

DOI: 10.16750/j.adge.2023.12.011

# 如何以评促建： 荷兰学科评价三十年回顾与反思

杨佳乐

**摘要：**学科评价是保障学科建设质量的关键环节，唯有平衡好内部评价与外部评价的关系，方能形成良性的质量保障闭环。但当外部评价凌驾于内部评价之上，极易引发评建倒挂、以评代建，抑制学科建设主体自主性。荷兰经过协会评价、标准评价与战略评价三个发展阶段，学科评价逐渐从制度化、标准化走向特色化。重点通过允许学科评价口径自主选择、扩大学科评价指标遴选范围和调整学科评价结果呈现方式等举措，激发学科建设主体自主性，从而实现以评促建。回顾荷兰学科评价三十年的制度变迁，有必要反思评价口径与学科边界、评价指标与建设维度、评价结果与卓越话语之间的关系，重建学科评价秩序。

**关键词：**学科评价；学科建设；荷兰；标准评价；战略评价

**作者简介：**杨佳乐，中国社会科学院中国社会科学院评价研究院助理研究员，北京 100732。

## 一、问题的提出

学科评价顾名思义，是以学科为评价对象所开展的事实判断与价值判断活动。根据评价主体划分，既有学科建设单位自行组织的内部学科评价，也有政府、营利或非营利的第三方机构开展的外部学科评价。内外部形形色色的学科评价一定程度上发挥着保障学科建设质量的积极功能。借用伯顿·克拉克经典的三角模型，内部学科评价意味着学术力量充当学科建设质量“守门人”，外部学科评价则可视作学科建设主体回应国家力量与市场力量问责的表现。在学科分化初期，学科评价只是学术界内部之事，因而当时只有学科内部评价，基本没有学科外部评价。然而学科作为一种资源依赖型组织，其发展需要不断从外界汲取有形的人力、物力和财力资源，以及无形的声誉资源，国家力量与市场力量通过控制资源供给，对学科组织的影响程度日益增强，对外部学科评价的关注度与日俱增，与此同时争议却也随之而来。一个争议焦点是，外部学科评价以绝对优势压倒内部学科评价，主导着学科建设方向，学科建设主体的自主性得不到有效激发，学科评价与学科建设之间出现分离<sup>[1]</sup>，以评代建、评建倒挂

现象普遍存在，以评促建难以实现。在学术、国家与市场三股力量中，由于学科身处学术共同体与科层制组织的二维空间<sup>[2]</sup>，使得评价学科时学术力量与国家力量的博弈尤为激烈，且呈现国家力量强势主导而学术力量被动应对态势，具体表现有三。

其一，外部学科评价的评价口径固化，可能抑制学科建设主体发展学科交叉与交叉学科的主动性，继而对创新造成负面影响。世界一流学科生成于制度化与去制度化的内在张力之中，制度化为学科知识的生产、传承提供合法性，去制度化则为学科的特色发展、创新发展创造条件<sup>[3]</sup>。制度化的学科评价却与学科交叉、交叉学科存在着天然矛盾<sup>[4]</sup>。学科评价建立在学科高度分化、学科边界清晰的基础上，而学科交叉、交叉学科往往意味着学科整合和边界模糊。英国与经费拨款挂钩的学科评价虽然专门制定了学科交叉成果的评价办法，却还是因为评价口径问题抑制了不确定性更高的创新研究，并且对人文社会科学领域的负面影响显著大于自然科学领域<sup>[5]</sup>。其二，外部学科评价指标存在“测不准”问题，可能诱使学科建设主体被动跟随评价指挥棒，异化建设行为，偏离建设方向。指标是现代学科评

基金项目：中国社会科学院“青启计划”项目“科研人才评价制度：变革逻辑、创新效应与优化建议”

价制度的核心要素,但如同量子力学领域存在“海森堡测不准原理”(Heisenberg Uncertainty Principle),学科评价中也存在评价行为与评价对象之间的“测不准”问题,一旦将某一指标纳入学科评价指标体系,就可能引发学科建设主体不惜一切代价追求该指标高绩效表现的非理性行为,那么该指标的评价效度将会大打折扣,未被纳入评价体系的指标则被选择性忽视,从而悖离学科评价的原初目标。其三,外部学科评价青睐定量评价结果,可能引致学科建设主体向排名靠前的机构看齐、被动模仿,继而加剧学术漂移,形成组织趋同。定量的学科排名结果形塑起高等教育系统的丛林法则与马太效应,排名靠前的学科组织资源丰富,排名落后的学科组织则发展受限,甚至面临被裁撤风险,学科生态遭到破坏。迫于生存危机,学科建设主体不得不整齐划一地对标本学科领域处于第一梯队的头部机构,无暇顾及自身特色。

追根溯源,外部评价压倒内部评价是学科评价主体不平等性的一种表现<sup>[6]</sup>,倘若学科建设主体的自主权得不到根本保证,以评促建自然也无从谈起。在全球范围内,荷兰为平衡内外部学科评价,激发学科建设主体的自主性贡献了一个富有研究价值的个案。1993年至今,荷兰的学科评价制度历经三十年的探索完善,逐步形成了以学科建设主体为中心,学科内部评价为基础,专家外部评价为补充的荷兰模式,为该国知识创新营造了良好的制度环境,2023年荷兰在全球创新指数排名位列世界第7名<sup>[7]</sup>,其在全球科研与创新版图中的实力不容小觑。同时荷兰也在全球学术评价改革浪潮中扮演先锋角色,积极响应国际和欧盟学术评价改革倡议,加快推进学术评价改革落实。荷兰是世界上开放科学程度最高的国家之一,2017年发布了全球首个国家级开放科学计划。2022年荷兰乌得勒支大学宣布在教职员工招聘和晋升决策中放弃使用影响因子,更加看重团队贡献及促进开放科学等方面的努力<sup>[8]</sup>。荷兰如何以评促建充分调动学科建设主体的自主性?从回顾与反思荷兰学科评价三十年的变迁中可窥探一二。

## 二、荷兰学科评价的制度变迁与阶段特征

荷兰的学科评价制度自1993年正式建立至今

已经走过了整整三十年,大致可以划分为三个阶段:第一阶段是1993—2002年的协会评价阶段,第二阶段是2003—2021年的标准评价阶段,第三阶段是2021年至今的战略评价阶段。

### 1. 协会评价阶段:学科评价的制度化

荷兰学科评价的制度化肇始于政府和大学之间关系的变化。20世纪80年代之前,该国并未建立起国家层面的学科评价制度<sup>[9]</sup>。直到1983年,荷兰政府引入有条件的经费分配方案,根据大学的科研质量与社会影响力分配财政经费,经费开始分为教学经费和科研经费两部分。为争取更多的科研经费,大学必须不断提高其科研绩效。伴随着尼夫(Neave)所言的“评估型国家”(Evaluative State)的崛起,荷兰政府开始重新谋划与大学之间的关系,试图通过评价高等教育系统的输出质量来代替传统的政府官僚法令管制形式<sup>[10]</sup>,由重过程监控的先验性评估向重产出质量的后验性评估过渡。随后在1985年颁布的政策文件《高等教育:自治与质量》中,荷兰政府明确表示不再自上而下直接管理大学事务,而是采取远距离调控(steering from a distance)的大学治理哲学,通过设置高等教育系统运转主要参数的间接方式引导大学发展方向,扩大高等教育机构的办学自主权<sup>[11]</sup>,政府与大学之间的关系随之发生转型。因为荷兰政府认为,赋予大学自主权,加强大学自我管制将有助于大学提升科研质量和科研效率,进而激发创新活力<sup>[12]</sup>。

那么大学在扩大办学自主权后的学科建设成效究竟如何?要回答这一问题就需要建立全国性的评估体系,对大学的学科建设情况进行评价以应对问责,学科评价作为一种“治理术”也逐渐成为荷兰政府调控大学学科建设方向的政策工具。于是在1992年,荷兰议会通过《高等教育与科研法案》,提出大学必须在独立的外部专家的参与下,实施周期性的学科评价,制定相应的评价规范来确保有效的自我管理,为保障科研质量负责。为回应该法案,大学与荷兰教育部达成一致意见,同意由荷兰大学协会(VSNU)负责制定国家层面的学科评价方案,以外部专家评价作为内部质量控制的有益补充<sup>[13]</sup>,并开始在一定学科先行先试。1993年2月,首先选择机械工程、生物学、心理学和历史学四个学科进行试点,标志着

荷兰国家层面的学科评价由此实现制度化。次年根据试点结果对评价方案予以修正,随后按照修正后的评价方案对其余学科展开评价。1998年荷兰启动第二轮学科评价,评价方案与第一轮评价基本保持一致。

### 2. 标准评价阶段: 学科评价的标准化

由于荷兰的学科评价以外部专家评议为主要评价方式,尽管同行评议能够在一定程度上克服量化评价的弊端,不过同行评议也并非金科玉律,时常面临评价过程易被操纵、评价结果可比性差、主观性强等抨击,并且当时的荷兰学术界存在多头评价问题,因此迫切需要统一评价数据填报格式,尽可能减轻参评单位的数据填报负担<sup>[10]</sup>,因而荷兰开始考虑推进学科评价的标准化。1999年,荷兰大学协会、荷兰研究委员会(NWO)和荷兰皇家艺术与科学院(KNAW)共同成立“学术研究中的质量保障”工作小组,在此基础上三方合作研制了标准评价协议(Standard Evaluation Protocol, SEP)。自2003年研发SEP以来,荷兰已于2003—2009年、2009—2015年、2015—2021年实施了3轮评价。评价由VSNU、NWO和KNAW共同负责组织实施,评价周期固定为6年。

标准评价协议对学科评价的评价目的、评价原则、评价主体、评价客体、评价标准、评价流程等做出了明确规定,对研究型大学与科研院所同等适用,旨在实现全国范围内学科评价的标准化,尽可能统一评价标准,规范评价流程,相当于荷兰版的学科评价国家标准。具体而言,SEP自评阶段的评价主体为各参评单元(research unit),专家评价阶段的评价主体为大学或科研院所的董事会,以及由外部专家组成的评价委员会,二者之间属于委托代理关系。评价客体在SEP中称为参评单元,由大学或科研院所的董事会认定。评价标准虽然在几轮评价中有所调整,但科研质量、社会影响力和可持续发展能力均为核心标准。评价流程大致包括评价准备、单元自评、专家评价、公布结果四个阶段。在评价准备阶段,由大学或科研院所的董事会负责确定参评单元和评价委员会的职责范围(terms of reference),随后组建评价委员会并任命评价委员会秘书。在单元自评阶段,参评单元需依据既定模板撰写自评报告。进入专家评价阶段,评价委员会专

家将综合考虑参评单元提供的自评报告以及实地考察情况,在实地考察后的20周内向董事会递交最终的评价报告。董事会根据专家给出的评价报告出具立场声明(position document),并在机构官方网站上公开专家评价报告和立场声明,以接受政府问责与社会监督。每个评价阶段都规定了时间节点。

### 3. 战略评价阶段: 学科评价的特色化

在SEP 2021—2027中,战略取代标准成为荷兰学科评价关注的焦点,英文缩写尽管仍为SEP,但字母“S”所指代的内容已由标准(standard)转变为战略(strategy)。相较于标准评价,战略评价以评促建的功能导向更为突显。评价目的由统一标准、规范流程转变为找准定位、鼓励特色,进一步弱化了学科评价的行政化色彩,充分尊重参评单元的主体地位,注重激发参评单元紧密围绕发展战略规划学科发展的内生动力,而非被动地迎合评价指挥棒。

战略评价强调基于每个参评单元自身的发展战略展开评价。在自评报告中,参评单元需详细阐述其过去6年的战略目标(strategic aim)和战略计划(strategic plan),以及实现战略目标和战略计划的战略过程(strategic process),并提供翔实的支撑材料。支撑材料可以采用指标或案例的形式,但禁止使用期刊影响因子和H指数以避免评价谬误。因为期刊影响因子属于成果层面的评价指标,H指数属于个人层面的评价指标,而SEP的评价层次定位于参评单元。战略目标的制定应聚焦其对科学知识和经济社会的贡献。战略计划可围绕如下要素展开:①如何组织科研人员;②科研成果发表渠道的选择;③社会影响力创造渠道的选择;④鼓励其成员申请的资助类型;⑤认为谁是其科学和/或社会合作伙伴;⑥如何与这些合作伙伴互动;⑦如何将开放科学纳入科研过程;⑧如何规划博士政策与培养;⑨学术文化的类型;⑩引导和选拔人才、营造多元化环境的人力资源政策<sup>[14]</sup>。战略过程可从谁参与了战略制定、如何监测战略规划的实施、战略目标或战略规划是否有所调整以及调整的原因等方面陈述。参评单元未来6年的战略规划需使用战略分析工具SWOT框架,逐一呈现发展优势(strengths)、劣势(weaknesses)、机遇(opportunities)和挑战(threats),并与国际或国内同行对标,找准发展定位。

由于参评单元的发展战略特色各异,不易取得广泛共识,因此战略评价特别需要吸纳多方代表、秉持多元视角,更加全面地评价参评单元的战略目标、战略计划与战略过程,故而战略评价阶段评价委员会的专家构成从学术界专家和非学术界专家进一步拓展到博士生和处于职业生涯初期和中期阶段的科研人员,并且尤为强调专家性别、国籍、文化背景和专业背景的多元性。

### 三、荷兰以评促建的重点举措

#### 1. 允许学科评价口径自主选择

20世纪末 VSNU 制定的评价协议也曾规定荷兰全国层面统一的学科评价口径,但随后发现将不同战略定位的同一学科机构进行比较,无异于拿苹果和橘子相比,几乎不具备可比性<sup>[15]</sup>。荷兰政府意识到必须赋予高等教育机构更大的自主权,以增加高等教育系统的适应性、灵活性与创新性,并将滥觞于军事领域的“战略”、发展于商业领域的“战略规划”等概念引入高等教育领域<sup>[16]</sup>。于是从 SEP 2003—2009 开始,荷兰打破了固定的大学学科评价口径,不再由 VSNU 组织全国范围的学科评价,而是将评价口径的选择权赋予学科建设主体,由各高等教育机构自行决定参评单元,只要战略定位清晰并满足基本的规模条件(至少拥有 10 名取得终身教职的全职科研人员)与时间条件(至少成立 3 年),均可作为参评单元。若参评单元的学科交叉特征明显,则遴选具有丰富学科交叉经验的专家进入评价委员会,以保证评价的信效度。作为研究型大学和科研院所定期开展学科评价的纲领性指南,SEP 从制度上保证了学科建设主体选择学科评价口径的自主权,打破了纵向与横向双重边界。打破纵向边界是指评价口径覆盖了不同组织层次,既可以评价大学和科研院所,也可以评价院系、研究中心、研究所,或是科研团队;打破横向边界是指评价口径跨越不同知识领地,学科交叉或是交叉学科组织均可参与评价。灵活的评价口径一则有效解决了同一学科不同战略定位机构的可比性问题;二则巧妙回应

了固化学科边界、低估学科交叉等质疑,为用好用学科交叉催化剂、激发科研创新原动力预留了充足的制度空间。

以荷兰莱顿大学为例,该校自主决定对考古学、治理与全球事务、区域研究、历史学、伊斯兰研究、创作与表演艺术、语言学、艺术与社会、哲学、法学、医学、环境科学、高级计算机科学、生物学、物理学、化学、数学、天文学、科学技术研究、文化人类学与发展社会学、教学法与教育科学、政治科学和心理学等学科领域,以及非洲研究中心与教学研究生院 2 个跨院系的研究机构进行评价<sup>①</sup>,参评单元既有传统院系,也有跨院系的研究中心。其中,治理与全球事务、区域研究、伊斯兰研究、非洲研究等都具有明显的学科交叉特征,归入既有任何一个主学科进行评价都无法准确刻画学科交叉图景,自主确定评价口径则尽可能降低了评价偏差,为探索交叉学科评价和学科交叉的研究机构评价积累了有益经验。

#### 2. 扩大学科评价指标遴选范围

三十年来,荷兰的学科评价内容随着国际趋势、欧盟倡议与国内实际不断与时俱进,几经调整。在协会评价阶段,主要评价学科的下述 4 个维度:①科研质量;②科研生产力;③科研影响力;④长期生存性。在标准评价阶段早期,评价维度基本保持稳定,SEP 2003—2009 的评价维度涉及:①科研质量;②科研生产力;③科研影响力;④活力和可行性。SEP 2009—2015 则进一步规定了评价维度下的评价指标,形成了便于量化的指标体系。具体而言,科研质量维度下设学术影响力、学术领导力、学术声誉、组织资源和博士生培养 5 个指标;科研生产力维度下设生产力战略、各类科研产出两个指标;社会影响力维度下设经济、社会、文化影响力指标;活力和可行性维度下设发展战略、SWOT 分析、稳健发展条件指标。不过评价指标体系一旦固定,指标“测不准”的问题也随之而来,特别是在人文社会科学领域,如何选取科学合理的指标来评价科研质量与社会影响力等均成为实践难题。2009 年荷兰

<sup>①</sup>评价报告见:<https://www.universiteitleiden.nl/en/research/about-our-research/quality-and-integrity/research-assessments/assessment-reports#archaeology,governance-and-global-affairs,humanities,social-and-behavioural-sciences,african-studies-centre-leiden-ascl>。

国家人文学科未来规划委员会发布报告《可持续发展的人文学科》，提到能够客观界定和区分人文学科质量的指标还很匮乏<sup>[17]</sup>。该委员会由此向 KNAW 建议开发简洁、清晰、有效的人文学科质量指标。这一建议随后得到采纳，荷兰专门成立了人文学科质量指标委员会负责开发相关指标。

受《旧金山科研评价宣言》《莱顿宣言》、英国“负责的度量”（Responsible Metrics）等启发，SEP 2015—2021 指标体系设计思路发生较大变化。一则取消了科研生产力，只保留科研质量、社会影响力、可行性、博士生培养、科研诚信和多样性维度。部分原因在于荷兰试图通过此举缓解科研人员巨大的科研成果发表压力，同时也传递出重质量轻数量的评价导向<sup>[18]</sup>。二则不再规定统一的评价指标体系。各评价维度下仅提供“一揽子”指标（baskets of indicators）<sup>[19]</sup>清单，从而扩大了评价指标的遴选范围，学科建设主体自主决定与自身发展战略相契合的指标（或案例）。指标清单通常由第三方机构、大学或科研院所开发。如 VSNU 研发了《科学研究的定义协议》（Definitie afspraken Wetenschappelijk Onderzoek）、《价值指标框架》（Raamwerk Valorisatie-indicatoren）及《价值指标选择清单》（Keuzemenu Valorisatie Indicatoren）。ERiC 项目组研发了《学术研究社会影响力评价指南》（Evaluating the Societal Relevance of Academic Research: A Guide）。QRiH 项目组研发了《人文学科科研质量指标》（Quality Indicators For Research in the

Humanities）。莱顿大学开发的指标清单遵循学科知识生产和交换的结果、知识利用的效果、知识开发的收益三重逻辑，依次从各学科与学术领域、专业领域、商业/公共部门、社会公众的互动设计评价指标（见表 1）。

在战略评价阶段，与欧盟发布的《下一代指标：面向开放科学的负责任指标和评价》、欧洲大学协会和科学欧洲共同发起的《科研评价改革倡议》、荷兰政府制定的认可与奖励计划“为每个人的才能预留空间”相呼应，荷兰从时间和空间两方面，形成时间上跨越过去 6 年和未来 6 年、空间上跨越学术界内与学术界外的评价指标矩阵。具体而言，科研质量维度关注过去 6 年参评单元在学术界内的学术贡献、学术声誉和学术领导力。其中，学术贡献不限于学术发表，也包括工具研发、基础设施建设等情况。社会影响力维度关注过去 6 年参评单元在学术界外的经济、社会、文化、教育等方面的影响力。可行性维度则关注未来 6 年参评单元在学术界内外的学术贡献和社会贡献的战略目标，以及达成战略目标所必须的组织管理、领导能力、资源储备等战略规划情况。此外，开放科学、博士生培养、学术文化、人力资源政策 4 个维度共同构成矩阵交集，可在评价科研质量、社会影响力与可行性时择机体现。

与 SEP 2015—2021 的评价维度相比，SEP 2021—2027 除持续关注博士生培养外，新增了开放科学，并将科研诚信拓展至涵盖开放、安全、包容、科研诚信的学术文化，将多样性拓展至涵盖组织多

表 1 莱顿大学评价指标清单

	学术领域互动	专业领域互动	商业/公共部门互动	社会公众互动
知识生产和交换结果	发表/合作发表（文章、著作、评论等）；科研项目；博士论文指导；本科/硕士教学；学会参与	发表/合作发表（期刊论文、手册、专著等）；专业讲座	在商业或政府媒体上的发表/合作发表、访谈等；面向企业职员或政府公职人员的讲座、圆桌会议讨论等；合作项目	杂志等科普发表；公共讲座；开放获取；慕课
知识利用效果	学术同行对所开发或获取的科研工具、基础设施、数据库、实验室、模型、程序、软件、设计等的使用；引用	专业人员对所开发的科研工具、模型、程序、软件、设计等的使用	商业/公共部门对所开发的科研工具、模型、程序、软件、设计等的使用；对博物馆、标准化协会等的贡献；咨询/政策建议；对社会团体的贡献	对论坛、电视、社交媒体上公共讨论的贡献；能够增加社会价值的校友
知识开发收益	科研基金/学术获奖；学术任职/期刊编委；特邀学者；排名；学术联盟中的地位；科研经费	委托研究；专业培训；专业咨询；专业人员对科研设备的使用；创业实践	委托研究；专业培训；专业咨询；公司或其他团体对科研设备的使用；专利、执照、信息披露、税收	公共奖项；付费公开课；社会团体任职；就业情况

资料来源：Leiden protocol for research assessments 2021—2027: 13-14。

多样性和人才管理的人力资源政策。具体言之,博士生培养维度重点考察博士生教育内容、结构和质量,涵盖从招生到培养再到就业全过程。开放科学维度重点考察过去6年向其他研究者和社会利益相关者开放工作的程度,如是否根据可发现(findable)、可获取(accessible)、可操作(interoperable)和可重复(reusable)原则<sup>①</sup>使用和存储研究数据,是否公开研究数据、方法和材料,出版物是否开放获取,以及在未来6年的开放科学计划。科研诚信维度主要关注对《荷兰科研诚信行为准则》的落实情况。多样性维度可评价性别、年龄、种族、文化背景和学科背景的多样性。人才管理维度侧重评价人才聘任、培养、晋升、奖励、指导、职业发展支持等政策。

### 3. 调整学科评价结果呈现方式

从学科评价结果来看,在标准评价阶段,SEP 2003—2009和SEP 2009—2015规定评价专家需对全部评价指标给出定量评价结果。到SEP 2015—2021,评价委员会对科研质量、社会影响力与可行性三个维度既做定量评价,也做定性评价;博士生培养、科研诚信与多样性三个维度只做定性评价。量化评价结果的呈现方式也在不断变化。协会评价阶段共分为5个评价等级,由低到高依次为:较差(poor)、不满意(unsatisfactory)、满意(satisfactory)、良好(good)和卓越(excellent),但各个等级究竟应该达到何种

水平并没有具体说明。进入标准评价阶段,荷兰开始对各个等级的判断标准予以统一规定。2003—2009年与2009—2015年的两轮SEP延续五级分等,在卓越与良好之间增设优秀(very good)一级,同时取消较差一级,不满意成为最低等级。到2015—2021年,SEP去掉满意一级,调整为4个等级(见表2)。

然而设定量化的评价等级意味着建构起阶梯式的衡量标准,最高等级卓越逐渐成为各级各类组织孜孜以求的目标,犹如“无形之手”,引导参评单元朝着卓越的方向走向整齐划一,导致组织同形<sup>[20]</sup>。故而进入战略评价阶段,在以参评单元自身战略为中心的评价理念下,专家评价报告不再呈现定量的评价等级,仅给出定性评价结果以及未来发展建议,旨在打破定量评价建构起的卓越神话。各参评单元将学科评价视作质量保障的重要一环,以自身战略为参照系自主选择国内或国际对标机构,不必在单一赛道竞争卓越,从而克服定量评价“指挥棒”效应下被动的同质化,专注于结合专家建议更有针对性地加强学科建设,彰显学科特色。

## 四、荷兰学科评价的反思与启示

荷兰战略评价协议的形成并非一蹴而就,而是经过三十年不断调整、优化的结果,该国科学<sup>②</sup>政策制定机构也素有关注科研系统战略变化的传统<sup>[21]</sup>。通

表2 荷兰定量学科评价结果呈现方式变迁

等级	协会评价阶段	标准评价阶段		
		SEP 2003—2009	SEP 2009—2015	SEP 2015—2021
5	卓越	卓越(国际领先,可能具有重要影响)	卓越(国际领先,具有重要影响)	
4	良好	优秀(国内领先且具有国际竞争力,期待作出重要影响)	优秀(国内领先且具有国际竞争力,作出重要影响)	卓越(国际领先,对社会作出杰出贡献,为未来做好充分准备)
3	满意	良好(国际具有显示度,国内具有竞争力,可能作出有价值贡献)	良好(国际具有显示度,国内具有竞争力,作出有价值贡献)	优秀(国际具有显示度,对社会作出重要贡献,为未来做好较为充分准备)
2	不满意	满意(国内具有显示度,研究扎实但不令人振奋,将增进理解)	满意(国内具有显示度,研究扎实但不令人振奋,增进了理解)	良好(开展较好研究,对社会作出较大贡献,为未来做好准备)
1	较差	不满意(研究既不扎实,又不令人振奋,不值得追求)	不满意(研究既不扎实,又不令人振奋)	不满意(研究结果并不理想,社会贡献并不令人满意,没有为未来做好准备)

资料来源:根据荷兰历次评价协议整理所得。

<sup>①</sup>FAIR原则详见:<https://www.go-fair.org/fair-principles/>。

<sup>②</sup>在荷兰,科学(wetenschap)一词不仅包括自然科学,也涵盖人文社会科学。

过允许学科评价口径自主选择、扩大学科评价指标遴选范围和调整学科评价结果呈现方式等举措,荷兰较好地平衡了内外部学科评价,有效激发了学科建设主体的自主性,为实现以评促建贡献了荷兰智慧,也对我国重塑学科评价秩序,构建具有中国特色、世界水平的学科评价体系颇具启发价值。

### 1. 评价口径与学科边界

评价口径是学科评价的关键参数。在学科边界相对清晰的小科学时代,自上而下确定学科评价口径或许有助于保证学科评价标准化。但当人类知识体系发展到今天,大量新兴、交叉学科竞相涌现,学科交叉随处可见,学科边界不再清晰可辨,自下而上确定学科评价口径已成大势所趋。随着中国《研究生教育学科专业目录(2022年)》中交叉学科门类与相应一级学科的设置,以及大量跨学科研究中心的建立,既有以一级学科为口径的评价体系与学科发展趋势之间的落差越来越大,保持学科评价口径的灵活性成为必须直面的现实挑战。如何确定学科评价口径,关键要看学科边界如何划分。学科知识与组织的二重性决定了学科边界亦有知识边界与组织边界之分。一方面,大量新兴、交叉学科的出现意味着学科的知识边界日趋模糊;另一方面,学科评价又要求相对清晰的组织边界,实然状态下模糊的知识边界与应然状态下清晰的组织边界之间存在巨大张力。如何既能顺应知识边界模糊的客观现状,同时又能满足组织边界清晰的现实要求?荷兰的经验是,采取自下而上的方式,由学科建设单位自主决定参评学科口径。但需要注意的是,这里的自主并不是无限的自主,而是有条件的自主,即参评学科至少满足下述条件:①战略条件,即参评学科必须要有明确的发展战略。②人才队伍条件。学科归根结底要靠人才来建,人才是学科建设过程中最具活力的要素。荷兰对参评学科的人才队伍规模提出了要求。除此之外,人才队伍结构也应纳入考量范围,参评学科必须形成由学科顶尖人才—学科领军人才—学科青年人才构成的人才梯队。③时间条件。只有组建时间超过一轮建设周期的学科方可参评。

### 2. 评价指标与建设维度

面对评价指标可能存在的“测不准”问题,荷兰采用固定评价维度但不固定评价指标的设计思

路,以相对稳定的评价维度来引领学科建设方向,对照评价维度学科建设主体便可明确建设重点,至于具体的评价指标则只提供指标清单,由学科建设主体自主选择适合的指标,不必唯指标是瞻,从而克服机械地缺什么补什么的非理性行动逻辑。通过SEP评价维度的变迁可以发现,科研质量、社会影响力、可行性构成荷兰学科评价相对稳定的评价维度,最新一轮学科评价则加大了对于博士生培养、开放科学、学术文化和人才政策维度的重视程度,这些变化趋势对于中国优化学科评价不无启发。我国第五轮学科评估主要考察各一级学科在人才培养质量、师资队伍与资源、科学研究(与艺术/设计实践)水平、社会服务与学科声誉四大维度的整体水平<sup>[22]</sup>。“双一流”建设成效评价重点关注各建设学科在人才培养、科学研究、社会服务、教师队伍建设四个维度的建设成效<sup>[23]</sup>。借鉴荷兰经验,一是要综合考虑学科知识前沿、国家战略需求与世界普遍趋势,设计具有中国特色、世界水平的学科评价维度,充分发挥对学科建设的引领功能。在保持既有评价维度基本稳定的基础上,探索新增开放科学、学术文化维度,将科研诚信纳入评价范围,并将师资队伍拓展至更广义的人才队伍,在学科评价中落实《科学数据管理办法》《关于优化学术环境的指导意见》《关于进一步加强科研诚信建设的若干意见》《关于深化项目评审、人才评价、机构评估改革的意见》等文件精神。二是可委托第三方机构或鼓励学科建设主体研发适用于不同学科的指标清单,将分类评价真正落到实处。分类评价不仅要依据学科特性设置不同的评价指标体系,因为即使为同一学科,不同发展定位的组织之间差别依旧巨大,所以根本仍在于以参评单元自身战略为参照系,引导其自主选择赛道,彰显学科特色。

### 3. 评价结果与卓越话语

量化评价结果催生了卓越修辞在全球学术界的蔓延。美国国立卫生研究所定期为卓越中心(Centres of Excellence)提供资助;加拿大建有卓越中心网络(Networks of Centres of Excellence),英国和澳大利亚的科研评估分别命名为“科研卓越框架”(Research Excellence Framework)和“澳大利亚卓越科研”(Excellence in Research for Australia),德国通过

“卓越计划”(Excellence Initiative)为“卓越集群”(Clusters of Excellence)提供资助。这套卓越话语规训学术系统日益走向高度竞争、相互模仿,继而导致组织同质化<sup>[20]</sup>,倘若失去自身特色,那么量化的卓越充其量也只是“失去灵魂的卓越”。在学科建设语境下,与卓越类似的话语还包括一流。三十年来荷兰逐渐弱化定量评价,在最新一轮评价协议中则完全取消了定量评价结果,要求外部专家只给出定性评价结果并提供相应的改进建议,以此规避模仿性机制下学科建设单位之间因向卓越看齐而盲目攀比。不过不攀比并不等于不比较,在自评环节学科建设主体需自主选择对标机构,锚定自身战略比较学科发展面临的优势与不足、机遇与挑战。荷兰的经验表明,打破卓越神话下的组织同构可以从转变评价结果形式与运用方式入手,一是弱化量化的打分、评级或分档,重视定性判断和学科建设建议,以量体裁衣的建设建议尊重学科发展规律,增强学科自组织能力。二是引导学科建设主体形成相对理性的对标意识,区分“建出来的卓越”与“评出来的卓越”。卓越可以作为学科建设的战略方向,但借助量化手段评出来的卓越不应成为奋力追赶的唯一目标,对标机构的选择一定是基于学术共同体对于“建出来的卓越”达成的共识,而非“评出来的卓越”制造的等级。理想的学科评价秩序应始于学科内部的对标与自评,继而由外部专家判断上一轮建设周期学科战略目标的达成度以及下一轮建设周期学科战略规划的可行性,最后将专家评价结果用于完善学科建设方案,指导新一轮对标与自评。

#### 参考文献

- [1] 张凤娟,宣勇.从“评建分离”到“以评促建”:有用的学科评价应实现四个转变[J].江苏高教,2022(7):37-42.
- [2] 朱冰莹,董维春.学科评价省思:场域特性、价值趋向与制度构建——兼议一流学科建设评价[J].科技进步与对策,2019,36(8):146-151.
- [3] 孟照海.制度化与去制度化:世界一流学科建设的内在张力——以美国芝加哥大学社会学为例[J].中国高教研究,2018(5):20-25.
- [4] 崔育宝,李金龙,张淑林.交叉学科建设:内涵论析、实施困境与推进策略[J].中国高教研究,2022(4):16-22.
- [5] MANVILLE C, ANGELO C D, CULORA A, et al. Understanding perceptions of the research excellence framework among UK researchers: the real-time REF review[R]. Cambridge, UK: Rand Europe, 2021: 19.
- [6] 蒋悟真.科研评价权的理论阐释及体系建构[J].法学,2022(12):149-162.
- [7] World Intellectual Property Organization. Global innovation index 2023: innovation in the face of uncertainty[R]. Geneva: WIPO, 2023: 163.
- [8] WOOLSTON C. Impact factor abandoned by Dutch university in hiring and promotion decisions[J]. Nature, 2021, 595: 462.
- [9] VUGHT F A V. The new government strategy for higher education in the Netherlands: an analysis[J]. Higher education quarterly, 1989, 43(4): 352-356.
- [10] STEEN J V, EIJFFINGER M. Evaluation practices of scientific research in the Netherlands[J]. Research evaluation, 1998, 7(2): 113-122.
- [11] 褚艾晶.从严格管制到大学自治——20世纪80年代荷兰政府与大学关系转型研究[J].外国教育研究,2016,43(4):119-128.
- [12] KICKERT W. Steering at a distance: a new paradigm of public governance in Dutch higher education[J]. Governance: an international journal of policy and administration, 1995, 8(1): 135-154.
- [13] GEUNA A, MARTIN B R. University research evaluation and funding: an international comparison[J]. Minerva, 2003, 41(4): 277-304.
- [14] VSNU KNAW NWO. Strategy evaluation protocol 2021-2027[R]. Netherlands: Hague, 2020: 29.
- [15] DROOGE L V, JONG S D, FABER M, et al. Twenty years of research evaluation[R]. Hague: Rathenau Instituut, 2013: 4.
- [16] MAASSEN P A M, POTMAN H P. Strategic decision making in higher education: an analysis of the new planning system in Dutch higher education[J]. Higher education, 1990, 20(4): 393-410.
- [17] Commissie Nationaal Plan Toekomst Geesteswetenschappen. Duurzame geesteswetenschappen[R]. Amsterdam: Amsterdam University Press, 2008: 35.
- [18] 刘凡丰,李晓强.荷兰标准评价协议的指标变革及其启示[J].中国高校科技,2019(11):45-48.
- [19] WILSDON J, ALLEN L, BELFIORE E, et al. The metric tide: report of the independent review of the role of Metrics in research assessment and management[R]. London: HEFCE, 2015: 58.
- [20] MOORE S, NEYLON C, PAUL E M, et al. “Excellence R Us”: university research and the fetishisation of excellence[J]. Palgrave communication, 2017(3): 16105.
- [21] RIP A, MEULEN B J R V D. The patchwork of the Dutch evaluation system[J]. Research Evaluation, 1995, 5(1): 45-53.
- [22] 中华人民共和国教育部.第五轮学科评估指标体系框架[EB/OL].(2020-11-03)[2022-08-02].[http://www.moe.gov.cn/jyb\\_xwfb/moe\\_1946/tj\\_2020/202011/t20201102\\_497819.html](http://www.moe.gov.cn/jyb_xwfb/moe_1946/tj_2020/202011/t20201102_497819.html).
- [23] 教育部,财政部,国家发展改革委.关于印发《“双一流”建设成效评价办法(试行)》的通知:教研[2020]13号[A/OL].(2020-11-03)[2022-08-02].[http://www.moe.gov.cn/srcsite/A22/moe\\_843/202103/t20210323\\_521951.html](http://www.moe.gov.cn/srcsite/A22/moe_843/202103/t20210323_521951.html).

(责任编辑 黄欢)