

# 科研绩效拨款制度中的 国家科研评价体系研究

蒋 颖

---

**【摘 要】** 科研绩效拨款制度是国家科研评价体系发展的重要驱动力。国家科研评价体系辐射面广、影响力大，在其建设过程中应树立系统观念，对评价体系的基本要素进行精心设计，同时充分认识评价环境和元评价对评价体系建设及评价效果的重要影响，并最终将科学的设计思想反映在评价政策中。目前尚不存在国家科研评价体系“最优”模式，但30余年的国际实践提供了丰富的建设经验，也提出了很多问题和挑战，可为完善我国科研评价体系提供参考和借鉴。

**【关键词】** 国家科研评价体系 科研绩效拨款制度 评价政策 评价环境 元评价

**【作者简介】** 蒋颖，中国社会科学院中国社会科学评价研究院副院长、研究馆员。

**【中图分类号】** G311      **【文献标识码】** A

**【文章编号】** 2097 - 1125 (2024) 12 - 0005 - 18

---

近年来，我国学术（科研）评价体系建设得到高度重视。习近平总书记指出，我国的“学术评价体系不够科学，管理体制和运行机制还不完善”，“要建立科学权威、公开透明的哲学社会科学成果评价体系”。<sup>①</sup> 2021年5月，中央全面深化改革委员会第十九次会议审议通过了《关于完善科技成果评价机制的指导意见》。建立和完善中国特色学术评价体系和评价机制，坚持正确的科研成果评价导向，创新评价方式，通过评价激发科研人员积极性，推动产出高质量成果，营造良好学术生态，是当前阶段的重要任务。

---

<sup>①</sup> 习近平：《在哲学社会科学工作座谈会上的讲话（2016年5月17日）》，《人民日报》2016年5月19日。

科研绩效拨款制度是国家科研评价体系与拨款制度的结合。推行该制度在很大程度上促进了国家科研评价体系的建设和完善。20世纪80年代以来,多个国家实施科研绩效拨款制度,积累了丰富的国家科研评价体系建设经验,同时也遇到了很多问题和挑战。国家科研评价体系建设极具挑战性,对评价体系的科学性、系统性有很高要求,评价结果辐射面广、影响力大。因此,系统地梳理和总结国家科研评价体系建设国际经验,有助于加深对国家科研评价体系建设规律的认识,可以为完善我国的学术评价体系提供参考和借鉴。

## 一、科研绩效拨款制度与国家科研评价体系

### (一) 科研绩效拨款制度

在传统上,现代公立大学和国立研究机构大多通过中央政府提供的“一揽子拨款”方式获得资助,同时也通过项目方式从各类国家或私营资助机构获得竞争性科研资助。20世纪80年代,随着新公共管理理论的发展,各国政府为提高公共部门效率,强调对公共部门进行明确的绩效评价,受公共经费支持的科学研究事业也逐渐被纳入绩效评价的范围。与此同时,全球科技竞争不断加剧,科学研究与创新对国家经济发展的重要性日益凸显,追求卓越科研成为国家的重要战略目标。在此背景下,与科研相关的机构整体资金的分配机制开始发生变化,从主要以历史数据和教育科研规模为依据进行经费分配转向部分或全部按科研绩效指标进行经费分配。

1986年,英国政府率先开展了针对大学的科研绩效评价,并以评价结果为依据对大学进行“一揽子拨款”。在随后的30余年中,英国政府不断改进评价方法,先后开展了8轮科研绩效评价。<sup>①</sup>受其影响,很多国家(特别是欧盟成员国)陆续开展了服务于科研绩效拨款的评价活动,在此基础上逐渐形成了科研绩效拨款制度(Performance-based Research Funding System, PRFS)。科研绩效拨款是来自公共资金并面向组织层面的一种竞争性(或选择性)资助,在全国或区域范围内实施,拨款依据主要是对科研成果的数量、质量和(或)影响的事后评价结果。<sup>②</sup>因此,项目资助、基于教学质量

<sup>①</sup> 参见 Building on Success and Learning from Experience, <https://assets.publishing.service.gov.uk/media/5a803df4e5274a2e8ab4f03d/ind-16-9-ref-stern-review.pdf>, 2024年10月10日。

<sup>②</sup> 参见 Thomas Zacharewicz, Benedetto Lepori and Emanuela Reale et al., Performance-based Research Funding in EU Member States—A Comparative Assessment, *Science and Public Policy*, Vol. 46 (1), 2019, pp. 105-115。

评价的拨款、基于机构内部评价的拨款及基于科研人员个人绩效评估的拨款等都不属于科研绩效拨款的范畴。设立科研绩效拨款制度通常有两方面的动机，一方面是在全球化和知识经济的背景下促进卓越科研和创新发展，提高国家竞争力；另一方面是揭示和提高公共经费的使用效率，落实对大学和研究机构的问责制，满足服务国家战略发展的需求。

自英国实行科研绩效拨款制度后，很多国家也开始采用类似的拨款方式。截至 2015 年，在欧盟 28 个成员国中，有 18 个国家采用了这种拨款方式。还有一些国家虽未正式实行科研绩效拨款制度，但在科研经费分配中引入了一些基于绩效的要素。<sup>①</sup> 欧洲之外的一些国家，如澳大利亚、新西兰等也实行了科研绩效拨款制度。在欧洲实行科研绩效拨款制度的国家较多，这与欧盟积极的推进政策密切相关。一些国际组织也关注科研绩效拨款制度，如 OECD 曾经举办会议对科研绩效拨款制度进行研讨。<sup>②</sup>

## （二）国家科研评价体系及其要素构成

科研评价是科研绩效拨款制度的基础和核心。一方面，科研评价是科研绩效拨款制度的基础，为拨款工作提供依据，作为该制度中最受关注的部分，科研评价是保证科研绩效拨款制度实施效果的关键。另一方面，科研绩效拨款制度涉及全国性经费分配，对科研评价体系有很高要求，同时也为科研评价体系建设提供良好支持和保障，极大地促进了科研评价体系的发展和完善。一些国家在推行科研绩效拨款制度过程中采用的评价体系，逐渐成为其国家科研评价体系的主体。本文将这类由国家主导并面向组织层面的全国性科研评价体系称为“国家科研评价体系”。

国家科研评价体系最主要的特点是由政府推动并通常与拨款相关，评价效力强大。它围绕国家政策目标，服务国家发展战略，具有高度的合法性和权威性，评价实施有一定的强制性，评价范围具有全国性，评价制度作为一种制度长期存在，评价活动按照一定的周期定期开展，评价结果往往被作为国家科研绩效拨款的重要依据。国家科研评价体系具有高度的引领作用，是国家科研活动的重要“指挥棒”。这意味着相较机构层面的科研评价体系，政府、学术界和社会对国家科研评价体系的系统性和科学性有更高的要求，一旦其评价体系不够完善，带来的负面影响就更大。

科研评价体系建设是近年来的热门话题。科研评价体系的基本要素包括

① 参见 Thomas Zacharewicz, Benedetto Lepori and Emanuela Reale et al., Performance-based Research Funding in EU Member States—A Comparative Assessment, *Science and Public Policy*, Vol. 46 (1), 2019, pp. 105 - 115.

② 参见 OECD, *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, Paris: OECD Publishing, 2010, pp. 1 - 187.

评价目标、评价主体、评价对象（客体）、评价标准及指标、评价方法。在此基础上，良好的评价体系需要一系列明确的规则，包括评价的基本原则和评价组织方法，由评价主体制定并被作为学术评价活动的行动依据和准则。叶继元提出，学术评价体系由评价主体、评价客体、评价目的、评价方法、评价标准及指标、评价制度组成。<sup>①</sup>从定义来看，制度是“人类共同遵守的人为设定的办事规章或行动准则”。<sup>②</sup>诺斯把制度定义为一个社会的“博弈规则”，并将之分为正式规则和非正式规则。<sup>③</sup>正式规则通常以明文形式存在，由政府或相关权威机构制定和实施；非正式规则包括思想信仰、意识形态、习惯行为、社会习俗、传统文化、伦理道德规范等，没有明确的书面形式，在短时间内不易改变。而政策是“党政机关、企事业单位和各种社会组织为达成一定目标、结合特定情况所制定的行动准则和方案”。<sup>④</sup>因此，评价政策就是评价主体为达成评价目标、结合特定情况制定的评价活动的行动准则和方案，也就是评价制度中的正式规则。在评价制度中也存在非正式规则，如评价文化。为清晰地分析评价体系的构成，本文将二者区分开来，将属于正式规则的评价政策作为评价体系的一个重要组成部分。

现有评价体系的研究和设计大多只考虑评价体系的内部因素，但对国家科研评价体系而言，影响其评价效果的不仅是评价体系本身的设计和实施，而且包括一些外部因素，如评价体系所处的社会、经济、文化和科研环境，此外，评价文化也会在空中通过评价主体、评价客体和其他利益相关者对评价活动的效果产生正面或负面影响。这解释了不同的国家采用同样的评价方法和相似的评价政策何以有时会产生截然不同的评价效果。因此，评价环境也应作为外部条件纳入国家科研评价体系中，成为其重要的组成部分。

好的评价体系不是僵化的，而是可以不断改进和完善的有机体。评价活动是评价体系实现的方式，一个连续运转的国家科研评价体系通常会持续开展多轮评价活动。为保证评价体系的有效性，在每一轮评价活动之后都应开展元评价——对已发生的评价活动的再评价，通过及时进行再评价和总结，肯定有效做法，发现问题和短板，提出对评价活动的具体改进意见。因此，可将元评价作为评价体系的一部分，以保证评价体系具有自反性。

总之，国家科研评价体系是一个复杂而有巨大影响的系统，不仅包括一般评价活动涉及的内部因素，而且必须考量相关的外部环境及持续发展机

① 参见叶继元：《学术“全评价”体系论》，社会科学文献出版社2021年版，第107~118页。

② 经济学名词审定委员会编：《经济学名词》，科学出版社2020年版，第57页。

③ 参见[美]德勒巴克主编：《新制度经济学前沿》，张宇燕译，经济科学出版社2003年版，第14页。

④ 管理科学技术名词审定委员会编：《管理科学技术名词》，科学出版社2016年版，第521页。

制。因此，笔者认为一个完善的国家科研评价体系应由评价体系的基本要素、评价政策、评价环境和元评价等方面组成。其中，评价体系的基本要素包括评价目标、评价主体、评价对象、评价标准及指标、评价方法 5 个方面，构成了评价体系的框架；评价政策为评价活动提供了行动准则和具体方案，是决定评价体系质量的关键因素；评价环境对评价活动形成软约束；元评价保证评价体系具有自反性，以不断自我优化和改进。

## 二、国家科研评价体系的要素分析

在科研绩效拨款制度下，很多国家基于本国国情开展了内容丰富、各具特色的科研评价活动。本文基于前文提出的国家科研评价体系的框架，对各国的国家科研评价体系进行梳理，期望总结国家科研评价体系的共性特征及其他国家的经验教训，为完善我国科研评价体系提供参考和借鉴。

### （一）评价政策

国家科研评价政策在总体上有以下几个方面的特点。

第一，指导性。作为国家层面的政策工具，在科研绩效拨款制度下的国家科研评价政策体现国家意志，服务国家战略，引导科研机构的战略行为。这表现在评价政策并非仅考量学术价值和社会影响，而是将国家需求和国家科学发展规划融入评价政策之中并放在重要位置。例如，很多国家都将促进卓越科研和创新发展作为国家科研评价的目的，并设立相关的评价标准，对高质量成果赋以较大的指标权重。

第二，系统性。在科研绩效拨款制度下的国家科研评价政策必须是一个可以落地实施的完整体系。它既包括宏观的法律、法规，又有详细的操作方法及评价规范，从评价的意义和目的开始，贯穿评价活动的设计、组织、实施及效果评价的全过程，最终所有环节都要落实到操作细节上，可直接指导评价实践。英国、挪威、芬兰等国都制定了全面、完善的国家科研评价政策体系，并将之在网站上公开发布。

第三，周期性。在科研绩效拨款制度下的国家科研评价活动具有明确的周期，相关的政策规定也可以进行周期性调整和改变。国家科研评价活动的实施周期因评价方法的不同而有较大差异，在个人或机构层面收集信息并进行同行评议的国家通常需要 4 年至 5 年甚至更长的评价周期，而那些利用公开数据的量化评价体系的评价周期可短至 1 年。在新一轮国家科研评价开始前通常会轮对上一轮评价进行充分评估以发现其中的问题，并在即将开展的评价活动中对评价政策进行调整和优化。国家科研评价体系通常有自己的名称，如果其内容发生重大改变，就会被重新命名。例如，英国的国家科研评价体系就经历过“科研选择性评估”（Research Selective Exercise, RSE）、“科研水平评估”

(Research Assessment Exercise, RAE) 和“科研卓越框架”(Research Excellence Framework, REF) 三个阶段,不同阶段中的评价方法、评价标准和评价指标都有实质性变化。

第四,权威性。在科研绩效拨款制度下的国家科研评价政策引导性强、影响力大、社会关注度高,在国家总体科研评价体系中具有引领作用。一方面,国家科研评价政策可以推进国家科研评价活动顺利进行。科研绩效拨款制度对国家科研评价政策的形成具有驱动和保障作用,既是推进科研评价的重要动力,也是科研所需经费、人力等条件的有力保障。例如,挪威等国在科研绩效拨款制度的推动下建设了高质量的国家当前科研信息系统(Current Research Information Systems, CRIS),为量化评价提供可靠的数据。另一方面,相较机构层面的评价政策,国家科研评价政策具有更大的影响力,“指挥棒”效应更加明显,这也意味着如果出现政策失误,其负面影响可能被放大,后果将更为严重。因此,国家科研评价政策的制定必须经过科学论证和仔细考量。

## (二) 评价体系的基本要素

### 1. 评价目标

评价目标决定评价结果的应用范围,直接影响到评价标准和评价方法的选择。在科研绩效拨款制度中评价的主要目标非常清晰明确,就是通过评价国家公共资金资助的科研系统的整体科研质量,为科研绩效拨款提供直接依据。挪威等一些国家的国家科研评价体系只有上述这一个评价目标。另一些国家的国家科研评价体系除了为科研绩效拨款提供依据,还有其他目标,如英国 REF 的主要目标有三个:为分配科研资助提供依据;为对科研公共投入问责提供投资效益依据;为国家和社会提供高等院校的声誉评价信息。<sup>①</sup>多重目标使评价政策的影响范围更大,同时也使评价方法更为复杂。

### 2. 评价主体

评价主体可分为评价活动的组织者和评价者两个层面。由于科研绩效拨款包含对公共经费的考核和分配,故而其评价活动的组织者通常为政府或其下属的相关部门,其中主要是教育部门,也有受国家委托的相关基金会或第三方机构。科研绩效拨款的评价者通常由学术共同体成员组成,也可以分为多个层次。很多国家专门设立了科研评价委员会,并随着评价目的和职责的变化对机构进行调整或更新。例如,意大利在1996年设立大学体系评估监察组(Osservatorio),1999年将其更名为“全国大学评价委员会”(CNVSU),2003年设置研究评估指导委员会(CIVR),2008年成立高校及研究机构国

<sup>①</sup> 参见刘彦林:《英国社会科学评价模式——以科研卓越框架(REF)为例》,内部研究报告,2023年。

家评估署（ANVUR）以取代 CIVR 和 CNVSU，并进一步明确了评价的目的和范围。<sup>①</sup> 科研评价委员会通常设置专家指导组和学科专家组，专家指导组负责协调和指导学科专家组的工作，学科专家组完成各学科的具体评价任务。例如，芬兰的科研评价委员会设置了 1 个指导小组和 23 个专家小组；<sup>②</sup> 英国 REF 2021 设置了 4 个主要小组和 34 个子小组，其中有 900 名科研成员和 220 名研究用户直接参与专家评价工作。<sup>③</sup>

### 3. 评价对象

评价对象可以包括大学（研究机构）、院系、研究小组、个人等不同层级，多数科研绩效拨款制度中的评价对象处于中观层级，即大学或院系的层级。在挪威、芬兰、丹麦、比利时佛兰德斯地区等国家和地区的科研绩效拨款制度中的评价对象均为公立大学或公立研究机构。一些国家的科研绩效拨款制度以大学中的院系或专业为评价对象，如英国以大学的学科单元为评价对象，意大利、葡萄牙与英国类似。波兰在 2018 年《科学和高等教育法案》颁布之后，其国家科研评价体系的评价对象从大学的院系、研究所变为大学或研究所内的学科。虽然国家科研评价涉及各个机构或学科的成员及其成果，但是最终的评价结果仅在机构或学科等组织层面进行比较，相关数据通常不作为个人晋升或考核评价的依据。只有很少国家同时面向个人开展科研评价，如西班牙国家科研评价体系 Sexenio 的评价结果与个人工资挂钩。<sup>④</sup>

### 4. 评价标准及指标

大部分国家的国家科研评价标准都包括科研产出（科研成果的数量和质量）、科研环境（外部课题经费、人才培养、科研合作、国际化、多样性等）、社会贡献等内容。其中，科研产出是评价的重点。近年来，为突出科研成果质量的重要性，多国在科研成果的影响力评价方面开展了深度探索。例如，英国 REF 采用提交案例的方式开展科研成果的影响力评价，将科研成果的影响分为对社会福利、商业和经济、生产的影响等 9 大类，并在评价文件中提供了判断影响力范围和重要性的参考指标。<sup>⑤</sup>

评价指标是评价标准的具体化。面向不同评价方法的指标设计差异较

① 参见耿海英：《意大利科研质量评价体系》，内部研究报告，2024 年。

② 参见 Elina Pylvänäinen and Janne Pölönen, *Publication Forum Review of Ratings in 2022*, Helsinki: The Federation of Finnish Learned Societies, 2023, p. 4.

③ 参见 REF 2021, Key Facts, [https://2021.ref.ac.uk/media/1848/ref2021\\_Key\\_facts.pdf](https://2021.ref.ac.uk/media/1848/ref2021_Key_facts.pdf), 2024 年 10 月 10 日。

④ 参见 Diana Hicks, *Performance-based University Research Funding Systems*, *Research Policy*, Vol. 41 (2), 2012, pp. 251 - 261.

⑤ 参见刘彦林：《英国社会科学评价模式——以科研卓越框架（REF）为例》，内部研究报告，2023 年。

大,如科研产出的量化指标数量有限,通常以期刊论文、图书及其章节、会议论文、专利等成果的发表或被引指标为主,而同行评议方式下的产出指标范围可以更加广泛,如英国的科研成果指标除了常见类型的指标,还包括研究报告、展览和表演、艺术作品、数字人工制品等多种指标。

### 5. 评价方法

评价方法是评价的关键。当前主要的科研评价方法可分为同行评议和以文献计量学为主的量化评价两种类型。大量的理论研究和实证分析对定量和定性评价方法的优劣进行了比较,并对定量方法的使用条件和限制达成共识,形成《科研定量评价莱顿宣言》。<sup>①</sup>与此同时,学界普遍认识到不同层面的评价对象适用的评价方法不同:评价对象越偏向宏观层级,量化评价方法的可靠性越高;评价对象越偏向微观层级,同行评议方法的可靠性越高。很多研究发现,对中观层面的评价对象(如大学及其院系和学科)的量化评价和同行评议结果一致性很高,但前者在评价效率和成本节省方面显著优于后者。<sup>②</sup>因此,采用量化评价方法的国家科研评价体系大多基于期刊社或出版社等出版渠道的出版物进行评价。在实行科研绩效拨款制度的国家中,英国、法国、意大利、立陶宛、葡萄牙、新西兰等国采用以同行评议为主的评价方法,比利时、捷克、丹麦、爱沙尼亚、芬兰、克罗地亚、挪威、波兰、瑞典、斯洛伐克等国采用以量化评价为主的评价方法。

各国在科研评价实践中对评价方法不断改进,在采用量化方法的评价体系中吸收了同行评议的一些做法,对部分学科在采用同行评议方法时也参考了量化指标。例如,一些国家学习英国的做法,限制每名研究者可提交的成果数量,实施类似代表作提交制度的政策,以此控制片面追求成果数量的趋势;英国在进行科研评价时为部分学科提供文献计量学数据供同行评议参考,形成知情同行评议的模式。同行评议方法也被拓展。例如,英国 REF 中评审专家的范围已由传统的学术专家扩展至更广泛的用户和利益相关者等群体,在 REF 2014 中外部专家占比已达到 23%。<sup>③</sup>这意味着英国 REF 使用的同行评议已逐渐演变为更大范围的专家评议。

各国科研绩效拨款制度采用的评价方法并不是固定的,实际上很多国家都探索过不同的科研评价路径,试图发现最适合本国的评价方法并建立评价

① 参见 Diana Hicks, Paul Wouters and Ludo Waltman et al., *Bibliometrics: The Leiden Manifesto for Research Metrics*, *Nature*, Vol. 520, 2015, pp. 429 - 431.

② 参见张琳、Gunnar Sivertsen:《科学计量与同行评议相结合的科研评价——国际经验与启示》,《情报学报》2020年第8期,第806~816页。

③ 参见茹宁、闫广芬:《非学术影响评价:英国 REF 科研影响评估的创新性评析》,《国家教育行政学院学报》2020年第9期,第24~34页。

体系。澳大利亚科研绩效拨款制度就经历了量化评价—同行评议—量化评价的演变过程，在这个过程中不断根据评价需求和评价目的的变化调整评价方法、完善评价过程。1995年，澳大利亚确立基于量化指标的科研绩效拨款方式，造成片面追求科研成果数量而忽视科研成果质量的问题。为解决这一问题，澳大利亚政府将科研绩效拨款建立在澳大利亚卓越科研评价体系（Excellence in Research for Australia, ERA）的基础上，采取同行评议的方式进行科研评价，激励大学科研人员更加注重科研成果的质量。随着科研成果质量的提升，大学科研的社会贡献问题成为新的焦点。2017年，澳大利亚政府又适时调整和精简科研绩效拨款制度，不再将其与学科评价结果挂钩，而是采用量化公式拨款的方式，将激励的重心转向大学与产业互动领域。<sup>①</sup>

### （三）评价环境

评价环境包含诸多因素，不同国家的评价环境有很大差异，其对国家科研评价体系的影响程度也不同。对国家科研评价体系直接产生影响的评价环境因素，包括评价激励的强度、科研环境与评价文化、相关的教育评价制度以及与利益相关者的协商。

#### 1. 评价激励的强度

基于评价结果对评价对象进行的激励，如绩效拨款或其他针对机构及个人的物质奖励，反过来也会影响评价效果，甚至会给评价体系的长期发展带来影响。一方面，适度的激励会激发科研积极性和创造力，但需要注意的是，激励强度越大，评价体系中的不足也越容易被放大，进而越容易影响到评价的有效性和合理性。例如，从科研绩效拨款占整体科研经费的比例来看，各国科研评价激励强度的差异很大，挪威的这一比例为2%，英国的这一比例为25%，比利时佛兰德斯地区的两个主要基金的这一比例分别为50%和75%，而捷克曾一度完全采用绩效拨款方式分配科研经费。<sup>②</sup>相应地，挪威、英国、比利时的国家科研评价体系较为完善，运行平稳，而捷克采用的完全基于量化指标的国家科研评价体系则导致科研经费分配不合理，引发了学术界的强烈批评。<sup>③</sup>另一方面，激励强度越大，机构将绩效最大化的动机越强烈，越容易引发博弈行为，进而会扭曲评价政策的意图，导致短期主义、机会主义、个人寻租甚至学术不端现象，使科研评价偏

① 参见张强、周志刚：《有限绩效与多维评价：国外科研绩效拨款机制及其实践——以英国、澳大利亚、新西兰为例》，《外国教育研究》2022年第6期，第111~128页。

② 参见 Marc Luwel, *Performance-based Institutional Research Funding in Flanders, Belgium, Scholarly Assessment Reports*, Vol. 3 (1), 2021, pp. 1-24.

③ 参见 Barbara Good, Niki Vermeulen and Brigitte Tiefenthaler et al., *Counting Quality? The Czech Performance-based Research Funding System*, *Research Evaluation*, Vol. 24 (2), 2015, pp. 91-105.

离激发科研积极性和创造力的初衷。

## 2. 科研环境与评价文化

科研环境会直接影响科研评价方法的选择。如果国家学术共同体规模不够大,难以找到足够数量的满足回避规定的“小同行”,同行评议的实施就有难度。科研成果的发表语言也会影响科研评价指标和方法的选择。因为国际引文数据库对英文期刊收录较全面,所以英语国家在进行定量评价时可以直接使用国际引文数据库的指标,但非英语国家则需要特别注意国际引文数据库是否充分收录使用本国语言发表的科研成果,以及本国语言成果与国际语言成果之间的权重比例关系等问题。例如,波兰的国家科研评价体系曾在收录波兰语成果方面经历了从排斥到过度收录等阶段,最后确定了择优收录的原则,对波兰语期刊的评价采用了同行评议的方法,而不是英文期刊评价的量化方法。<sup>①</sup>

评价文化对评价方法的选择及评价效果也有较大影响。是否有良好的同行评议传统和较为廉洁公正的评价文化,以及对文献计量学指标的接受或排斥程度等因素,都会影响评价方法的选择和评价结果的科学性与有效性。例如,丹麦基于期刊的BFI评价体系受到学者的抵制。这是因为丹麦自然科学和医学领域的学者已经习惯于采用其他文献计量指标,如被引频次和H指数等测度方式,所以他们认为BFI缺乏新意,甚至是一种倒退;而人文学科的学者认为研究的“计量化”与学术工作的性质格格不入。<sup>②</sup>这种排斥心理是BFI最终停止使用的原因之一。

## 3. 相关的教育评价制度

从国家科研评价体系与国家科研教育事业全面协调发展的角度看,科研绩效评价的效果还与国家其他相关制度的建设有关,要充分考虑制度之间的统筹协调。大部分国家科研机构的主体是大学,而除了科学研究,大学的另外两个主要功能是教育和社会服务。承担这些功能的主体都是教师,因此对科研的激励可能会影响教育和社会服务,要保持这些功能之间的平衡就必须关注科研评价政策与另外两方面评价政策之间的统筹和协调。在对大学进行科研评价的同时,也要考虑其与教育评价政策及社会服务评价政策的呼应,以保证大学的全面健康发展。英国在大学评价方面进行了卓有成效的探索,发展出三个“卓越”系列的大学评价框架,包括对科学研究进行评价的“科研卓越框架”(REF)、对教育进行评价的“教学卓越框架”(Teaching Excellence Framework, TEF)以及作为评判大学社会服务活动新手段的“知识交换框架”(Knowledge Exchange Framework, KEF),三者共同构成一个完整的大学外部评价体系。

① 参见蒋颖:《波兰绩效评价资助体系中的学术评价》,内部研究报告,2024年。

② 参见 Poul Erik Mouritzen and Niels Opstrup, *Performance Management at Universities: The Danish Bibliometric Research Indicator at Work*, London: Palgrave Macmillan, 2020, pp. 37-39.

#### 4. 与利益相关者的协商

评价活动利益相关者的理解和支持是评价活动顺利开展的必要条件。国家科研评价活动的利益相关者包括作为评价对象的大学、研究机构、科研团队和科研人员，以及科研成果的用户和服务对象等群体。良好的评价体系应包含与利益相关者广泛协商的过程。与利益相关者进行充分沟通和广泛协商，收集大学和研究机构的意见，就评价目标和评价标准达成一致，这些环节对顺利推进科研评价活动和提升科研评价效果是非常有必要的。一些国家早期的科研评价活动缺乏广泛协商的环节，如澳大利亚综合指数 CI，它的评价方法和评价结果因此受到广泛的批评。在后续的 ERA 体系设计过程中，澳大利亚研究委员会与大学中的相关人员进行深度交流，发现后者对程序的公平性和透明度有强烈的要求。最终，ERA 方案被学术界接受。<sup>①</sup>这也意味着评价体系保持开放和畅通信息沟通渠道十分重要。

#### （四）元评价

元评价是按照一定的理论框架和评价标准对评价体系本身进行的再评价。该概念由 M. 斯克里文在 1969 年首次提出，在教育评价中有较多应用，后来逐渐扩散到科研评价领域。在科研评价领域开展元评价可以保证评价效果、完善评价机制、提高评价质量。元评价有自我评价、专家评价、第三方评价等多种形式，有些国家还邀请国外专家或相关机构进行元评价，以便保持更加公正客观的立场。元评价的内容通常包括对评价体系的各要素、评价过程、评价结果和存在问题的分析，以及对评价的改进意见等。英国的国家科研评价在元评价方面做得非常出色，在每一轮评价活动结束后，都会开展一系列的评估活动，并公开发布相关报告。例如，REF 2014 评估报告包括管理人员、评价专家、被评价机构、第三方机构（如兰德公司）等具有不同身份的人员或机构撰写的评价报告或反馈意见，并针对评价成本、收益及跨学科研究情况等特殊问题开展专项评估。英国国家科研评价工作的稳步推进与其对评价活动的全面总结和对评价政策的及时调整改进有密切关系。挪威也曾经邀请丹麦奥胡斯大学的学者代表挪威高等教育机构协会对被成为“挪威模式”的科研绩效拨款评价体系的实施效果进行评估，并及时针对在评估中发现的问题进行改进。<sup>②</sup>

① 参见 Ellen Hazelkorn, *Assessing Europe's University-based Research*, Luxembourg: Publications Office of the European Union, 2010, <https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1017&context=cserrep>, 2024 年 10 月 10 日。

② 参见 Kaare Aagaard, Carter Bloch and Jesper W. Schneider, *Impacts of Performance-based Research Funding Systems: The Case of the Norwegian Publication Indicator*, *Research Evaluation*, Vol. 24 (2), 2015, pp. 106 - 117。

### 三、科研绩效拨款制度下国家科研评价体系建设的效果及挑战

科研绩效拨款制度下的国家科研评价体系对国家的科学研究和科研生态具有广泛而深刻的影响。因为涉及经费分配和机构声誉，所以科研绩效拨款制度下的国家科研评价影响力大，受到大学、研究机构等评价对象的高度重视，“指挥棒”效应显著。然而，如果评价体系不合理，其造成的负面影响也不容忽视。国家科研评价体系的构建效果集中体现在科研绩效拨款制度的效果中。后者是各国普遍关心的重要问题，欧盟及部分欧盟之外的国家对已实施的科研绩效拨款活动组织了评估和检视，一些学者也通过问卷调查或文献计量的方式对此进行了分析。鉴于数据获取困难、各国情况差异较大、影响因素众多、实施效果难以测度等原因，对科研绩效拨款制度进行总体评价较为困难。即便如此，从既有研究中仍然可以看到，服务于科研绩效拨款制度的国家科研评价体系在总体上得到各国官方的肯定，同时也存在一些问题和挑战。

#### （一）科研绩效拨款制度的总体效果

相关研究显示，大部分国家的科研绩效拨款制度带来的不利影响在可接受范围内，作为一个政策工具，该制度总体而言是有益的。

2016年，欧盟发布研究评估报告，对欧盟成员国及部分其他国家实行科研绩效拨款制度的效果进行分析比较。结果发现，就该报告审议的各项指标而言，已实行的科研绩效拨款制度未产生强烈的负面影响，该制度的效果总体而言是积极的。<sup>①</sup>在欧盟科研绩效拨款制度相互学习活动<sup>②</sup>中的12个实行科研绩效拨款制度的国家都认为它们从该制度中受益，表现为科研质量和生产力提升，大学更积极地采取以质量为导向的政策，政策透明度有所提高，博士生和年轻科研人员的发展前景得到改善。<sup>③</sup>

① 参见 Koen Jonkers and Thomas Zacharewicz, *Research Performance Based Funding Systems: A Comparative Assessment*, Luxembourg: Publications Office of the European Union, 2016, <https://core.ac.uk/download/pdf/38632345.pdf>, 2024年10月10日。

② 相互学习活动 (Mutual Learning Exercise, MLE) 是欧盟“地平线 2020”计划支持的项目，旨在关注会员国和联合国感兴趣的具体研究与创新挑战，促进相关项目的良好实践交流。

③ 参见 European Commission, *Performance-based Funding of University Research: Summary Report*, Brussels: European Commission, 2018, [https://projects.research-and-innovation.ec.europa.eu/sites/default/files/rio/report/MLE%2520PRFS\\_Summary%2520Report\\_0.pdf](https://projects.research-and-innovation.ec.europa.eu/sites/default/files/rio/report/MLE%2520PRFS_Summary%2520Report_0.pdf), 2024年10月10日。

Checchi 等对 31 个国家 20 年间的数据进行分析，认为科研绩效拨款是一种非常有用的工具，可以长期指导大学系统并提高国家科学研究的整体影响力。<sup>①</sup> 英国对 REF 的评估结果表明，科研评价活动增强了科研人员对研究的严谨性、新颖性和重要性的认识及对国际合作重要性的认识，有助于在学科层面制定研究规范，引发对科研质量和卓越性的关注，提高学术界对影响力问题的重视程度，同时，评价带来的声誉为大学吸引了学生、员工和外部资金。<sup>②</sup>

## （二）科研绩效拨款制度对学术产出量与质的作用

在科研绩效拨款制度实行期间，各国的学术产出普遍增加。例如，挪威于 2003 年开始实施被称为“挪威模式”的科研绩效拨款评价活动，从 2004 年到 2012 年，人均论文发表数量增长了 26%。<sup>③</sup> 波兰在实行科研绩效拨款制度后，在保持科研质量的同时，论文的数量得到提升。<sup>④</sup> 该制度在英国、澳大利亚和西班牙实行后，都带来论文产出的增加。<sup>⑤</sup>

在科研绩效拨款制度实行期间，各国科研质量的提升情况不尽相同，一些国家的科研质量得到明显提升。英国对 REF 的评估结果显示，30 多年来，英国的 RAE 和 REF 对英国科研质量和生产力的持续提升提供了支持，英国在被引用次数排名前 1% 论文中所占的份额从 1996 年的 11% 增加到 2012 年的 16%，这在一定程度上可归因于科研绩效拨款资金的驱动和支持。<sup>⑥</sup> 但也有些国家在科研成果数量增加的同时科研质量出现了下降。澳大利亚在 20 世纪 90 年代采用的量化评价体系中没有区分出版物质量，结果论文发表

① 参见 Daniele Checchi, Marco Malgarini and Scipione Sarlo, Do Performance-based Research Funding Systems Affect Research Production and Impact?, *Higher Education Quarterly*, Vol. 73 (1), 2019, pp. 45 - 69。

② 参见 Building on Success and Learning from Experience, <https://assets.publishing.service.gov.uk/media/5a803df4e5274a2e8ab4f03d/ind-16-9-ref-stern-review.pdf>, 2024 年 10 月 10 日。

③ 参见 Kaare Aagaard, Carter Bloch and Jesper W. Schneider, Impacts of Performance-based Research Funding Systems: The Case of the Norwegian Publication Indicator, *Research Evaluation*, Vol. 24 (2), 2015, pp. 106 - 117。

④ 参见 Luwel Marc, Performance-based Institutional Research Funding in Flanders, Belgium, *Scholarly Assessment Reports*, Vol. 3 (1), 2021, pp. 1 - 24。

⑤ 参见 Kaare Aagaard, Carter Bloch and Jesper W. Schneider, Impacts of Performance-based Research Funding Systems: The Case of the Norwegian Publication Indicator, *Research Evaluation*, Vol. 24 (2), 2015, pp. 106 - 117。

⑥ 参见 Building on Success and Learning from Experience, <https://assets.publishing.service.gov.uk/media/5a803df4e5274a2e8ab4f03d/ind-16-9-ref-stern-review.pdf>, 2024 年 10 月 10 日。

数量虽然增加，但增量多分布在影响力较低的期刊上，致使澳大利亚的论文引用影响力降低。<sup>①</sup> 还有一些国家，如挪威<sup>②</sup>和波兰<sup>③</sup>在论文产出增加的同时论文引用影响力没有降低。

不同的量化指标显示了不同层面的科研质量的提升情况。Cecchi 的分析表明，科研绩效拨款制度对提升顶级期刊论文数量的激励效果不明显，但会提高篇均被引次数。<sup>④</sup> 这意味着科研绩效拨款制度可以提升高质量论文的平均影响力，但对促进卓越科研成果数量增长的效果有限。

### （三）科研评价政策及拨款制度对科研产出增长的激励效应随时间衰减

科研绩效拨款制度虽然增加了出版物的数量，但是这种影响会随时间推移而降低，并在几年后消失，<sup>⑤</sup> 这表明科研评价活动的激励效果不具有可持续性。例如，捷克于2008年开始实行科研绩效拨款制度，2009年到2011年捷克大学总体科研产出的年增长率分别为65%、30%和12%，该制度促进了科研生产力的提升，但提升速度却显著放缓。<sup>⑥</sup> 在人数保持稳定的情况下，科研机构的生产力不会无限制地提升。捷克在实行科研绩效拨款制度之前，科研生产力相对较低；在实行科研绩效拨款制度之后，因科研评价对象受到激励，科研成果产出量最初会大幅增加，而一旦科研生产力原有的潜能被激发出来，科研成果产出量的提升速度就会逐渐放缓，呈现边际效用递减现象。因此，为保持激励效应，就必须对科研评价政策进行经常性调整，不断挖掘各方面的潜力，这也使科研评价政策的周期性调整成为必然。

① 参见 Jesper W. Schneider, Kaare Aagaard and Carter W. Bloch, What Happens When National Research Funding is Linked to Differentiated Publication Counts? A Comparison of the Australian and Norwegian Publication-based Funding Models, *Research Evaluation*, Vol. 25 (3), 2016, pp. 244 - 256。

② 参见 Jesper W. Schneider, Kaare Aagaard and Carter W. Bloch, What Happens When National Research Funding is Linked to Differentiated Publication Counts? A Comparison of the Australian and Norwegian Publication-based Funding Models, *Research Evaluation*, Vol. 25 (3), 2016, pp. 244 - 256。

③ 参见 Luwel Marc, Performance-based Institutional Research Funding in Flanders, Belgium, *Scholarly Assessment Reports*, Vol. 3 (1), 2021, pp. 1 - 24。

④ 参见 Daniele Checchi, Marco Malgarini and Scipione Sarlo, Do Performance-based Research Funding Systems Affect Research Production and Impact?, *Higher Education Quarterly*, Vol. 73 (1), 2019, pp. 45 - 69。

⑤ 参见 Daniele Checchi, Marco Malgarini and Scipione Sarlo, Do Performance-based Research Funding Systems Affect Research Production and Impact?, *Higher Education Quarterly*, Vol. 73 (1), 2019, pp. 45 - 69。

⑥ 参见 Barbara Good, Niki Vermeulen and Brigitte Tiefenthaler et al., Counting Quality? The Czech Performance-based Research Funding System, *Research Evaluation*, Vol. 24 (2), 2015, pp. 91 - 105。

#### （四）不完善的科研评价体系和拨款制度会带来较大负面影响

相较机构层面的科研评价体系，与科研绩效拨款相关的国家科研评价体系具有更大的影响力，科研评价政策不合理、指标体系简单片面、实施过程有纰漏等，都可能会带来较大的负面影响。有些国家早期在科研评价中采用简单的基于数量指标的计算公式统计科研产出，忽视了其他因素，导致对科研成果数量的片面追求。例如，捷克从 2008 年起采用简单的基于数量指标的评价体系分配全部科研经费。简单追求科研成果数量和高比例的绩效经费对科研机构及科研人员的管理、研究和成果发表行为产生了很大影响，鼓励了机会主义，一些权重很高但相对容易产出的成果数量或容易被操纵的出版物数量增加最快。该政策还对科研机构造成了严重影响，由于绩效拨款占整体科研经费比例太高而各机构每年的绩效得分变化较大，故而科研机构年度经费稳定性极差。2008 年至 2010 年，捷克一些科研机构的拨款份额相对增加了 200% 以上，而另一些科研机构的拨款份额则出现大幅度下降，这使科研机构无法进行中长期规划而只能高度注重短期利益，对国家科研系统发展极为不利。由于上述原因，该政策在实施之后受到广泛而强烈的批评。<sup>①</sup>

#### （五）国家层面的科研评价政策在基层存在不当使用

在国家科研评价体系存在的问题中，有些是容易发现并可以纠正的，还有一些潜在的缺点和问题并不直接显现于政策的实施效果中，而是以较为隐蔽的方式起作用，但可能会破坏国家的学术环境和学术风气，长此以往将产生严重的后果。其中，最值得关注的是国家科研评价政策在基层的不当使用问题。

国家科研评价政策通常适用于宏观和中观层面的评价对象，其采用的量化评价指标不适用于微观层面的评价对象。然而，在现实中基层科研机构很容易将国家层面的科研评价政策直接用于评价个体研究者和具体科研成果，以便提升自身的评价等级，从而获得更多经费支持。例如，一些国家在评价科研机构或学科时，通常按期刊评价结果给在高质量期刊发表的论文以更高的权重或分数，这在宏观和中观层面是合理的，但是如果将之用于评价个体研究者或具体科研成果，就会产生“以刊评文”的不合理现象。很多国家的科研评价政策明确规定国家科研评价的指标或结果不应被用于评价科研人员个体，如意大利的 VQR 评价体系在其评价政策中明确提示“VQR 的结果不能也不应该被用来评价科研人员个体”，<sup>②</sup>但实际上很难阻止基层机构显

<sup>①</sup> 参见 Barbara Good, Niki Vermeulen and Brigitte Tiefenthaler et al., Counting Quality? The Czech Performance-based Research Funding System, *Research Evaluation*, Vol. 24 (2), 2015, pp. 91 - 105.

<sup>②</sup> Andrea Bonaccorsi, Two Decades of Experience in Research Assessment in Italy, *Scholarly Assessment Reports*, Vol. 2 (1), 2020, p. 9.

性地或隐性地将这些指标和结果用于评价个体研究者及用于决定研究者的晋升和考核结果。政府希望通过科研绩效拨款制度引导大学和研究机构向国家期待的方向发展,由于评价方法适用性的限制,国家科研评价体系采用的评价方法不适用于评价个体研究者,但大学和研究机构往往容易简单照搬国家科研评价体系的评价指标,从而带来非预期的负面效果。这个问题普遍存在且难以破解,是科研绩效拨款制度下量化评价体系面临的重大挑战之一。

#### (六) 国家科研评价体系不存在普遍适用的“最优”模式

国家科研评价体系建设是一个复杂的系统工程,评价指标的设计、评价政策的制定和执行、经费分配的模式和激励强度、国家的科研规模和评价文化以及学术界的接受程度都会对科研评价的实施效果产生重要影响。目前,没有任何国家的科研评价实践可被称为“最优”模式,不同国家之间可以相互学习和借鉴,但不能简单复制他国的科研评价模式。在理论研究方面,以同行评议为基础的国家科研评价模式得到较为一致的认可,但在实践方面,坚持采用量化评价模式的国家仍多于采用同行评议模式的国家。英国 REF 系列科研评价活动被世界广泛认可,其中的很多理念和要素被其他国家借鉴和吸收,但采用英国科研评价模式的国家很难取得同英国一样的评价效果。这不仅需要良好的评价文化、庞大的学术同行群体、合理的政策设计和细致的组织实施,而且需要耗费巨大的评价成本和漫长的评价周期。英国 REF 2014 的评价成本高达 2.46 亿英镑,评价周期已从之前的 5 年延长到目前的 8 年。<sup>①</sup>“挪威模式”是基于文献计量学评价指标的较好实践,具有简约、高效的优点,同时也具有量化评价模式的缺点。丹麦、芬兰、比利时等多个国家参照“挪威模式”建立了本国的科研评价体系。2021 年底,丹麦在采用该体系 12 年之后最终放弃了这条路线,其原因既有评价指标体系设计与实施方面的问题,也有评价目标不够明确、国内学术界强烈反对、与利益相关者沟通不足等方面的因素。

未来也不会有普遍适用的国家科研评价体系“最优”模式,各国应当结合本国的评价需求、评价文化、科研规模、评价投入等多方面的情况,在评价效果和评价成本之间有所取舍,探索并完善适合本国情况的科研评价体系。

### 四、对建立完善我国学术评价体系的启示

目前,我国虽然没有明确实行科研绩效拨款制度,但是也存在类似的管理机制。党的二十届三中全会提出,要“深化科技体制改革”,“坚持面向世界科技前沿、面向经济主战场、面向国家重大需求”,“推动科技创新力

<sup>①</sup> 参见刘彦林:《英国社会科学评价模式——以科研卓越框架(REF)为例》,内部研究报告,2023年。

量、要素配置、人才队伍体系化、建制化、协同化”，“提升国家创新体系整体效能”，“完善中央财政科技经费分配和管理使用机制”。<sup>①</sup> 2015年国务院印发的《统筹推进世界一流大学和一流学科建设总体方案》提出，要“创新财政支持方式，更加突出绩效导向，形成激励约束机制”，“建立健全绩效评价机制……在相对稳定支持的基础上，根据相关评估评价结果、资金使用管理等情况，动态调整支持力度”。这些国家重大科研体制改革和战略部署都蕴含着对国家层面科研评价体系的需求。30多年来，多个国家在国家科研评价体系的建设和实践中积累的丰富经验和深入思考对进一步改善和优化我国国家层面学术评价体系、加深我们对科研评价的理解和认识大有裨益。本文建议从以下三个方面对我国国家层面学术评价体系进行完善和优化。

### （一）完善评价要素，规范评价过程

制定合理、系统、操作性强的学术评价政策，该政策既要包括宏观原则，又要有实施细则。对长期性的学术评价要保持评价政策的系统性、连续性，还要根据前期存在的问题对评价政策进行适度的动态调整。当前，我国国家层面学术评价活动的总体政策虽然可以做到及时公开发布，但是存在评价政策不够系统完善、缺乏评价细则以及具体评价政策公布不够及时的问题，应尽可能及时通过公开渠道正式发布相关评价政策，提升评价政策及评价结果的透明度。

国家层面的学术评价体系必须围绕国家政策目标，服务国家发展战略，面向经济社会发展需要，坚持正确的政治方向和学术导向，充分发挥学术评价的“指挥棒”作用。要制定具体而明确的评价目标，紧紧围绕评价目标制定评价标准，选择合理的方法开展评价。例如，我国“双一流”建设提出要坚持以一流为目标、加快走向世界一流。在评价过程中，应明确“世界一流”的定义。不应将大学排行榜作为“一流”的标尺，而应制定更加科学和国际公认的“世界一流”标准。首先，要深入研究“一流大学”和“一流学科”的内涵和范围，确定评价标准；其次，要确定若干用于比较的尺度，即评价指标；最后，“世界一流”是比较的结果，因此，要确定比较对象，收集比较对象的发展情况和相关数据，以便将我国高校和学科放在世界层面进行比较，从而了解我国高校和学科是否已达到世界一流水平和仍需加强的方面。

### （二）重视评价体系的系统性和完整性

国家科研评价体系影响范围广，在评价体系建设过程中，不仅要认真考虑评价政策、评价主体、评价对象、评价标准及指标、评价方法，而且要将评价的外部环境纳入评价体系进行综合考量。

应开展国家科研评价体系外部环境的影响因素研究，详细了解影响我国

<sup>①</sup> 《中共中央关于进一步全面深化改革 推进中国式现代化的决定（二〇二四年七月十八日中国共产党第二十届中央委员会第三次全体会议通过）》，《人民日报》2024年7月22日。

学术评价的重要外部因素，并了解其影响强度及影响机制。应加强与科研评价对象及利益相关者的协商。从各国的实践经验来看，评价活动利益相关者的支持是评价活动顺利开展的必要条件。评价对象及利益相关者对评价的理解是评价活动成败的关键之一。同样的评价政策在不同国家实施，有的较为顺利，有的却遭到强烈的反对。在建设中国特色学术评价体系过程中，必须深入调研，广泛听取大学与研究机构、出版领域、管理部门、评价专家及科研服务对象的意见，明确评价目标，在评价标准、评价方法方面基本达成共识。应提高学术评价体系中评价方法和评价数据的透明度，确保评价机制公平、公开。应加强与国家其他相关制度的统筹协调，综合考虑各种外部因素对学术评价工作的影响，如处理好高校学术评价与教学评价的关系，以及做好研究机构中学术评价与智库评价的协调与结合。在我国，要特别注意保持合理的科研绩效激励力度，避免形成过度激励。例如，地方政府高度重视“双一流”建设，很多省份在“十四五”规划中都将推进“双一流”建设作为重要的工作目标。国家的相关经费支持、地方的各种配套资源和政策倾斜，再加上评估结果对学校声誉的影响，各种有形和无形资源构成巨大的激励和压力，推动高校全力以赴满足甚至迎合“双一流”评估的要求。在这种情况下，非常容易出现营私舞弊、弄虚作假的现象。因此，要根据发展目标和发展阶段确定激励的形式及力度。

### （三）科学开展元评价活动，加强反思与改进

在我国，规范的元评价活动相对较少。为持续改善和优化我国学术评价体系，在每一次重要的学术评价活动结束后均应开展规范的元评价活动，从评价设计者、评价参与者、评价对象及第三方机构等多个角度广泛征求意见，分析评价效果及存在的问题。应按照学术评价的目标审视和检查学术评价活动是否达到目的、是否存在问题或潜在的不良影响，并提出改进意见。

关注科研绩效评价活动对科研生态的影响。开展出版模式的监测及分析，了解科研人员的出版行为及其变化情况，通过文献分析和调查方法对评价活动的效果给予客观评价。对评价对象进行调研和访谈，从科研人员和管理者的角度了解其意见和建议，并了解评价活动对不同年龄段学者的影响，其中应特别关注年轻学者对评价的态度及感受。这些反馈都应作为制定后续学术评价政策的重要参考。

加强对学术评价的反思，充分认识学术评价可能带来的问题，避免过度评价和过度激励，关注国家层面的学术评价政策对基层评价活动的带动和影响，预防和纠正评价方法的不当使用。关注基于评价的激励机制，分析我国各类学术评价体系是否出现边际效用递减的现象。认识评价及激励的有限作用，观察并分析学术评价效果，当激励效应随时间衰减时，就应适度调整学术评价政策。

（责任编辑：任朝旺 李润东）