是目标，如果把评估作为一个目标，并去追求这个目标，那么问题就大了。在这个问题上，政府是有责任的。首先，政府的责任就是要考虑如何发挥评估的杠杆作用，合理调整资源配置，以评促建，以评促改。其次，在评估的过程中，评估者和被评者都必须对评估工作和学科发展的规律性予以充分的认识和尊重，即对学科自身发展的一个逻辑有一个清醒的认识，评估不能急功近利，不能把一个非常复杂的问题简单化。在评估之前，评估者以及各方利益相关者应该多方共同协商一个评估标准，标准绝不是由一家制定的，学校、学者、政府，各个利益主体之间不一定能够很快得出共识，但是协商讨论的过程本身也是评估的一个部分。第三，评估标准的制定既要考虑我国高等教育现有的发展水平，也要参照国际标准。例如，在双一流建设过程中，什么是一流大学？什么是一流学科？这里不能自说自话，需要学术共同体的普遍认可和承认。不是说自己是一流就是一流的。自我评价是一回事，让人家心里完全承认你是一流则是另一回事，后者更难。

问题八：两个“绝对化”倾向影响评估健康发展

樊秀娣：施晓光老师讲得非常对，评估未必就是排名，评估只是一项评价活动。教育部学位中心第四轮学科评估开展至今反响之强烈不出意料、情理之中。因为此轮评估带有较浓厚的官方色彩，又处于国家“双一流”建设实施方案和配套政策尚未明朗时期，因而对此轮评估的理解和定位就变得相当敏感和武断，认为评估结果挂钩学科去留、挂钩“双一流”建设、挂钩财政拨款的说法烽烟四起，至今也未见权威部门予以明确说明。现实中也确实出现有的学校直接对弱势学科关停并转的现象。如此种种，着实令人担忧并质疑评估的初衷和目的。比如，动静如此之大，是否真实体现了本轮学科评估的目的和意图？而本轮评估的评估指标、技术及最终的评估结果的科学性又如何？现今参与评估的各方似乎都心有默契，那就是本轮评估的结果将直接挂钩国家“双一流”建设，而且学科排名相差一、二，国家投入却可能“失之毫厘谬以千里”。在如此“致命”效应预期下，学校会否因用力过猛而做出非理性抉择？学科团队会否因过度焦虑而无心苦练内功？如是，那么保证质量、以评促建岂不是句空话？弄不好，还可能产生负面效应。审视各方表现，我们可以发现，在本轮评估中弥漫着一种“绝对化”的态势，它影响着本轮评估的健康发展，审视整个学科评估活动，主要有两个突出问题：

一是把本就片面的评估结果绝对化。任何以定量形式表现的教育评估结果都是相对的，复杂的学科评估更是如此。而本轮评估的最终结果不仅会以量化的形式出现，而且还会以排名的形式列出，这样精确的量化结果能获得被评者由衷的认可吗？答案显然是否定的。就本轮评估的指标而言，如A刊论文的计量，显然就难以承担对教师科研成果评价的重任，即便后期有大幅度的修正和完善，仅就价值逻辑推断，学术论文中蕴含的科研成果水平是否可由它发表的期刊决定？答案显然也是否定的。还有，本轮评估指标中，有不少是需要人为打分的主观评价指标，且占的比重非常大，但它们的客观性和可信度恰恰又是被评者担心和质疑的。再有，师资队伍中的人才指标，因在各类人才评选中不同类别的高校机会不均等，现在要以此来判断师资水平的高低，很多学校是不买账的。事实上，本轮学科评估主管部门在评估思路、规则、指标、手段等方面进行了重大调整，力求真实反映各校各学科的质量和水平，但靠显性的指标和结果来量化学科间的绝对差距，是难有作为的。

把评估结果绝对化，导致一些学校和团队为了评估结果而做出非理性的举动。本来评估规则设计的意图是通过积聚资源、优化配置来加强学科建设和提升实力，按照此

逻辑应该是：因为有了相关资源和成果，所以就能保证学科的水平；而现实的逻辑是：因为能够筹措到这些资源和成果，所以就能使评估名次上去了。如此“因果倒置”，出现了各校为了评估结果和名次，把大量精力花在现有资源和成果的“排列组合”上，还有不少学校内部的学科之间为了争抢资源和成果而闹崩，有的学校干脆关停并转那些在本轮学科评估中无望进入前茅的学科，全然不顾学科发展的自然规律和生态效应，这对于学科发展和学生培养哪有益处？对此仅以长官意志、行政权力来归因恐怕也不全面。学科评估应该是建筑起一条对学科质量和人才培养监控和保障的底线，把它作为学科水平名次之争在国外未见。

二是把本就有限的评估功效绝对化。学科评估最大的功能和效应无疑是服务于学科的发展和建设，而学科的水平关乎学校人才培养和科研产出的质量和效益，为此，学科建设和学校水平互为因果、相互影响都是有道理的。从这点出发，把学科评估和“双一流”建设关联起来也是有道理、成立的。但学科评估和其他外部评估一样，都是在一定参评对象、评估标准下的状态呈现和评判，它的根本作用就在于帮助参评学科在相互比较中知己知彼，从而能够更迅速、准确找到各自学科建设中的问题与对策，树立标杆，保证质量和提高水平。但学科发展本身的构成要素、渊源脉络和学科间的相互关系等，并不是学科评估能够反应和解决的问题，因此直接把评估结果作为其他相关政策制定的唯一依据显然是粗糙和不科学的。

就拿资源配置和学科排名的关系来说，国家若只是以学科排名或评估结果为依据来进行资源分配，将资金集中到学科排名靠前的几所大学，其合理性，有待商榷。因为这些大学的学科本身已发展得很成熟了，再不断地给予大量投入，这样产生的效率可能不会太高，效果也不会很明显，还会导致大学之间的失衡，即好的大学越来越好，弱的大学越来越弱。一流大学和一流学科对社会资源的“吸引”能力本身就是其水平和良好“造血”机制的体现。试想，如果中国的世界一流大学和学科单是靠国家的资金堆积而成的，这样的成就会有多少人由衷地诚服？而这样的“高度”又有多少可持续性可言？当然，这并不是说国家不要对“双一流”建设的大学和学科重点投资，我只是反对一味依据学科排名的掐尖投资。

问题九：由行政部门认定A刊的做法不妥

马陆亭：我来自政策研究机构，从政策方向来谈几个基本观点：

第一，评估本身作为一个质量保障手段是符合方向的，评估管理实际上是从直接管理向间接管理的过程，评

估是对的，不能因为评估工作中出现了问题而质疑评估本身。公立大学公共资源的投入应该可以而且需要评估，这是我的基本观点。

第二，不能赋予评估太多附加职能。大家为什么如此看重学科评估？因为赋予的职能太多了，在过渡时期，咱们国家在行政管理中出现了很多这样的问题，包括跨世纪人才、新世纪人才、长江学者、院士，一旦当选后，可以带来无穷的资源。由此，现在的学科评估，大家潜在的认识是跟未来的“双一流”建设挂钩的，所以过度重视进而产生“恐慌”应对现象。评估本身是对的，但是对附加的东西看得太重，这是我们制度设计上出现了问题。

第三，政府是由不同部门组成的，而且不同部门下还有事业单位性质的中心，确切地讲就不是政府部门了。当然外界说你挂着教育部的抬头，就代表政府，有时是这样的，因为承担了部分转移了的政府职能。我本人是从事政策研究工作的，也来自于这样的事业单位。作为一个事业单位的人，内心当然希望说的话有公信力。各部门在抓工作的时候，有可能出现抓权现象，这时候就受到部门利益和个人利益的纠葛。比如说作为教育部的统称和教育部里面的司局和直属机构其实是不同的，司局为什么有时候抓的问题非常细呢？这时候就可能涉及到部门和个人利益了，作为教育部整体就不涉及到这种利益，但是作为一个教育部的部门会涉及到部门利益。政府由部门组成，部门再派生出若干机构。我们在今后的改革中，需要意识到如何协调和克服部门自身的利益冲动等这些问题。

第四，为了在评估中取得好结果，一些学校做材料，甚至数据造假。这样的行为严重违背了教育的基本精神，即使为公也不行！因为学校是教育机构，学校为了利益而作假，学生看到后自然就不会因为个人私利造假而脸红，社会的诚信怎么办？这是个重大的原则问题。

第五，在现有的情况下，不管是政府还是政府派生的部门，都不应该定义哪些杂志是最好的。当然政府部门会感到很委屈，说我们定的这些杂志是有道理的，程序上没有问题，都是经过专家论证的，但是其客观后果确实是做好事帮倒忙。当政府部门认定它是最好的时候，这些杂志无形中就会炙手可热，接下去大家就会去公关，这对整个学术的生态发展是不利的。如果政府或者政府的部门一定要去认定的话，我个人认为认定一个更宽泛的范围，因为更宽泛的范围就是把那些不好的淘汰出局。

陈学飞：还有一点不得不说的是，一开始推出的方案，评价学科水平，仅靠若干杂志发表的文章数量，以此来衡量这个学科水平，其结果便是引导大家去发表文章，而且一定要在规定的所谓A刊发表文章。但是，这能真正反映学科和学术发展的真实水平吗？

叶祝弟：所谓以刊物级别作为评价文章质量的标准，

无疑于买椟还珠。这份A刊名单推出后，确实引起了比较大的争议和批评，后来虽然匆匆收回，但是其引发的反响还是比较大的。这里我将学界的评论收集、整理为九个方面的问题，供大家参考。第一，教育部学位中心推出A刊，虽然初衷是“鼓励优秀成果优先在国内发表，增强我国期刊的影响力，改进学术论文评价方法，引导学术论文由数量评价向质量评价转变”，但是这样的初衷不过是一厢情愿，由行政部门出面遴选A刊，实际上是利用行政的强力在学术刊物中人为划出一个所谓的第一阵营，这不符合学术期刊发展的规律，真正好的学术期刊是在长期的发展过程中自然淘洗、遴选而形成的，标准只有一条，即学术共同体的口碑。第二，鉴于A刊的指挥棒效应，高校管理者必然将教师入职、晋升、考核与此挂钩，加剧高校内部的竞争和失衡。第三，A类期刊一旦与高校老师的学术生存权挂钩，必然引发发文者趋之如鹜，容易滋生公关等学术腐败行为。第四，学术论文的质量并不能机械地由所发刊物的级别来衡量，学术论文的创新最好由学术共同体来认定。虽然权威期刊能保证总体质量，但是它也会错杀一些真正具有锋芒和创见的论文。第五，虽然学科评估是以评价学科建设为主，但是今天人类面临的是复杂性、综合性的难题，依靠某一学科的力量并不能解决复杂的难题，人们越来越意识到问题导向、跨学科、多学科联合的重要性，这方面综合性刊物可以弥补专业刊物单一性、视野窄的局限。综合性刊物和专业刊物之间应该是互补的关系，两者缺一不可，而A刊中很少有综合性刊物入选，这不利于跨学科研究，更不利于复杂问题的解决，根本上不利于学科发展。第六，就A刊目录而言，为了强调国际话语权，可以引导自然科学、社会科学类论文在国际期刊上发表，但是对于中国人文学术类成果，国内刊物最有发言权，用国际刊物来评价中国人文学术成果，有失妥当。第七，就学科内部而言，还存在二级学科不同专业的平衡问题，那些边缘学科、交叉学科、新兴学科很难顾及，不利于学科的协调发展。第八，此次公布的A刊，地方社科院（社科联）主办的刊物鲜见入选。在中国学术期刊方阵中，地方社科院（社科联）主办的刊物因其刊发文章质量较高等享有不错的声誉，但是此次却很少有此类期刊入选，这不利于学术生态的平衡发展。第九，学术评价方式应该多元，期刊发表只是一种，学术评价应该是一个多元、综合的体系，成果评价的导向应该是为解决重大现实问题提供思想和理论资源，而不是仅仅自娱自乐或解决晋升等问题，更不能隔膜于社会现实。

问题十：评估“进高校”干扰了高校正常教学工作

周川：除此之外，还有一些劳力费神、匪夷所思的行为。比如，我们搞学科评估，每个教授要填大量的表，包括

你培养优秀的毕业生要提供两到五个，并填写他们300字的事迹。如何看待这种劳力伤神的行为？大家已经有很多的讨论。中国的高校如今也面临着越来越多的评估，这或许是一个必然的趋势。这些评估，既有外部的，也有内部的；既有评高校整体的，也有评学科专业的；既有行政性的，也有社会自发的；既有排名性质的，也有诊断性质的。这些评估，形形色色、五花八门，我们不能笼统地说是好还是坏，是福还是祸。好或坏，福或祸，在很大程度上取决于评估本身，取决于评估的目的、指标、主体、方式等要素。

评估的方式，也就是“如何评”的问题，也是很重要的。就评估“进不进入高校”这小小的一点来说，就大有讲究。我们可以将评估分为两类不同的方式，一类是不进入高校的评估，另一类则是进入高校的评估。这里所谓“进入”和“不进入”，不仅仅是指评估的“人”进入或不进入高校，更重要的还在于评估的“事”进入或不进入高校。不进入高校的评估，不仅评估的“人”不进入高校，而且评估的“事”也不进入高校。比如当前流行的许多“大学排名”，虽然其科学性广受质疑，但它在做排名的过程中，“人”和“事”都是不进入高校的。它们主要按照设定的指标，收集高校的公共信息和数据，通过对这些公共信息和数据的处理进行排名，整个工作过程在高校之外进行，不惊动高校，也不需要高校填表格做材料。这类不进入高校的排名，科学不科学另当别论，但它至少不骚扰你，也不强求你认可、接受，你愿意把这些排名当回事就当回事，你不把这些排名当回事也完全可以不当回事，完全可以不把它放在眼里。排不排，是他的自由；当回事还是不当回事，是你的自由。

然而，进入高校的评估，情况就大不一样了。评估的“人”进来了，“事”当然也就同时带进来了；有时候“人”虽然没有进来，但“事”却一样不少地进来了。一般来说，能够进入高校的评估，都是有来头的。在当前的高教管理体制下，能够进入高校的评估，说白了都是行政权力主导的评估，或者是借助于行政权力进行的评估，他们都是行政性的评估。这种评估的特点是，整个评估过程由行政部门（或准行政部门）发起、操纵，一是涉及范围广，往往是全国性或区域性的；二是自上而下长驱直入，层层贯彻，层层施压；三是搞运动的方式，评估的过程无异于一场运动，或者说是一场战役。这一类进入高校的行政性评估，“运动程序”不外是：行政部门或准行政部门下发评估红头文件；高校进行“迎评”（抑或“迎战”）动员；从校方到院系再到学科甚至个人，层层签订“军令状”；而后，教师填写相关表格并提供相关材料（比如“权威刊物论文”、“优秀毕业生事迹”之类）；院系汇总，编制院系一级的材料；校方审核并模拟初评，上下同力将评估材料“涂脂抹粉”打扮得“花枝展俏”；最后将一大堆表格和材料提交上级评估部门。由于这种行政性的评估往往与资源配置联系在一起，高校把这些评估看作是“生

死之战”也是事出有因，因此，每当这类评估进入高校，评估本身就成为高校的一项突击性工作，甚至是“压倒一切”的“重中之重”工作，不惜牺牲大量的人财物力，不惜牺牲正常的教育教学秩序，不惜牺牲教师宝贵的时间和精力，其评估过程，实际上很容易变成“瞎折腾”的过程，变成对高校正常教育教学工作粗暴干扰的过程，更不要说“大学精神”、“学术诚信”所受到的伤害了。

评估本来就是一把“双刃剑”，始终利弊相随。评估目的、评估指标、评估主体、评估方式等要素本身的品质，都决定着评估的性状和利弊。就“进不进入高校”这一点而言，在当前的体制下，假设其他评估要素大体相当，那我们就有理由认为，不进入高校的评估是相对比较“良性”的评估，而强行进入高校的评估，是相对比较“不良”的评估。换句话说，在其他评估要素相当条件下，两害相权取其轻，我们可以容忍各类不进入高校的“不良”评估，但不能容忍各类强行进入高校的“不良”评估。

当然，由于行政权力的强势，行政性的评估要进入高校，是由不得高校的事情，不是高校能够左右的。但是，从高校的一方来说，如何来应对进校的评估，其实也是有高下之分的。是采用行政手段，层层下压，全员参与，劳教师之力去填表格做材料一直做到天昏地暗、怨声载道，还是将应对评估的工作交给一个内设机构去操作？这也是检验一所高校“现代管理水平”的一个重要判据。

二

陈学飞：既然存在这么多问题，那么到底如何科学评价学科发展和水平？国际国内有没有什么经验值得学习？各位专家可否提供建设性意见？

建议一：学科评估不可急功近利，更不能搞招尖投资

王英杰：第一，评估很有必要，但是现在的问题是推出比较仓促。第二，为什么说现在比较仓促，是因为大家都会因为自己的利益而改变对评估的的总体认识。我们真正要推进学科评估，我们需要搞清楚，为什么评估？评估的目标是什么？由谁来评估？评估的结果如何使用？把这些问题弄清楚了再来评估。我想，与大学评估相关的各方应该找准自己的位置，发挥各自的作用，为完善大学评估做出各自的贡献。

政府应该明确自己的作用，不应该自己直接开展大学或学科评估，要制定政策鼓励和促进第三方评估机构的建立、健全和发展，同时对建立的第三方机构进行必要的监管，使其依法依规开展评估。政府还应该制定政策集中规范大学的数据统计口径，要求大学定期发布准确的数据，从而为评估提供数据依据，逐渐使评估建立在可靠的数据

基础之上。大学应该把定期发布全面和准确的数据作为一项经常性的工作和一项必须履行的社会责任。大学、公众和决策机构要正确地解读评估结果。大学应该将评估作为一个诊断工具,寻找自己的不足,改进自己的服务。

建议二:推进管、办、评分离, 多管齐下培育第三方学科评估机构

周川:王老师的发言我很赞同。我来回应一下谁来评的问题。就“学科评估”而言,“谁来评”的问题也是至关重要的。“学科评估”的专业性特别强,由综合性的准行政部门来组织实施就很不合适,最好的实施主体应该是各个专业学会。专业学会是最有资格、最有权威来组织实施“学科评估”的机构,例如,高校数学学科(或院系)的评估由中国数学会来组织实施,高校化学学科(或院系)的评估由中国化学会组织实施,高校心理学科(或院系)的评估由中国心理学会来组织实施,如此等等。由专业学会组织“学科评估”,最大的好处就是专业性和权威性。评估采用什么指标,以什么方式进行?是诊断还是排名?是定量评价还是专家投票?诸如此类的关键问题,只有专业学会的专家才有可能做出最适合本学科特点的选择,评估结果才有可能有更大的適切性和权威性。除此之外,由专业学会组织“学科评估”,也有利于高等教育评估主体的多元化,发挥专业学会在学科建设和发展中应有的作用,有利于消减行政权力对学科专业事务的干预,这也是符合当前行政体制“放改服”的改革大方向的。

熊庆年:与“谁来评”相关的第三方评估不可能从天上掉下来,如何走出困境?我以为,当下之要务,是多管齐下培育第三方评估。

首先是改造准官方的教育评估机构。不管怎么说,已有的这些机构还是一支专业化的力量,他们在原有的体制框架下,取得了权威地位,建立了信用,有业务资源的积累。可以在事业单位改革的总体框架下,将这些评估机构与教育行政部门剥离,不再具有隶属关系,取消编制管理,使之成为独立的非营利的法人机构。重建这些机构的治理结构,建立有政府教育行政部门、高等学校、专业学者、产业界和社会各界人士代表参加的董事会,实行董事会领导下的院长或所长负责制。其运行资金的来源,可以借鉴英国高等教育质量保证署(Quality Assurance Agency in Higher Education,简称QAA)的模式,一部分来自直接的公共财政资助,这种资助应是不附加条件的,并相对固定的;一部分来自高等院校参加评估交纳的费用。当然,也鼓励社会捐助。这种改造的好处是,可以相对保持这些机构的大体稳定,保持与过去政策的连续性,又使其独立的特质逐步建立起来。

其次是发展学术社团评估组织。我国高等教育有庞大的社团组织,中国高等教育学会下有分支机构、省市机

构、行业分会、大学分会122个。这些组织中有大量的政府和高校在岗及离岗的专业管理人员、研究人员,由这些社团来组织评估活动,可以保证专业性和与高校的有效互动。但是,社团并不能简单地就赋予其第三方的功能。不言而喻,现有的这些社团都具有官方性,也是高度行政化的。从中国高等教育学会到所属各分支机构和分会,负责人往往都是由相关教育行政部门安排的,有的还有事业编制,有财政的拨款。学会运行也基本上是沿袭行政运作的方式。所以,要在整体的社团改革中去“行政化”,使其回归学术本原,转变为真正的学术共同体,具有民间性和独立性,自主地开展学术活动。在这个前提下,学术社团成立专门的评估组织,利用专业的资源开展评估活动,才可能成为可信的第三方。

再次,要大力培育高校间的自治性评估组织。近年来,一些高校结成联盟,以共同应对现实的挑战。这类组织的民间性、平等性、协商性特点显著,离第三方性质比前两者近得多。通过这类组织来培育出第三方评估机构,也是现实可行的。美国的经验告诉我们,这类组织有强的自治性,组织内部能够较好地达成共识以谋求发展,形成特定的标准以实现自律。以北美研究型大学联盟(Association of American Universities,简称AAU)为例,什么大学可以成为其成员,有自己的一套标准,并以此来开展评估。本世纪初,它接受了两所大学为其成员,也劝退和除名了两所大学。它持有的标准和规则具有很强的约束力,评估活动也得到所有成员的认可。美国各式各样的大学联盟很多,各种各样的评估也很多,形成了不同的圈子和不同的评估,从而构成了多样化的评估文化和发展生态。美国的经验对发展我国的第三方评估具有借鉴意义。

最后,要积极扶持民间非营利组织。尽管数量极少,但是我国社会民间组织还是在教育领域里生长出来了,21世纪教育研究院就是这样一个组织。近些年来,它也开展一些评估活动,在高等教育领域产生影响的,是2015年开始发布的《年度高校信息公开情况评价报告》,2016年已经第二次发布,对最受公众关注的教育部直属高校2014年的信息公开情况进行了第三方评价。与其同时发布同类报告的还有中国社会科学院法学研究所,其《中国高等教育透明度指数报告》是对政府信息公开工作开展第三方评估中的一部分。法学研究所有“公家”的性质,而21世纪教育研究院完全是民间的,它可以完全不看政府的“脸色”,独立开展评估,十分可贵。对这类民间组织政府应当通过购买服务来给予政策上的支持,使它们能够生长、发展,在高等教育评估中发挥越来越大的作用。

事实上,民间还有一类组织在参与高等教育评估,比如麦可思数据有限公司,是家提供高等教育管理数据与解决方案的专业公司,每年编著出版《中国大学生就

业报告》(就业蓝皮书),为高校、社会大众、用人单位和部分政府机构提供数据服务。它自诩是第三方权威数据机构,但不管怎么说,它是营利的公司,我们把它作为第三方机构的理由是不充足的。前文提到的汤森路透、爱思维尔,也都是数据公司,它们是不是能够“中立”“客观”地评估,需要我们睁大眼睛。政府购买它们的服务也应该谨慎。

能不能培育好第三方评估机构,政府是关键。除了转变角色、转变职能之外,政府最需要做的是创造良好的法治环境。要为评估立法立规,为第三方建立法制保障,为评估活动立规矩、定规范。另一方面,要强化政府的监督责任,系统地建立制度,对第三方实施严格的认证,以及有效的问责,促进第三方评估公信力的提高。另外,还要进一步大力推进信息公开、公共数据共享,使得第三方“有米下锅”。总之,要着力推进教育的管、办、评分离,把培育第三方当作教育治理体系与治理能力现代化的一个有机部分。

建议三:建构融合评估的事实判断与价值判断于一体的学科评估体系

卢晓中:我觉得首先最重要的是弄清楚学科评估价值立场问题,即将为什么评这个问题弄明白了,学科评估的正当性问题也就迎刃而解了。如何进一步明确与学科评估的目的与功能相一致的价值立场,从评估目标与过程等方面融合评估的事实判断与价值判断来建构起学科评估体系,这也是走出具有中国特色的大学学科评估之路的基本路向。

如果说“双一流”建设是以学科评估的结果为前提和依据(由于国家最终的政策并不明朗,这里只能用“如果”来表达两者关联的可能),那么从中至少可解读出两方面意蕴:一是学科及其建设是大学及其建设的基础,大学发展须以学科发展为重。这也与长期以来教育部主张并实施学科评估,而不赞成大学的综合评估排名的立场相一致(教育部曾多次声称,不赞成目前国内形形色色的对大学进行综合排名的大学排行榜,认为此类综合性的大学排行榜不科学)。二是学科评估能够有助于学科发展乃至大学整体的发展,关键在于怎样评。实际上,任何一项评估都涉及到为什么要评、怎样评、评的结果如何用等评估诸要素,其中,“怎样评”在整个评估中占据主体地位,因此,在怎样评中解决好评估的科学性和精准性,对于引导大学发展的正确方向和体现制度设计的公平公正是十分重要的。

上面前一点解读主要涉及到我们究竟需要的是“双一流”建设,还是着重于“单一流”建设(一流学科建设)。应当说,着重于一流学科建设,给更多高校拥有了更多选择的可能,由此也更能体现高校分类发展的现代理念与时代趋势,更符合高等教育大众化的发展规律与特征。世界

一流大学是世界一流学科建设的必然结果。如果在一流学科建设的同时又通过一种类似综合性的大学评估排名的方式来确定一些大学作为一流大学建设目标,姑且不论这种确定方式是否科学(比如与大学这一特殊社会组织的性质和现代高等教育发展规律是否相符等),极有可能使如此重要的国家高等教育政策仅仅成为少数大学来玩的游戏,这对于我国高等教育形成良性的发展生态和高等教育发展系统的合理构建未必是有益的。

后一点解读则是如何解决好“怎样评”的问题,也就是要明确学科评估的科学性、客观性和精准性问题,这里需要探讨的是学科评估中的事实判断与价值判断问题。一方面,学科评估过程所追求的科学、客观、准确更倾向于事实判断;而另一方面,学科评估的目的与功能取向(诸如导向、管理、监测、诊断等目的与功能)又决定了学科评估中价值判断的必不可少。应当说,无论是何种评估都会涉及到价值判断与事实判断问题。强调评估的事实判断,是从评估本身的客观性而言的,即价值中立或价值无涉,尤其是评估过程中尽量排除或减少人为(如评估者等)的主观影响。那么,在什么情况下评估可以有价值判断呢?用日本比较教育学者冲原丰的话来说,就是在首先声明了自己的价值立场后可以在事实判断的基础上进行价值判断。在学科评估中明确自己的价值立场往往是通过确立学科评估的目的与功能,进而具体化为学科评估的标准与指标等,来体现出一种价值判断。事实上,依据不同的价值立场确立的评估标准与指标往往不同,自然,评估的结果也相去甚远。比如在国际上颇具影响的三大大学排行榜中,由于所持的价值立场相异,最终得出的评估结果也有一定的差异。如进入《泰晤士高等教育》世界大学排行榜前100名的中国大学有4所,进入QS前100名的中国大学有9所,而在上海交通大学的排行榜中,却没有中国大学进入前100名(2016年排行榜已有2所大学进入前100名)。

由此可见,学科评估中的事实判断与价值判断是密切相关的,很难将某一种学科评估简单归为事实判断或是价值判断,而往往是兼而有之,也就是说事实判断中有价值引导,而价值判断中则常常以事实判断为依据。值得提及的是,明确学科评估的价值立场是十分重要的,唯有如此,方可适当兼融学科评估中的事实判断与价值判断。此外,在兼融事实判断与价值判断时避免出现“休漠的铡刀”现象,即把事实如此当做本应如此,这是在确立学科评估目标与指标时需特别给予关注的地方。

目前正在进行的大学学科评估,从评估指标、评估过程、评估方法等,较之以往都有了许多改进,评估也试图从方法和技术层面处理好科学性与精准性、操作性(如对毕业校友信息的采集等),国际化与本土化(如社科领域的A刊设立等)等关系。但也毋庸讳言,本次学科评估

所秉持的价值立场并不明确，或者过于宏大笼统，抑或语焉不详。究其原因，我以为这与本次学科评估的目的与功能尚不清晰是有关联的，一方面学科评估的目的与功能受制于瞄准世界一流大学和一流学科的目标，而如前所述，此目标所关联的价值立场是非常多元的，以致无所适从；另一方面，为避免学科评估过于功利化导向，如可能导致的各大学争名次或排位，本次评估又宣称淡化学科排名，着重于诊断与监测的目的与功能。特别是评估组织方对评估的结果如何使用这一关键问题一直秘而不宣，这究竟是出于工作策略的考量，还是目前尚难以确定，或是两者兼而有之？诸如此类的疑虑问题也都反映出本次学科评估的目的与功能尚不清晰。如果有第五轮学科评估，我希望方案设计师首先要将这些问题想清楚。

建议四：政府管好底线，尊重丛林生态，培育自主生态，坚持二八分治，理性把握边界，适度有限作为

董云川：真正的大学学科构成了复杂多变的丛林生态，大、小、强、弱各适其位，自然筛选之后必然形成其独特的生长样式。比如，我们自己的教育学科连“省级”都算不上，但反过来，整个省域范围内高等教育的事务除了教育厅直接行使的行政职能之外，许多事情都有我们施展作为的身影。换句话说，我们是一个小学科，但是我们的作用有效辐射于区域教育体系之中。作为一个小学科来说，这难道不是其服务社会的价值所在吗？另一方面，我们所有的老师都潜心于研究生教育，认真地行使着育人的天职，但这一切并不直接显现于任何的排行榜中。我认为这才是教育或学术的本职，也正是学科教育以及教育学术的本真。

针对当下以及今后还要继续开展的学科评估，我的建议是：尊重丛林生态，培育自主生态，坚持二八分治，理性把握边界，适度有限作为。评估一定是需要的，但是，政府只应该管底线，不应该关心“忽悠”（上线数字指标）。上行下效，高等学校中出现的一些错误取向大多是“忽悠”惹的祸。

优秀的学者和名牌的大学当然都不是“评”出来的。要不然怎么会有出“类”、拔“萃”的表述。近年来的评估造成了一种假象，就是当一所学校、一个学科或者一个学者能够不偏不倚地被框进一个贴满了指标的大柜子里，即具备了所谓“优秀”的条件时，它就成为好学校、好学科和好学者了。因此而使得所有的人都去关心“贴标签”而忽视了“做事情”。当然，这些所谓“优秀的指标”原本都是有根有据的。就大学而言，本来是某个个体在探究真知的过程中超越创新、兴利除弊、新陈代谢之后的种种特征，后来被证明确实促进了学术和社会的进步发展，继而提高了办学水平，成就了自身的卓越品质。其后，一旦

这些事件或者人物之内隐的品质被忽略而外显的特征被固化为标准之后，挑选出来成为“优秀”的指标模板就已经具有了两面性——一方面以榜样的力量促进事业发展，一方面以框框套套的形式阻碍新事物的超越。一朵花因其芬芳独特而被称之为玫瑰，并非因为叫做玫瑰才拥有独特芬芳。爱因斯坦不是评出来的“名师”，哈佛不是评出来的“名校”，贝尔实验室不是评出来的“重点基地”，马克斯·韦伯的组织理论不是评出来的“精品课程”，奥本海默也没有被授予“优秀教学成果奖”。评估行为无疑可以实现督促监管，却不能对质量提升有什么直接的作用。探究发现并尊重实践事物本身发生发展的规律才是质量的真正尺度。评估只是裁衣用的那把尺，称重时的那杆秤，而不是制造优秀大学、优秀学科、优秀学者的万能机器。

总之，我的态度很明确，学科建设、质量管理是学校自己的事，少管才是智慧！

建议五：学科评估的立足点是学科建设，培育卓越的学科群体

施晓光：关于怎么评才科学的问题，我认为，任何评估手段都一定是服务评估目的本身。试问我们为什么要评估？如前所述，评估无外乎要实现以评促建，以评促改，达到质量不断提升，也就是说能够保证学校的健康良性发展。这里我们可以引进一个心理学上的原理：双因素原理。所谓双因素原理就是说对于事物激励的因素有两种，一种叫保障性激励。另外一种叫奖励性激励。用这个原理解释评估，就是说保障性激励可以对所有被评院校产生影响，即政府投入评估以后，对所有院校都产生促进作用，让大家觉得都有碗饭吃，不能让大家觉得因为评了别人伤了自我的感觉。在大家都得到评估的好处之后，再拿出更多的资源用于奖励性激励。表扬先进，示范后进。总之，政府在评估过程中的首要职责就是要先建立保障性机制，然后再考虑真正的鼓励问题。如果做不到这一点，最后谁对评估的结果都不会满意，就算通过评估拿到资源的学校也不满意，它们也会觉得被政府搞得精疲力尽，也会抱怨评估不具有可持续发展性。

评估机制需要有一个共用的指标体系，这个指标体系应该是科学的，经得起实践检验的。现在的评估都是工具化的，如学校获得多少奖项、有几个长江学者，国外根本不知道这是什么，与国际都不在一个话语体系中。因此，我们首先搞清楚，国际上学科评估的基本标准是什么。

另外，评估标准不是一成不变的，评估过程本身也需要不断改进。在评估的过程中，不论是评估者还是被评估者都需要不断发现评估中存在的问题，发现问题后要不断的改进，这样才能进行更好的评估。

马陆亭：施老师你刚才说，让人有碗饭吃，我不太同意这个观点。为什么不同意这个观点呢？说点我个人经历

中的体会,我发现国外的一些学校(在我们看来是一些很低端的学校),他们不是仅仅满足于有碗饭吃。当他们谈论起自己学校的时候,很是自豪,他们觉得所工作的学校一点都不逊于哈佛、牛津、剑桥。这不是个别人,而是一批人,他们每个人在跟我们介绍他们学校的时候,都充满了自豪感。我们为什么不为我们的学校充满了自豪感呢?所以我个人觉得,不能满足于有碗饭吃,而要让每个职业、学科、学校都能把自己建设得更好,并让其成员充满自豪感。我觉得这是我们目前所欠缺的。

施晓光:我同意马老师的观点。在评估的时候,绝对不能只有一套指标体系,学科评估很重要,但是学科评估不是对整个大学的评估,因为大学不是一个学科发展的。但是整体学校的发展,你把所有的资源都捆绑在一个学科上,学校该如何发展?在西方,已经没有 first class 的概念,取而代之的是卓越。每个学校都卓越了,还需要那么在意谁是第一吗?

胡建华:实际上,只有弄清楚什么东西能够反映学科水平,才能制定出好的评估指标。我觉得我们的评估指标体系有很大的问题,比如说,QS 现在也给中国的大学学科排名。但是这个学科排名指标很少,并且对我们学科的认识只是基于学科声誉或者是该学科所发表的文章。我以为我们的学科评估指标体系太复杂,而且有的东西本来就是不公平的竞争,但是却把它作为评估指标,这肯定是不公平的。比如说人才称号,部属院校和省属院校在申报的时候就是有差距的。像新世纪人才,部属院校是每个学校下指标,而省属院校全部下到省里面。长江学者也是。好比百米赛跑,有人是从 100 米处起跑,有人却从 80 米处起跑。所以说,这个指标本身就不合理,你再把它作为评价指标,这就无法评价出学科水平,我觉得学科水平的评价指标体系越简单越好,不能太复杂,我们应该找到能够反映学科水平最根本的东西。比如说产出,科研情况,这个实际上是反映学科水平的,或者是大家都公认的指标。

建议六:摒弃以论文论英雄的功利取向

陈学飞:我认为衡量学者水平,不在于发表多少篇文章,出版多少本著作,拿了多少项课题,拿了多少经费,而在于对本学科的理论有什么新的贡献,对国家和民族的理论创新和进步有什么贡献,这个应该是学科追求的一个基本目标和价值基石。此外,学者对整个社会的影响也应该用来衡量学者的水平,但是谁来评价学者对社会的影响?我以为应该由该学科最好的专家来评,由懂行的专家来评,由学术共同体来评。

马陆亭:对于学科评估,第一,我同意海峰教授的观点,学科评估水平公信力更强。为什么呢?这是因为相对

而言,学科更具可比性。第二,学科评估虽然有相对的对比性,但是我们绝对不能再走掐尖的道路。比如,为什么只选 100 个优秀学科?能不能选 1000 个?让第五名觉得同样有机会发展自己,第五名跟第一名的差距又有多大?并且很多时候是不可比的。

今后的学科评价如何开展?我的建议可以概括为以下四个观点:

第一,适当进行分类,比如说学术型的、应用型的,国家也是这样在推动发展的,政策要呼应起来,还有涉及国防的学校能以发表论文而论吗?

第二,重视产出,输入、产出的指标最好不要同时出现,因为有重叠现象,可分类设计典型的产出指标,鼓励学校更有效、有特色地开展学术工作;

第三,突出代表性成果,以鼓励大家潜心做研究、面向问题做研究、做真正的研究;

第四,不要轻易否定别的成果形式,比如说对刊物的选择,你可以筛选一批刊物给予高权重,但不能否认其他刊物上的成果,可以给它较低的权重嘛。一般而言,人们发表文章的逻辑是,先在一般刊物上发文章,慢慢地再到所谓的核心期刊、C 刊、A 刊上发文章,但现在连研究生都要直接在 C 刊上发文章,这其实是不符合人才成长规律的。一旦变异,就会产生种种“故事”,不利于良好学术生态的建立。

建议七:学科评估应分门别类制定标准,切忌一刀切

龚放:我认为,人文、社科、自然科学应该遵循不同的规律,学科评估也要分门别类制定不同的标准。前面董云川谈到丛林生态,我们学科的评估,实际上是要促进丛林生态的各个层次,乔木、灌木、植被各方面蓬勃的发展,用一句话就是“鹰击长空,鱼翔浅底,万类霜天竞自由”。但是我们现在的评估实际上是一个丛林法则,丛林法则是赢者通吃、弱肉强食,我们就评出第一,然后给予大量的奖励,最后必然损害整个中国学科的发展。有些学科是理论贡献,有些学科是解决问题。评估不能因为评出第一名,而把第二名及以后的,都给打击了。是促进丛林生态的繁荣,还是导致丛林生态法则的赢者通吃,取向完全就不同了。

美国学者约翰·S. 布鲁贝克曾经用“在学问的圣殿里有许多厅堂”来比拟大学学科的多样性。在现代大学,既有“通过在隔音的实验室里拨控制盘来验证真理的”学科,也有“通过在喧闹的城市、福利中心、诊所、法院等地方参与工作来积极验证真理的”学科,还有一些“在静寂的图书馆里通过钻研故纸堆来验证他们的思想”的学科;有些学科“仰望星空”,有些学科“关注脚下”。不同的学科有其不同的使命和功能,它们的历史经纬、发展方

向与优化路径也不尽相同。学科特点的多样性呼唤高校学科评价的科学性和合理性。尊重多样性,走出学科评价的“窄化胡同”,摈弃“以论文论英雄”的价值取向,是我们建立科学、合理的学科评价制度的第一步。在这方面,南京大学有过自己的经验和教训。

南京大学在 20 世纪 80 年代中期率先采用科学计量学的方法,重视 SCI《科学引文索引》等国际检索系统在理科科研、学术评价及博士生培养中的地位 and 作用,曾经取得了十分明显的效果和骄人的实绩。笔者曾经在 2007 年西澳大利亚大学国际学术研讨会上提出一个观点:“在肯定南京大学在 SCI 等‘学术榜’问题上的历史性贡献的同时,也要清醒地看到,SCI 等等仅仅是科学评价的工具和指标之一,它相对客观、科学,但决不是唯一的、十全十美的。SCI 也好, EI 或 ISTP 也好,‘核心期刊’、‘影响因子’或‘被引频次’也好,都只是从一个方面、一种角度来衡量研究成果的原创性和影响力;它们不能取代科学家们的‘同行评审’,不能取代对研究内容、方法与结论的深入探讨和多方评价。”而且,过于强调学术论文的数量和质量,过于强调 SCI、EI 和 ISTP 论文,还产生了始料未及的负面影响,抑制了、妨碍了相关工程技术学科的新建和发展。当时,南京大学校长曲钦岳已经意识到,现代工程技术学科的薄弱和空缺是南京大学学科发展的“短板”所在。因此他果断提议并在学校管理层形成决议:每年拿出留校和引进人才的不少于 5 个百分点,专门用于引进工科人才。这本是一个很有远见的设想,也是一个具有相当力度和可操作性的学科建设举措。但是,因为南大的理科思维、理科氛围太强势了,对 SCI 论文的奖励举措也影响甚大(尽管当时只是每篇 SCI 论文奖励 100 元),工科人才引进以后难以立足,难以施展才干,更别说渐成气候了。几年以后,曲校长派人专门就有关院系引进的工科人才发展状况进行调查,发现其中相当一部分已经离开了南大,而剩下的也争相去发 SCI 论文,大多被“规训”和“调教”成为理科人才了。

我所亲历的南京大学在学科建设上的成功与受挫说明,大学的学科发展必须从学科评价、学术评价的变革入手,必须形成多样化的学科评价机制和体系,形成不同版本、不同价值导向的“学术榜”。只有这样,才能营造一种“鹰击长空,鱼翔浅底,万类霜天竞自由”的环境和氛围,让不同学科都找到自己发展、生长的空间和可能。只有这样,才可能有真正的学术繁荣、真正的学科发展,才可能有跻身“世界一流”又彰显“中国特色”的学科与大学!

建议八:尊重高校办学自主权, 建立多方协商的对话机制

樊秀娣:针对目前学科评估中出现的一些问题和各方观点,个人认为,可以通过明确评估目的和理念、完善技

术与方法、规范组织与实施等来保证学科评估的理性发展,主要建议如下:

第一,把学科评估回归到学科建设本身。学科评估是学科建设的重要内容和途径之一。学科评估的目的是监测各大学学科建设的成效和发展水平,保证教育质量。通过各项指标,为各大学提供一把自我审核和相互比较的尺子,同时也为社会大众了解、监督大学办学提供一种参考。如果人为地赋予学科评估太多的职责,人为地把评估结果的使用绝对化,如把学科评估结果和国家“双一流”建设方案、财经资助政策等紧密“捆绑”,表面上抬高了学科评估的身价,事实上这是学科评估本身无法承载的,客观上也会加剧整个学界的浮躁和急功近利心态,这样的评估走向与其以评促建的评估初衷是相背道的。我们说,国家的各项教育政策、方案自有其关注的重点和价值判断,单靠一轮学科评估来解决大学发展和竞争中的重大问题本身就是有悖科学决策常理的。更何況,本轮学科评估正处于国家“双一流”建设方案尚未明朗时期,又缘于评估主管部门的官方色彩,所以处在风口浪尖的学科评估对结果的发布以及结果可能被各种“挪用”的状况更要审慎对待。离开“绝对化”表达形式,更能彰显评估活动自身的水准和价值。

第二,由学科共同体来实施学科评估。本轮学科评估由教育部下属的学位与研究生教育发展中心来组织实施,虽然学位中心自己定位为第三方机构,但以其操作的手法和过往的表现,高校和学界中没有人会把它当作第三方机构来看待或对待。事实上,比较复杂的评估问题,类似学科评估,在还未形成成熟模式和达成共识的开发阶段,尤其需要较为宽松、探索的氛围,由行业学科共同体组成的自律组织来实施更为合适。因为学科共同体内部更容易建立学科文化、语言和评估标准,同行更熟悉学科发展的基础和方向,同时行业内部出台的各类评估项目,沟通、交流、调整和修正也更为灵活和方便,总之,这种同行内部、平等对话的评估组织更容易获得被评方心理上的接纳。虽然目前这样的学科共同体评估组织还未形成,但同行学科共同体组织是松散存在的,也许现在正是培育和扶植第三方评估机构的好时机。

第三,建立多方协商对话机制。在现阶段,无论是哪种性质的评估机构开展学科评估,其指标体系、技术方法、程序步骤等都需要经受实际验证并不断修正与改善。如本轮评估指标体系中,出现了许多需要相关者主观评分的指标,而对此具体如何操作,对指标得分的影响很大。为此,所有涉及评估的各项要素应该在评估前就向社会大众公示,接受各方的评议和置疑。这里要强调的是,为了获得社会各方尤其是被评方的响应和支持,同时也保证评估活动的科学性和公平、公正性,在制定各项评估指标和操作方案的过程中以及在形成完整评估预案后,评估主办

机构都要留有专门听取各方诉求的空间。其中，协商对话机制是实现充分吸纳各方意见，聚焦问题，缩小分歧差距，保证各方在目标和行动上一致性和匹配性的有效策略与途径。我认为好的学科评估指标体系的形成应该让被评学校都有最充分的发言权。

第四，尊重高校的办学自主权。理论上说，本轮学科评估，是国家依法治教形势下管办评分离实践的体现，更要注重科学性、规范性和独立性，评估活动开展的宗旨应是服务于学校的自我质量保障和问责制度，所以评估活动的重点更应体现在诊断上。为了切实发挥学科评估的诊断、导向和激励作用，应该把评估的着力点放到学科的展现、交流、剖析和比较上来，立足于学科发展实际，提供多维度、多层面的价值资讯，支持学校根据自身优势和发展需要来统筹学科设置和建设，形成良好学科生态群。学科评估应尽量淡化或不做量化结果和排名，不去左右学校对学科设置的自主决定权。

总之，个人认为，不出定量的评估结果，不搞“绝对化”的学科排名，也许是把评估活动引向更健康发展的一种明智、可行的选择。

建议九：学科评估尽量不要“进校”， 不要影响高校正常教学秩序

周川：从技术上说，美国高校的“院校研究”，是很值得我们借鉴的。相比之下，美国高校面对的评估，比我国的高校可能要多得多。1980年代后期，日本著名学者喜多村和之教授研究指出，美国大学“强盛”的原因之一，就是“大学从里到外被种种评价包围着”。美国不仅有六大区域的高校认证机构，有《美国新闻与世界报道》的大学排名，此外还有数百个形形色色的专业协会，都定期地在对大学进行评估，更不要说美国高校还有大量的内部评估了。当然，在美国除了极少数高校认证机构之外，其他的评估都是不进入大学的；即便是一些认证活动可能要进入大学，其进入的方式，与我国的行政性评估也是根本不同的。但是，随着高等教育问责制度的不断发展，一些评估机构或问责主体，也开始逐渐影响高校，比如要求高校补充提供某些没有在公共数据库中反映出来的信息。美国高校应付这类事情的秘诀之一，就是它们发达的“院校研究”。美国绝大多数的高校，都设有“院校研究办公室”一类的机构，内中有数位至数十位不等的、专事“院校研究”的专业人员和辅助人员，他们的岗位职责就是研究本校的校情、校务，包括建设学校基本数据库（“院校研究办公室”一般都兼作学校数据中心）、分析研究学校教育教学及管理工作中出现的问题并提出解决方案、撰写学校年度报告、应对评估和问责机构的需求，等等。美国高校的“院校研究”，在专业性和工作效能方面，绝对是世界一流的。正是依靠“院校研究”专业的、高效率的工

作，所以美国的高校才能在形形色色评估的包围之中应付裕如。

目前我们最关注的，可能还是正在进行中的新一轮“学科评估”，因为我们许多老师刚刚从既伤身又伤神的“做材料”中挣脱出来。眼下，“材料”已经上报，结果尚未出炉，为此，利益相关者们心中惶惶不安，都是可以预料到的。我们这一轮“学科评估”，是由综合性的准行政部门来组织实施的，按照统一的指标体系来评估不同的学科，自上而下进入高校之中，因此，它也就免不了那种“进校”行政性评估的种种弊病了。下一轮评估，这可能是需要克服的问题。

樊秀娣：确实如此，本轮学科评估，不少学校为填写报表花费大量精力，基层中也不乏怨言，究其原因，主要是临时突击收集、整理数据费时费力。这反映出目前国内多数高校，即便是冲击世界一流的大学，大多都还没有合格的数据库。完整、准确、动态的数据库是有效评估的基础，一方面它能够保证评估数据的真实性，减少数据差错及人为作假的可能性；另一方面它能切实提升评估活动的价值。评估活动的开展和效用的体现很多依赖于数据的分析和展示，有效的数据能让人发现的问题，界定需要改进的领域，了解变化并能预测未来。没有数据，难以聚焦问题，评估就变成了应付性的行为，以评促建也仅是一句漂亮的口号了。国内高校数据库建设问题一直存在，教育部应该引导建立全国统一的高校数据采集系统，逐步累积有关高校运行的主要数据。

第十，谨慎使用评估结果，尽量避免评估结果与资源、利益的直接挂钩

王英杰：最后要特别提出的是，应谨慎使用评估结果，尽量避免评估结果与资源、利益的直接挂钩。政府如要将资源分配与学科评估结果建立一定联系时，要综合考虑教育投入的公平、公正、透明、和谐和高效等原则，不仅要奖励排在最前端的大学或学科（应该不是某所学校、某个学科而是一组大学、一组学科），而且要关注排在尾端的大学和激励多数大学。经合组织的教育部长们一致认为，“将评价结果与院校激励制度联系起来，包括提供额外的资金，可以成为质量保障和质量改进及推动院校间相互竞争的强有力杠杆。同时还应注意到，评价也可能导致国家内（院校间）差异的扩大，这对制定政策鼓励处于评价尾端的院校改进质量是一个巨大的挑战。”

叶祝弟：今天的讨论可谓意味深长，而又意犹未尽，相信它对中国的学科评估在理念和实践方面的充实和完善将起到积极的推进作用。谢谢大家！

整理 王鸿
编辑 杜运泉