

澳大利亚学科评估：发展动向与启示

易凡 余小波

(湖南大学教育科学研究院, 湖南省长沙市, 410082)

摘要: 以学科划分评估单元, 澳大利亚学科评估围绕“卓越框架”不断发展, 在评估主体上, 组建了独立、多元、专业的评估主体团队; 在评估内容上, 关注非学术影响力; 在评估方法上, 定性与定量相结合; 在评估流程上, 利用信息技术, 追求评估的透明、公开和规范; 在评估结果的使用上, 以世界标准划分评估等级, 为利益相关者提供信息。在澳大利亚学科苹果体系的新发展动向的基础上, 探讨其对我国学科评估的启示。具体而言, 我国学科评估应建立元评估审查机制, 持续改进评估; 评估内容上应关注非学术影响评估, 突出成果贡献; 评估方法上应结合信息技术、引入数据库信息系统; 评估流程上应增强评估的透明性, 规范评估行为。

关键词: 澳大利亚 ERA; EI 影响力评估; 学科评估

中图分类号: G4 **文献标识码:** A

在创新型国家大背景下, 众多发达国家为促进本国高校科研高质量发展, 建立了以学科为划分基准的科研评价体系, 各个国家学科评估的名称、形式、侧重点不同, 但其根本目的是一致的, 一般被称为 Research evaluation。“学科评估”是我国高等教育情境下的一个术语, 作为高校科研评价的热点和焦点项目, 广受政府、学界和社会关注。作为率先开展高校科研评价体系的国家, 英国和澳大利亚的评估制度一直备受关注。英国的学科评估可追溯至 20 世纪 80 年代, 从 RAE、RAE、再到 REF, 经过 8 次大规模的评估活动, 目前已完成 REF2021。澳大利亚于 21 世纪初宣布面向本国高等教育机构实行科研评价, 从 RQF 方案研制到 ERA 具体实施, 澳大利亚已开展了四轮 ERA 活动。2018 年, 第四轮 ERA 实施的同时, 澳大利亚研究委员会 (Australian Research Council, ARC) 开发了“参与和影响 (Engagement and Impact, EI) 评估体系。该轮评估体系创新性地加入了“科研参与”因素, 研究人员与最终用户的参与情况, 并展示大学如何将其研究转化为经济、社会、环境、文化和其他公共利益。ERA2018 与 EI 评估形成配套评估机制。

目前, 围绕英国历次学科评估的发展历程、组织实施、指标体系、评估流程、评估方法等已有诸多学者进行论述, 而澳大利亚学科评估关注甚少。以“英国 REF”为主题, 期刊来源类别设置为“北大核心”和“CSSCI”, 在中国知网上检索相关文献 49 篇; 而以“澳大利亚 ERA”为主题, 期刊来源类别设置为“北大核心”和“CSSCI”, 仅有 7 篇相关文献。笔者认为, 英国和澳大利亚学科评估体系的不断发展优化, 出现了相似的变革趋势和各具特色的评估体系, 都可为我国学科评估提供宝贵的经验参考: 英国和澳大利亚在注重“跨学科研究”“非学术影响”“环境”等热点难点问题上作出先期探索, 一定程度上代表国际上学科评估发展的新趋势。此外, 澳大利亚利用信息技术和信息系统, 规范评估流程, 减轻评估负担。因此, 本研究澳大利亚学科评估体系的新变化和创新点为基础, 探讨不同学科评估体系与特征对我国学科评估的启示。

一、澳大利亚学科评估体系的发展概述

20 世纪 70 年代-80 年代, 澳大利亚开始探索本国的科研评估体系, 在国际形势的变化

和澳大利亚政府政策的调整背景下，构建出不同的评估项目，呈现出阶段性特征，其发展阶段大致可以划分为：20世纪70年代-80年代：基于科研量的绩效评估；20世纪90年代：基于投入和产出的综合指数评估；21世纪初：基于英国经验的质量评估框架 RQF；2008年-2018年：基于卓越科研的 ERA 阶段；2018至今：基于参与和影响的 ERA+EI 阶段。

	RQF评估	ERA评估	EI评估
评估目的	全面评价科研成果的质量和影响	1.提供给政府、工商业界和其他社会团体更优秀的科研信息； 2.提供以学科为基准的国家层面的盘点； 3.鉴别出卓越表现的科研活动； 4.鉴别新型科研领域和优待深层次发展的领域； 5.在国内和世界范围内比较澳大利亚所有学科范围的科研活动	向澳大利亚政府和公众展示大学科研投入所带来的学术界之外的社会经济利益；进一步促进大学对科研影响转化的支持；明确大学科研参与的支持机制；明确研究机构目前的科研影响发生途径
评估对象	13个学科专业评价小组进行科研评价	22个学科类目划分为8个学科群	各高校2位代码研究领域为评估单元

图 1 澳大利亚评估项目

（一）20世纪70年代-80年代：基于科研量的绩效评估

20世纪70年代-80年代，澳大利亚联邦政府成立了就业、教育与培训部，部长道金森提出了对发放高校科研经费的改革，声明不再增加经常性拨款，改为增加科研资助，在高校之间引入市场竞争制度，无论曾是“大学”或是“高级教育学院”，高校之间都公平竞争财政拨款。“并规定以“国家教育体系”（Unified National System）取代高等教育“双轨制”，进行高校规模、结构、类型的整合。”

在这一阶段，澳大利亚主要基于“科研数量”来对高校科研水平高低进行评估，进而根据评估结果进行科研拨款。当时，澳大利亚联邦政府成立“澳大利亚研究委员会”（Australian Research Commission, ARC），负责科研经费的相关分配事项，逐渐形成了基于各高校的科研数量分配经费的机制。这种完全量化的评估模式使科研成果大幅增加，各高校均争相增加科研产出以获得更多的经费支持，但却忽视了科研成果的质量。

（二）20世纪90年代：基于投入和产出的综合指数评估

20世纪90年代，受矿业等资源产业繁荣的推动，澳大利亚经济呈现快速增长态势。但由于过度依赖传统资源行业而技术开发与创新应用不足但问题日渐突出。澳大利亚政府迫切需要推动国家创新和科学发展，以加强科研创新与技术开发应用^[1]。1995年，澳大利亚以“科研投入、科研产出”的综合指数法进行评估，其中，科研投入主要是高校科研经费占比80%；科研产出看重论文、著作、专利等数量指标。1999年6月，澳大利亚政府在《新知识新机遇》报告中强调，ARC应充分发挥自身促进国家科研水平的战略决策功能，将提高高校科研质量使其为国家创新作出更大贡献的任务放在首位。

在这一阶段，虽然对高校科研的“投入和产出”进行综合评估，但实际上，澳大利亚高校热衷于提高 Web of Science 数据库收录论文数量。研究数据表明，澳大利亚科研产出每年

[1] 刘志民, 李馨儿. 澳大利亚高校科研评价改革动向与启示[J]. 高校教育管理, 2020, 14(05):24-32.

提高 8%，发表 SCI 论文增长 2%^[2]。但从其指标占比来看，澳大利亚发表论文的平均引用影响（Citation Impact）却在下降。可见，这一阶段澳大利亚学科评估仍然是以量化评估为主，科研质量并未得到较大改进。这给澳大利亚政府敲响警钟。

（三）21 世纪初：基于英国经验的质量评估框架 RQF

21 世纪初，相比日渐增长的科研支出，澳大利亚的科研产出与影响力却处于混乱之中。早期基于科研量进行绩效评价和以经费投入、成果数量产出为指标的综合指数评价已不适应国家发展需求。为缓解科研成果质量参差不齐这一局面，2004 年，澳大利亚在英国、新西兰等国家的评估专家和国内学界知名学者的共同商讨下，参照英国评估经验，针对评估方法和程序等内容，先后向各界专家学者进行了几轮公开意见征询，历时三年，建立了评价科研质量及其影响力的“科研质量框架”（Research Quality Framework, RQF），由 13 个学科专业评价小组进行评估，RQF 原计划于 2008 年启动，但最终因政府更迭而未能实施。

这一阶段，RQF 的设计蓝图中，聚焦于科研质量与科研影响力，体现出评估理念的更新与评估标准的提高，关注科研影响力是在更广泛的范畴内对科研成果产生的社会效益进行综合考察，力图使科研成果更广泛应用于社会各界。

（四）2008 年-2018 年：基于卓越科研的 ERA 阶段

2008 年，新上台的陆克文政府引入了侧重评价科研质量的“卓越科研评价”（Excellence in Research for Australia, ERA），旨在鉴别澳大利亚高等教育领先的研究领域与优势学科，以学科为划分基准，盘点、评估全国范围内高等教育机构的优势研究领域和有待发展和加强的研究领域。相比于 RQF，在评价设计多个方面有所改进的 ERA 将会有更加卓越的评价效果，更能挖掘澳大利亚的潜在科研优势，迄今为止，澳大利亚已开展了四轮 ERA 活动：2009 年 ERA 试行，以物理和化学学科群、人文与艺术科学学科群为评估对象；2010 年第一轮 ERA；2012 年第二轮 ERA；2015 年第三轮 ERA；2018 年第四轮 ERA；2023 年将开展第五轮 ERA，已出台 ERA2023 评估框架，评估资料提交指南，评估期刊刊表，评估系统规则等相关文件。

这一阶段，ERA 卓越科研评估制度关注评估工作的综合效应，其实施理念也不断追求卓越，运用更高的标准，更完备的指标体系、更加科学细致的评估维度来体现卓越，例如在评估结果上，以世界标准作为评估标准，通过科学评估引导高校发展，以系统体系确保卓越质量、以质量为核心，满足不同主体需求等。

（五）2018 至今：基于参与和影响的 ERA+EI 阶段

自 10 多年前 ARC 首次发展出 ERA 以来，大学的研究景观和国家大学研究评估的背景已经发生了显著的变化。国际排名计划变得更加突出，政府和公众更加关注研究的影响，以及大学如何与行业、政府和社区利益相关者接触。因此，澳大利亚政府制定了全国性的参与和影响评估（Engagement and Impact assessment, EI assessment），EI 评估于 2018 年进行，与 ERA2018 形成配套评估机制，主要是评估研究成果对社会广泛的贡献，通过评估参与度和影响，鼓励大学与行业等研究最终用户之间加强合作。作为同伴练习，ERA 和 EI 评估研究表现的不同方面，每个方面都有自己的目标。评估组认为，继续使用同伴练习的方法是有价值

[2] 赵勇, 张灵阁, 韩明杰, 李晨英. 澳大利亚科研评价体系的演变、特点与启示[J]. 中国科技论坛, 2015, (12): 149-155.

的，分别评估研究质量、参与性和影响反映了研究的每个方面之间的差异，对澳大利亚的研究实践产生更详细的见解^[3]。可以说，2018年后，ERA和EI评估并行的体系，是澳大利亚学科评估发展的新阶段。

二、澳大利亚学科评估体系的核心内容

（一）评估目的

澳大利亚学科评估制度发展初始，评估目的是政府为了简单进行科研经费的分配；发展到RQF时，其目的是鼓励高质量科研的发展；而ERA的评估目的有5个方面，（1）促进卓越：严格评估研究质量，以促进所有领域和所有类型研究的卓越追求。（2）为决策提供信息：提供有关大学卓越研究和活动的丰富而强大的信息来源，以告知和支持大学、行业、政府和社区利益相关者的需求；（3）展示质量：向政府和公众提供澳大利亚大学研究质量的证据；（4）启用比较：允许澳大利亚大学之间以及所有学科领域的世界标准进行比较^[4]。

2018年，引入EI评估，该评估评估研究人员与最终用户的参与情况，并展示大学如何将其研究转化为经济、社会、环境、文化和其他公共利益。目的在于为促进大学和产业间更为紧密的合作。自2003年以来，澳大利亚学科评估一直致力于为所有与我们互动的人提供独特而专业的教育体验。加强了高等教育部门提供的教育服务，并为学生提供一个平台，让他们追求最适合他们的职业道路。首先，在学科评估目的的根本性功能上，从学科科研本身出发，澳大利亚学科评估最初目的是提高科研质量，鼓励发展高质量科研；其次，在学科评估目的显性功能上，通过评价发现优势学科和有潜力的科研领域，了解澳大利亚高校科研水平在世界上的排名情况；在学科评估目的的隐形功能上，为形成良好的科研文化，鼓励积极的科研行为；再次，在学科评估目的的经济功能上，为确保政府公共资金被投到杰出的科研中，实现资金效益最大化，澳大利亚高校科研评价制度从设置初始就与政府资金分配脱离不了关系；接着，在学科评估目的的社会功能上，促进社会与科技发展，无论是出于其他什么方面的考量，提高高校科研质量、发掘新兴的、有优势的科研领域，还是资金效益的最大化，最终科研评价制度的目的还是要落实到国家建设上，为构建创新而卓越的国家添砖加瓦。

（二）评估单元设置

澳大利亚学科评估针对评估学科的分类引入了“评估单元”（Unit of Evaluation, UoE）概念。澳大利亚学科分类系统中，学科被定义为澳大利亚新西兰统计局发布的2008年澳大利亚和新西兰标准研究分类（ANZSRC）中确定的两位数 and 四位数研究领域（FoRs）代码。ANZSRC是一个有三个级别的分层分类系统。最高水平是22个两位数的研究领域（FoR），其次是157个四位数，然后是广泛的六位数的代码^[5]。例如，13代表教育学学科大类，1301教育制度学。各机构以4位数代码的形式提交初始材料，这些初始数据资料汇总并划分位2位或4位数字代码的评价单元（Unit of Evaluations），

[3] ERA.EI.Review.Final Report 2020-2021[EB/OL].https://www.arc.gov.au/sites/default/files/era_ei_ac_report.pdf

[4] Excellence in Research for Australia[EB/OL].<https://www.arc.gov.au/evaluating-research/excellence-research-australia>

[5] 2位数字代码代表各个学科大类-4位数字代码代表学科大类下设置的具体学科类目-6位数字代码表示更加详细的学科类别

在具体评估中，澳大利亚 ERA 评估也是将 22 个学科类目依据学科划分出 8 个学科群，针对每一个学科群分设 8 个对应的评价委员会^[6]，对 22 个评估单元或 157 个评估单元进行评估的过程。它们都以学科分类为立足点，运用专业的评价主体对相应的科研领域进行评价。按学科分类评价主体、评价对象使得评价工作不仅具有针对性，还在一定程度上体现了专业性，确保了评价结果的公正性和可信度。高校可以选择自己较有优势的科研领域去参加评价，了解自己在所在学科领域处于什么地位，加深对所在科研领域的了解。而且，对于不同的学科也将采用不同的评价方法，确保了评价结果的合理性。

（三）评估主体

评估主体是实施评估活动中必不可少的一环。独立客观的评估主体是评估活动系统性和科学性的前提。笔者认为，评估主体包括两个方面。其一是评估机构，涉及到学科评估制度的管理、改进和监督环节。澳大利亚负责执行学科评估相关工作事宜的是澳大利亚科研委员会（Australian Research Council，简称 ARC）；负责对全程评估工作进行审查与复核对机构是评估审查委员会（Scrutiny Committee，简称 SC），此外，还设立了科研评估委员会（Research Evaluation Committee，简称 REC）协助相关评审工作。其二，评估主体更是评估实践中真正进入评估现场的个人或群体。前文所述，澳大利亚学科评估是以评估单元划分评估小组，进行专家评审的过程。除制定相关评估制度的专家组和执行具体评估工作的专家组外，澳大利亚学科评估同行评议专家小组，以更好地确保评价结果的准确性与可信度。此外，为使评估专家构成在性别、年龄、种族、宗教、区域等方面的分布更趋于合理。澳大利亚评审专家名单及其简历都会在 ARC 的官方网站上向媒体和社会公布，如 ERA 2018 研究评估委员会由来自澳大利亚和海外的 150 名杰出和国际公认的研究人员组成，他们在各自领域和研究评估方面具有专业知识。

综上所述，澳大利亚学科评估评估机构带有半官方性质，是政府和高校之间的“缓冲带”。在评估过程中，政府、社会和学校均有话语权和参与权，遴选专家时需要注重专家的代表性和多样性，除国内外知名学者外还聘请一些工业、商业代表，根据科研成就、知识深度、学术地位、沟通能力与合作精神等遴选评审专家，有效地保证了评估结果的客观性。

（四）评估指标体系

最新的澳大利亚学科评估体系实际上是 ERA 与 EI 两项评估系统并行的，在 ERA 中，包含研究质量指标和研究活动指标。其中，研究质量是根据引文分析或 ERA 同行评审来考虑的。研究活动指标为评估 UoE 提供了背景。在符合条件的研究人员的背景下，研究活动是根据研究成果和研究收入来考虑的^[7]。EI 评估体系则包括“参与”和“影响”两个维度。关于“参与”指标，在文字材料上，高校需要提交评估单元（学科）参与活动的整体情况，包括该学科的发展战略及目标，在数据材料上，高校需提交相关定量数据，主要涉及该评估单元与学术界外部的研究用户提供的研究经费和研究生数量。关于“影响”指标，高校需提交主要包含“影响叙述”“影响方式叙述”的影响研究报告。在“影响叙述”指标上，

[6] 八个评估研究委员会分别是生物和生物技术科学（BB）、人文与创意艺术（HCA）、经济和商业（EC）、教育与人类社会（EHS）、工程和环境科学（EE）、数学、信息和计算科学（MIC）、医学和健康科学（MHS）、物理、化学和地球科学（PCE）

[7] ERA 2023.Submission Guidelines[EB/OL].https://www.arc.gov.au/sites/default/files/2022-07/era_2023_submission_guidelines_0.pdf

需提交该项科研成果的影响对象、影响程度报告以及提供相关证明。在“影响方式叙述”指标上，需详细阐明在该项科研成果产生“影响”的过程中，高校制定的学科发展战略与对该项科研成果的人力、物力、财力支持机制^[8]。

结合表 2，发现澳大利亚学科评估指标体系侧重于学科建设的影响，内部参与和外部环境的作用。一方面，关注到高校科研活动对学术界之外的影响和贡献；另一方面，全面掌握高校自身的科研制度建设和实施情况，致力于建立多维度的科研成果影响力指标，实现以评促建，最终推动高校与各产业界协同创新发展。

表 1 澳大利亚学科评估指标体系

	评估指标	评估内容	评估方式
澳大利亚 ERA2023	研究质量	根据出版简介、引用分析、ERA 同行评审和同行评审澳大利亚和国际研究收入进行审议	同行业专家评议和计量评估相结合
	研究活动	根据研究产出、研究收入和其他研究项目在合格研究人员概况范围内进行审议	案例提交，定性分析和定量评估相结合，同时结合专家小组、成果应用方以及第三方评价机构的共同审核
澳大利亚 EI2018	参与	参与叙述：科研活动的整体概况 参与指标：研究用户的现金支持数量；每名全职员工接受的 HERDC 研究收入；HERDC 第一类基金中由研究用户资助的项目占全部基金的收入数量、项目数量比例；研究商业化收入；与研究用户联合培养的研究生数量	定量指标+定性解释性声明
	影响	影响叙述：影响（包含关联研究 15 年）： 影响方式叙述：影响方式	

（五）评估流程

如图 2 所示，澳大利亚学科评估分为四个阶段，每一个阶段又由一个或几个小阶段组成。第一阶段，评估材料的提交。参评高校准备评估材料，按 ERA 技术包中的数据格式要求填至存储库中，将材料以规定格式上传至评估系统（System to Evaluate the Excellence of Research, SEER），ARC 将通过评估系统核验评估材料的完整性和一致性，最终参评高校校长将以认证形式提交确认最后的评估材料。第二阶段，评估材料的分配。评估小组 REC 将较难设定标准的评估单元分配给同行评议成员，将有详细标准的评估单元分配给 REC 成员。第三阶段，评估阶段，同行评议小组和 8 个 REC 小组成员将独自进行初步评估，得出初步结果后成员之间互相交换意见，最后复审评估结果，开会商议确定最终评估结果。第四阶段，评估报告的反馈，将由 ARC 结合同行评议报告和 REC 报告将评估结果写进国家报告中并发布。

[8] EI 2018 FRAMEWORK[EB/OL].https://www.arc.gov.au/sites/default/files/2022-07/era_ei_2018_framework_0.pdf

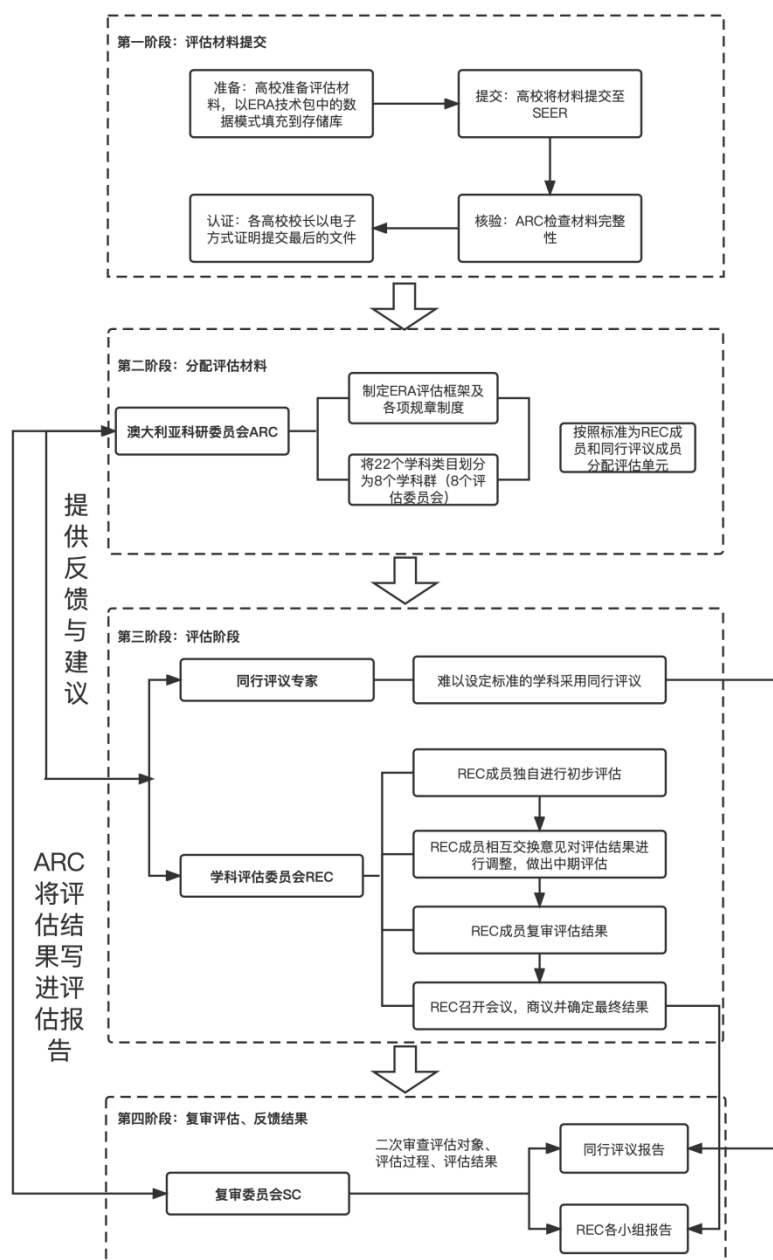


图 1 ERA 评估流程图

从澳大利亚学科评估的流程发现：每一个流程小节之下又有一个到几个不等的小阶段，各个阶段有不同的评价工作需要完成，但它们又相互协调、配合，保证了 ERA 评价工作井然有序的进行。整个程序都充分地体现了评估的客观性和公开性。在评估开始前，评估机构就会在网上公布各个高校学科所要提交的资料，ARC 会发布《ERA 提交指南》、《SEER 技术规范》等政策文件指导高校提交科研评价所需材料；REC 主席和成员的相关信息会公布到网站上，供公众监督，确保评价主体的资格性与专业性。同时，每个学科的评估委员会也会公布学科评估的方法和标准。评估结束后公布各个高校参评学科的得分情况和提交的参评资料。每个提交的每个部分都由具有适当专业知识的小组成员和/或评估员进行了审查，体现出严格的规范性。

（六）评估结果的使用

在结果的呈现上，澳大利亚 ERA 均采用五个评价指标等级进行评价指标，最高等级为“表现优异且远远高于世界标准”。以世界标准来评估，澳大利亚更加明确的将学科评估放在世界的大范围内，注重国际认可程度。

在结果的使用上，澳大利亚 ERA 评估结果报告也反映了每所大学每个学科的表现。它提供了澳大利亚研究的观点，从量子物理学到文学。它强调了我们在至关重要领域所有学科的研究优势。政府、大学和其他利益相关者使用 ERA 数据和结果。例如，为学生、行业和其他利益相关者提供基于学科的信息，许多大学将 ERA 结果纳入其年度报告和战略计划。大学利用 ERA 成果来促进他们在国内外的研究优势^[9]。澳大利亚 EI 评价结果报告则体现了大学对学术界以外产生的影响，为澳大利亚高校、政府和产业界之间多元合作提供了丰富信息。

三、澳大利亚学科评估发展趋势与启示

（一）建立元评估审查机制，持续改进评估

随着知识生产模式的现代转型，学科评估作为评估知识生产活动的根本制度必须同步发展变革，需遵循知识本身生产逻辑，才能发挥评估指挥棒和风向标的作用。在元评估审查方面，英国在 2014 年就根据相关的元评估报告与咨询论证结果对首轮科研评估进行了审查，在此基础上对下一轮的评估细则进行了修订与完善，以期突出强调公平、平等、透明的原则，提高科研评估体系与知识生产创新活动之间的兼容性。英国和澳大利亚于 2020 年正式启动最新一轮元评估^[10]。双方都通过问卷调查，评估各自评估项目的目标达成情况、评价过程、评价指标、评价方法和评价影响。作为补充，还通过访谈形式深度了解科研人员的个性化情况。澳大利亚研究委员会（ARC）于 2021 年完成了对澳大利亚卓越研究（ERA）和参与和影响评估（EI）的全面审查。审查使 ARC 能够：简化和简化其程序利用技术和大数据的最新发展，确保这些计划继续反映世界最佳实践，以及满足以下持续的需求：大学部门，政府和，公众对澳大利亚大学的研究质量、影响和参与度进行有力评估。ERA EI 审查咨询委员会的 ERA EI 审查最终报告向 ARC 提出了 22 项建议^[11]。尽管 ERA 和 EI 在许多方面都处于世界领先，但持续改进确保评估过程保持稳定和有效，能为所有学科和大学提供一个公平的竞争环境^[12]。

斯塔佛尔比姆认为元评估是描述、获取和应用有关评估的效用、可行性、正确性和准确性及其系统性、能力、诚信、尊重和社会责任的描述性信息和判断信息的过程，以指导评估并公开报告其优缺点。元评估符合公共、专业和机构利益，以确保评价提供合理的结果和结论；评估实践继续改进；并确保机构管理高效、有效的评价系统^[13]。从英国和澳大利亚学科评估实践经验来看，实施有效的元评估是应答社会问责和改进评估实践的国际经验和现实

[9] Excellence in Research for Australia[EB/OL].<https://www.arc.gov.au/evaluating-research/excellence-research-australia#what-do-the-era-outcomes-tell-us>

[10] 杨佳乐. 如何开展科研元评价：来自英澳的启示[J]. 科技管理研究, 2022, 42(17): 31-37.

[11] ERA EI Review[EB/OL].<https://www.arc.gov.au/evaluating-research/era-ei-review>

[12] ERA.EI.Review.Final Report 2020-2021.[EB/OL].https://www.arc.gov.au/sites/default/files/era_ei_ac_report.pdf

[13] Daniel L Stufflebeam. The metaevaluation imperative[J]. American Journal of Evaluation, 2001, 22(2): 183-209.

选择。只有通过元评估及时审查和反思,才能使学科评估质量得到公众认可,评估中的问题被理性对待,并成为下一轮评估创新的起点^[14]。在我国第五轮学科评估实践中,组织学科利益相关者交流,通过与多所高校开展交流座谈会,征求对评估的意见和建议。这一举措已体现出我们对元评估的初步探索,但参照国际评估的流程和标准发现,我国学科评估方案中缺少元评估环节和评估结果使用的指导说明。

我们需要认识到元评估的重要性和积极意义:在评估中的形成性元评估,帮助评估人员规划、进行、改进、解释和报告他们的评估研究。在评估后进行的总结性元评估有助于受众了解评估的优势和劣势,并判断其优点和价值。要将形成性元评估和总结性元评估相结合。首先,需要建立起专业的元评估小组,可参考澳大利亚审查小组的构成,邀请来自大学人员、评审小组成员、专家等针对对学科评估后学科的整改效果、学科评估的质量和学科评估产生的影响进行系统的调查和分析,形成学科评估反思和改进的元评估质量报告。其次,在评估过程中,可参考英国 REF2021 中建立投诉与调查机制,提供一个独立的审查过程来保证学科评估的公平公正,有利于获得基于证据的专业化改进建议。最后,评估后要促进元评估结果的交流和使用。评估结束后,深入分析各高校各学科的变化将能促进评估结果的充分交流和合理使用,而非停留在数字化的表象。

(二) 关注非学术影响评估, 突出成果贡献

在当前破除“五唯”,强调大学社会贡献的大背景下,思考大学科研的非学术影响有助于我们重新审视大学科研的目的与价值,有助于我们进一步完善学术评估机制,有助于科研工作者重新承担起作为一名知识分子所应承担的公共责任。英国 REF2021 较 REF2014 将“影响”这一指标比重提升至 25%,澳大利亚也在 2018 年引入 EI 参与和影响力评估。英国和澳大利亚关注学术以外影响的科研评价改革,关注高校科研人员与外部环境都合作互动以及科研的社会影响,建立了精细化、立体化的“非学术影响”指标体系,一定程度上代表了国际学科评估发展的新趋势。在空间维度上,“社会”的外延可以从当地社区、地区、国家、国际甚至到人类社会的意义。在时间维度上,“影响”的时间界定不仅是完成时态,还可以是未来时态,即包括已经产生和预测未来会产生的影响,可以是短期的影响,也可以是中期、长期或长远的影响。在具体内涵上,“非学术影响”指科研在学术共同体之外的所有影响,包括科研对“社会”“经济”“政治”“文化”等产生的影响。在我国二三轮学科评估中,学科社会服务情况在学科简介里只是一笔带过,似乎只是个象征性的存在,到第四轮学科评估时,社会服务贡献作为一个独立的二级指标,对其考量也变为用事实说话式的经典案例。

英澳经验对我国学科评估的启示在于:一是要界定清楚非学术影响的范畴。在“影响力”范围界定方面,英国将科研成果对教师、学生以及教育教学活动产生的辐射影响纳入了评估范围。结合本土实际,在评估指标设计上既注重科研成果对社会生产、生活产生的积极作用,也考虑对高校自身建设与发展所做的贡献。二是要完善“影响案例”+模版的评估方法。目前我国高校往往通过反复打磨修饰,形成的“模版”很难具有“可视、可量、可听、可用”的特点。评价维度有社会影响定性描述、社会影响指标量化、决策咨询产生的效益、利益相关者影响力调查、案例影响力等。必要时可开发影响力发现工具,发挥大数据在评价中数据提取、数据挖掘和可视化表达作用。三是要延长非学术影响评价周期。非学术影响评

[14] 赵立莹. 问责与改进:我国第四轮学科评估元评估[J]. 学位与研究生教育, 2018, (02): 32-38.

价往往不能在短期内呈现出来，一般需要一定的转化周期，特别是对经济、生态的科研成果形式呈现需要几年或者更长时间，在案例提交规定方面，可允许在评估周期内产生辐射影响的研究重复提交，以解决影响力“滞后性”问题。

（三）引入数据库信息系统，减少评估负担

澳大利亚学科评估的一个主要战略是更有效地利用外部数据收集和自动化技术进行评估，以信息科学方法和技术为评估保障，重视评估全过程的管理。以澳大利亚评价系统平台 SEER (System to Evaluate the Excellence of Research) 为例，澳大利亚 ERA 评估中的 SEER 平台是一个存储数据的平台，存储每一轮评价过程中产生的数据，为澳大利亚政府、商业界、其他行业提供高校科研较为全面的信息。更重要的是，SEER 是一个数据处理平台。高校将数据上传成功后，SEER 平台能够对提交材料进行验证，查询科研成果是否被重复提交；保存有效数据并将它划分为评价单元，分配给同行评议小组以及科研评价委员会；生成、预审数字化报告，存入科研机构资料库，最后对提交材料进行总结和验证，SEER 会验证数据格式，查询科研成果是否被重复提交等。

从澳大利亚的实践来看，通过建立学科信息数据库的方式有效缓解了学科评估给高校信息收集带来的压力；通过评估系统自动化应用，节省了评估过程中的人力、物力，且避免了一些人为疏漏问题，提高了评价及数据管理的效率。在我国学科评估过程中，高校填报和提交数据往往是在行政权威介入下开展，造成了系列连锁反应：一是高校主动提供的学科数据不全面。当行政权威介入数据获取分析和评价的整个过程时，许多高校迫不得已在参与学科评估的同时兼顾其他名目繁多的科研评价，未能及时主动提供较为全面的数据，甚至一些高校仅能够提供简单的数据咨询服务；二是报送材料避“问题”重“成就”。在行政压力下，有的高校还可能在申报材料时对学科存在的问题和困难避而不谈，仅提及该学科建设发展方面的成就，这也直接影响了评估结果的真实性和准确性。因此，借鉴国际经验，在我国全国学科大数据服务平台的基础上，也可尝试涵盖各高校近四年的各项学科科研基本信息，以应对下一轮学科评估的需要，从而缓解高校和学科的数据填报压力，帮助高校和学科更好地投入到主观材料的填报工作之中。此外，目前我国学科数据平台提供学科分析服务和信息支撑服务的力度还有待加大。在教育部学位中心官网上仅能看见数据库简介和入口，但是如何登陆如何使用缺少说明。

（四）增强评估透明和公开，规范评估行为

评价程序的客观性是得出公平的评价结果的前提，严格的把控评价程序中的每一个环节，为得出公平的评价结果奠定基础。各阶段信息公开的透明与及时情况是评价制度规范程度的重要表现。如澳大利亚十分强调学科评估的信息公开，在评估程序、标准、指标体系、大纲等全面公开，接受社会监督。

当前我国教育部学科评估中，前四轮各参评高校的评估结果在其教育部学位中心官网上已公布，但关于各项指标的数据收集、数据分析处理过程以及评估专家信息等仍然处于信息封闭状态。评估环节的透明度和评估信息的公开度决定了评估结果的科学性和社会公众的认可度。因此，我国在增强评估透明和公开，规范评估行为方面还需进一步加强。具体而言：一是在评估前期阶段，关于学科评估方案的设计原则、方案初稿、高校、学科行业协会、研究用户代表等主要利益相关者的书面意见，评估试点结果报告等各阶段相关信息等需要及时在教育部学科评估官网上公开，以便公众监督。二是在评估实施阶段，我们需要扩大评估文

件的公开范围,如全部学科评估通用方案、各学科差异性提交材料的规定以及分学科专家评估详细规则的文件等。三是在评估结果公开阶段,在公开各高校学科评级结果的基础上,我们需进一步发布各学科优秀的科研影响案例,展现高校科研的社会贡献,增进社会公众的理解,满足社会公众的信息需求。只有增强评估各环节信息公开的及时性与透明度,我国才能进一步提升高校学科评估的公信力和结果的可信度,更好地发挥学科评估体系服务社会的作用。

四、结语

在知识生产模式转型、政府对科研投入的社会价值追求产生以及公众对科研信任水平有待提升的现实需要下,澳大利亚为推进“创新型国家”发展战略,促进国家科研竞争实力提升和科研资源使用效率提高,不断创新学科评估体系。以学科分类为立足点划分评估单元;以评估单元划分评估专家小组;注重遴选专家的多元性构成;关注非学术影响评估和良好科研环境的构建;引入信息系统减轻评估负担;公开评估所有阶段的信息等等,ERA、EI2018中的诸多改进和创新渐有发展成一种全球趋势之可能,为我国学科评估发展提供参考和借鉴。

当然,澳大利亚 ERA、EI 体系仍存在一些争议和质疑。因此,正如澳大利亚借鉴英国卓越框架时,结合了本国的科研文化特征,我们在借鉴澳大利亚学科评估时需要理性审视,不能是简单的拿来主义,要在本土情境中完善中国特色化的学科评估体系。

参考文献

- [1].刘志民,李馨儿.澳大利亚高校科研评价改革动向与启示[J].高校教育管理,2020,14(05):24-32.
- [2].赵勇,张灵阁,韩明杰,李晨英.澳大利亚科研评价体系的演变、特点与启示[J].中国科技论坛,2015,(12):149-155.
- [3].ERA.EI.Review.Final Report 2020-2021[EB/OL].https://www.arc.gov.au/sites/default/files/era_ei_ac_report.pdf
- [4].Excellence in Research for Australia[EB/OL].<https://www.arc.gov.au/evaluating-research/excellence-research-australia>
- [5].ERA 2023.Submission Guidelines[EB/OL].https://www.arc.gov.au/sites/default/files/2022-07/era_2023_submission_guidelines_0.pdf
- [6].EI 2018 FRAMEWORK[EB/OL].https://www.arc.gov.au/sites/default/files/2022-07/era_ei_2018_framework_0.pdf
- [7].Excellence in Research for Australia[EB/OL].<https://www.arc.gov.au/evaluating-research/excellence-research-australia#what-do-the-era-outcomes-tell-us>
- [8].ERA EI Review[EB/OL].<https://www.arc.gov.au/evaluating-research/era-ei-review>
- [9].ERA.EI.Review.Final Report 2020-2021.[EB/OL].https://www.arc.gov.au/sites/default/files/era_ei_ac_report.pdf
- [10].Daniel L Stufflebeam. The metaevaluation imperative[J]. American Journal of Evaluation, 2001, 22(2): 183-209.
- [11].杨佳乐.如何开展科研元评价:来自英澳的启示[J].科技管理研究,2022,42(17):31-37.
- [12].赵立莹.问责与改进:我国第四轮学科评估元评估[J].学位与研究生教育,2018,(02):32-38.
- [13].Impact case study database now available[EB/OL].<https://ref.ac.uk/guidance-and-criteria-on-submissions/news/impact-case-study-database-now-available/>

Australian subject evaluation: development trend and enlightenment

Yi Fan, Yu Xiaobo

(Hunan University Research Institute of Education Science, Changsha / Hunan, 410082)

Abstract : According to the division of assessment units by disciplines, Australian discipline assessment has been developing around the "framework of excellence", and an independent, diversified and professional assessment subject team has been established in terms of assessment subjects; In terms of evaluation content, pay attention to non-academic influence; In terms of evaluation methods, qualitative and quantitative methods are combined; In the evaluation process, information technology is used to pursue transparency, openness and standardization of evaluation; In the use of the evaluation results, the evaluation level is divided according to the world standard to provide information for stakeholders. On the basis of the new development trend of the Australian discipline apple system, this paper discusses its implications for the discipline evaluation in China. Specifically, China's discipline evaluation should establish a meta-assessment review mechanism to continuously improve the evaluation; The evaluation content should focus on the non-academic impact assessment and highlight the contribution of achievements; The evaluation method should combine information technology and introduce database information system; The evaluation process should enhance the transparency of the evaluation and standardize the evaluation behavior.

Keywords: Australia ERA; EI impact assessment; Subject evaluation

作者简介 (可选) :易凡, 女, 湖南大学教育科学研究院硕士研究生; *通讯作者: 余小波, 男, 湖南大学教育科学研究院教授、博士生导师。